

Pricing Coastal Amenities and Flood Hazards

Zhenshan Chen Assistant Professor, Department of Agricultural and Applied Economics, Virginia Tech, Blacksburg; zhenshanchen@vt.edu

Charles A. Towe Associate Professor, Department of Agricultural and Resource Economics, University of Connecticut, Storrs; charles.towe@uconn.edu

ABSTRACT *Coastlines bring precious amenities to residents, but they also have significant and growing risks from storms and rising sea levels. Disentangling these amenities and risks is of ever-increasing importance to the public and policy makers. Evaluating the most tangible aspect of coastal living in New England, we present a scalable way to construct parcel-specific measures of coastal amenities. Estimates suggest that good coastal amenities, represented by unobstructed ocean views, add about 33%–69% to property prices. Coastal hazard effects are separated from the amenities: the flood zone discounts in coastal Connecticut and Massachusetts are generally low. (JEL Q51, Q54)*

1. Introduction

Numerous studies have focused on the effect of climate-change-related risks on coastal housing markets, but less effort has been made to understand the other side of the story. Considerably more coastal homes are expected to be subject to frequent flooding in the near future,¹ so the supply of enjoyable coastal amenities will decrease. Has the market been responding to this? In a coastal housing market study, disentangling and valuing the hazards and amenities require proper data on both, rather than relying on approximate measures derived from coastal proximity or elevation (Bin et al. 2008; Beltrán, Maddison, and Elliott 2018) because hazards and coastal amenities

are highly correlated with coastal proximity and each other. Without a clear knowledge of the confounding coastal amenities, one can hardly claim validity in the hedonic pricing of a certain coastal hazard (or the related risk). Therefore, the importance of identifying the value of coastal amenities lies not only in the enormous amount of associated dollar value but also in the urgent need to understand the market perception of and responses to coastal hazards and related policies.

Aggregating a comprehensive set of coastal amenities, including ocean view, waterfront status, coastal proximity, land cover, and many others, this article seeks to estimate the hedonic value of coastal amenities and disamenities for coastal properties in Connecticut (CT) and Massachusetts (MA). As the northeastern United States holds one of the richest coastal populations² and will experience a drastic increase in coastal flood risk (Gori et al. 2022), the importance of valuing coastal amenities is particularly evident in this study.

The lack of studies on coastal amenity values is mainly caused by the lack of proper data. Ocean view is a major part of coastal amenities, and efforts have been made to get the ocean view variables. Benson et al. (1998) carried out the onerous task of measuring the ocean view of about 5,000 properties via personal inspections in Bellingham, Washington, and found that high-quality ocean view added property value by 58.9%. As high-quality geographic information system (GIS) technology and data have become increasingly accessible, Bin et al. (2008) used the viewshed analysis tool to calculate the ocean view of about 1,000 properties in New Hanover County, North

¹Because of rising sea levels, 36% more people around the world will be subject to annual flood events by 2050, according to Kulp and Strauss (2019).

²The wealthiest concentrations of high-income households are in the Connecticut towns of Bridgeport, Stamford, and Norwalk, outside New York City, according to Bee (2013).

Carolina. Other hedonic studies using viewshed analysis (not necessarily for ocean view) include Paterson and Boyle (2002), Sander and Manson (2007), and Schmitz (2008), and they unambiguously suggest that favorable views significantly increase property prices. Blessed by more advanced GIS tools and large-scale remote sensing and property data, this article builds an improved and scalable ocean viewshed analysis procedure. As a result, we compute and value ocean views for more than 55,000 coastal properties in CT and more than 332,000 in MA.

We focus on the special flood hazard area (SFHA), or flood zone, status as the coastal hazard indicator.³ We attempt to disentangle the price effects of coastal amenities and flood zone status. A meta-analysis study shows that based on peer-reviewed hedonic studies, coastal properties located in the designated flood zone “command higher prices” (Beltrán, Maddison, and Elliott 2018),⁴ which is hard to interpret and might leave an impression that coastal hedonic studies are generally unreliable. The commonly perceived issue is that the missing coastal amenities are closely correlated with the key flood risk indicator (e.g., Bin et al. 2008), confounding the hedonic estimates. With the coastal amenities gathered in this study, we are able to show the hedonic prices on flood zone status in coastal CT and MA. In assigning the flood zone indicators,⁵ we refine the typical procedure by involving multiple waves of major map releases,

individual map amendments,⁶ and building footprints.⁷

Our property and transaction data mainly come from Zillow’s Transaction and Assessment Database (ZTRAX). Going through a careful data-refinement process (as detailed in Section 2 and [Appendixes](#)), the dataset for final analysis includes single-family residence sales in the coastal towns of CT and MA from 1998 to 2020. ZTRAX data provide at least three significant benefits for this study. First, as a dataset that is accessible free of charge, it enables researchers with limited resources (like ourselves) to work on property-level research questions with a great degree of freedom in the choice of study region and time. Second, without the large-scale and comprehensive information on property and transactions from ZTRAX, it would be immensely difficult for us to construct the final dataset covering a major part of the New England coastline. Third, not typically seen in similar property or transaction datasets, the mortgage information in ZTRAX makes it possible to differentiate the transactions with a mortgage loan and those without, which is important to consider in the context of flood insurance policies.⁸

To mitigate the specification error and improve covariate overlap, we use a quasi-experimental research design. The first step of our identification strategy is a matching algorithm that pairs houses in the flood zone

³Throughout this article, consistent with FEMA definitions, SFHA denotes high-risk flood zones (100-year floodplains, including zones A, AO, AH, A1-30, AE, A99, AR, AR/A1-30, AR/AE, AR/AO, AR/AH, AR/A, VO, V1-30, VE, and V). The flood zone will be the synonym of SFHA, unless suggested otherwise. See more details in the Flood Disaster Protection Act of 1973, available at <https://www.fdic.gov/regulations/laws/rules/6000-2400.html>.

⁴After filtering out publication bias, the authors suggest the price effect of flood risk in coastal regions is about 36.3%.

⁵Errors in flood zone assignment will lead to attenuation bias (i.e., bias the flood risk discount estimates toward zero and inflate standard errors). The typical procedure to assign flood zone status is taking one cross-section of flood map and intersecting the flood zones with address points, which does not consider multiple waves of major revision, individual map amendments, nor the building footprints.

⁶“Individual map amendment” refers to the cases where one homeowner applies for a map amendment to remove the house from the flood zone and is approved. The homeowner gets a letter indicating the amendment (letter of map amendment), but the flood map is not usually revised to present a “hole” reflecting this change for a separated address. Instead, the letter is placed in the public record (together with flood maps) for insurance or other usage. One caveat here is that historical letters of amendment are not always displayed in the public repository. Knowledge from the practitioners helped a lot in understanding this.

⁷In practice, the flood zone status is assigned based on building footprints instead of address points (using address points tend to be the common practice in previous flood zone studies).

⁸Property owners (or buyers) in SFHAs are required to purchase flood insurance up front if they borrow money from federally regulated lenders (for more details, see Di-Vincenti 2006), and this mandatory purchase establishes a big discrepancy in the insurance penetration rate between owners holding a mortgage and those without one.

with similar houses outside. In an attempt to mitigate leftover covariate discrepancies postmatching, the second step estimates a hedonic regression model on the matched sample with fixed effects capturing unobservable census tract specifics, tract-level shocks over time, and overall temporal shocks. To provide support for the identification strategy, we test the robustness of the estimates to different data refinement processes and specifications. Regarding the mandatory flood insurance take-up for with-loan buyers, we require exact matching on mortgage loan status.

The results suggest that ocean view and being on the waterfront significantly increase coastal housing prices. Specifically, we find that a good set of coastal amenities (represented by unobstructed and wide ocean view) add about 33.38%–69.05% to the coastal property prices on the coasts of CT and MA. On the other hand, the flood zone significantly lowers coastal housing prices in two of the six coastal regions (about –3% in New Haven County, CT, and –7% in Bristol County, MA). The average housing price discounts vary considerably in different counties, and we find price premiums associated with flood zone in Dukes and Barnstable Counties, MA.

This study contributes to the literature in the following aspects. First, it entails a data check and refinement process based on ZTRAX data,⁹ which results in a final dataset significantly improved from typical datasets analyzed in coastal hedonic studies.¹⁰ Second, by calculating ocean views for more than 387,000 coastal properties in the northeastern United States, it is the first study that analyzes the value of ocean amenities on a large scale. More important, it provides a scalable procedure to calculate coastal property viewshed (i.e., not necessarily just ocean view), which

could be conveniently extended to the other coasts in the United States given national-level property data like ZTRAX and proper light detection and ranging (LiDAR) layers. Finally, this study could also be seen as the beginning of reassessing coastal hazard effects across the United States, with the help of improved coastal amenity measures.

2. Data and Variables

To consistently estimate the hedonic values of ocean amenities and the flood zone, we spent a tremendous amount of effort building a high-quality dataset. The data consist of two major categories: the property and transaction data, and the GIS data. More details of data sources are found in [Appendix Table A1](#).

Property Data

The property data are acquired from ZTRAX. We gathered tax assessment records and deed records (including mortgage, foreclosure, and transaction records) spanning 23 years (1998–2020) for single-family residences in the focal area, 21 coastal towns in CT, and 66 in MA (shown in Figure 1).¹¹ Recognizing the heterogeneity in these housing markets, we define county groups (CTCGs or MACGs) from core-based statistical areas (CBSAs) and conducted analyses in each county group.¹² To ensure the quality of the sales information (mainly prices and mortgage status), we gathered sales and mortgage data from

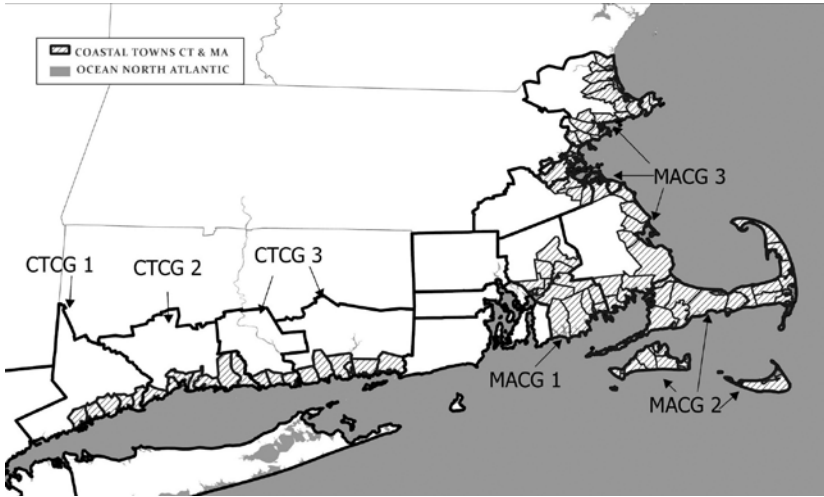
¹¹From west to east, the coastal towns are Norwalk, Westport, Fairfield, Bridgeport, Stratford, Milford, West Haven, New Haven, East Haven, Branford, Guilford, Madison, Clinton, Westbrook, Old Saybrook, Old Lyme, East Lyme, Waterford, New London, Groton, Stonington. The three coastal towns to the very west (Greenwich, Stamford, Darien) are not included because of data limitations and their distinctions from the rest (i.e., having extremely high socioeconomic status).

¹²The county group defined here is to ensure the matched pairs (to be illustrated in the matching methodology) are generally in the same “market.” The definition is based on CBSAs, while merging some counties across CBSAs when observations in one CBSA is very limited and the counties are close to each other: Middlesex is merged with New London into one county group, and Dukes, Nantucket, and Barnstable are merged, too.

⁹The process includes calibration of the coordinates, purging non-arm’s-length transactions, checking transaction prices and mortgage status against administrative data, assigning flood zones based on building footprints, and revising the flood zone status based on major and individual map amendments.

¹⁰For short, compared with this dataset, a typical data set has no viewshed variables, contains rough indicators of flood zone status (as explained in n. 5), does not have a price check process against closing documents, and does not involve an as-detailed process revising the address points to building centroids.

Figure 1
Study Area: Coastal Towns in Connecticut and Massachusetts



administrative data repositories as quality benchmarks,¹³ checked ZTRAX transaction records, and revised if necessary (see [Appendix B](#) for details). We developed a process to purge non-arm's-length transactions (i.e., with prices that do not reflect fair market values), described in detail in [Appendix C](#). To avoid influences from house flippers, we dropped transactions for the same property that are within 120 days of each other. The transaction prices are deflated to the 2017 dollar value according to the Consumer Price Index from the U.S. Bureau of Labor Statistics. The resulting transaction data were merged with tax records in the same year so that the property-level attributes can be observed in the year that the price was generated.

Geographic Data

The GIS data consist of a collection of geographic layers for coastal CT and MA. The

precision of all other geographic analyses depends on the precision of property locations,¹⁴ for which two groups of data are acquired: parcel polygons with addresses from the local authorities and building footprints acquired from Connecticut Environmental Conditions Online (CT ECO) and Microsoft Building Footprints. After being revised via Google geocoding API, the address points from the property data are overlaid on the parcel polygon layers and the building footprint layers. We matched the address points to the nearest building centroids, then revised or dropped the wrongly matched address points so that the resulting address points fall into the corresponding building footprints (i.e., important when conducting viewshed analysis). The detailed coordinate revision process is reported in [Appendix D](#), and by-town statistics of original and revised coordinates are presented in [Appendix Table D1](#) for CT.

Another important part GIS data is the LiDAR layers. LiDAR layers in this study are downloaded from the National Oceanic and Atmospheric Administration (NOAA), which details the post-Hurricane Sandy structure and bare earth elevation of a region roughly

¹³These auxiliary sale and deed records are gathered from Vision Government Solutions' Assessor's Online Database (VGS) and Connecticut Town Land Records. The VGS records are mainly used to check and revise the prices in ZTRAX; the process is detailed in [Appendix B](#). We do not directly use the records from VGS because the attached information for each property is limited (e.g., only five ownership change records are available). The Land Records are mainly used to check the quality of the mortgage status provided by ZTRAX, which seems to be quite accurate.

¹⁴As noted by many studies, the coordinates of address points in ZTRAX are not quite accurate.

within the 1 km buffer of the coastline.¹⁵ Based on the LiDAR data and building footprint data, viewshed analysis is conducted for single-family residents (more than 55,000 for CT and 332,000 for MA) within the LiDAR spatial domain.

To construct the flood zone indicator, data with time-appropriate flood zone information are crucial to our analyses. We collected waves of flood maps from the Federal Emergency Management Agency's (FEMA) national flood hazard layer (NFHL) (three waves dated September 2012, March 2017, and April 2019 for CT, and two waves dated March 2017 and August 2022 for MA).¹⁶ We matched the transaction data and the flood zones based on the flood map effective dates, which vary from town to town.¹⁷ To be consistent with the actual flood zone designation process,¹⁸ we intersected the building footprints (instead of the address points) by flood zones and decided the flood zone status by checking whether the building was wholly outside the flood zones. Knowing that the majority of flood zone amendments on individual properties may not be reflected in the flood map, we collected the Letters of Map Amendment from the state's FEMA office and revised the flood zone status accordingly.¹⁹

¹⁵The LiDAR layers are timed 2012 for CT and 2013 for MA. The spatial domain of this LiDAR dataset is the exact selection criterion for the set of properties on which we conduct viewshed analysis.

¹⁶This is a FEMA geospatial database containing effective and historical flood hazard data.

¹⁷The exact rule of this assignment, CT for example, is if the transaction date is before the 2013 wave of remapping, it is matched with the 2012 NFHL map (the first wave of the digitized map); if the transaction date is after the 2013 dates of remapping but before the 2017 dates of remapping (or there is no 2017 remapping in that town), it is matched with the 2017 NFHL map; if the transaction date is after the dates of 2017 wave of remapping and the transaction is in the four towns with 2017 remapping, it is matched with the 2019 NFHL map. For MA, the remapping effective dates start from September 2017 and persist until 2022.

¹⁸If a part of the house is in the flood zone, the whole house is in the flood zone. If a part of the house is in a flood zone type that is considered of higher risk, the whole house is in that flood zone type (e.g., if the house incorporates a part in zone AE and a part in zone VE, the whole house is designated as a zone VE house).

¹⁹Starting in 2010, the letters for studied coastal towns are publicly available online. Earlier letters were acquired from the local FEMA office and digitized by the authors.

Other geographic datasets include the waterbody layer from the U.S. Geological Survey, the shoreline and estuary data from NOAA's continually updated shoreline product, the landcover data from National Land Cover Database, the highway exit layer produced by Tele Atlas North America, the streets and roads layers from the U.S. Census Bureau TIGER/line geodatabase (USCB TIGER), the school district layers (i.e., elementary, secondary, and unified school districts²⁰) from USCB TIGER, and the census tracts with demographic data from American Community Survey (ACS USCB).

The property-sale dataset was then matched with the geographic variables and went through a data cleaning process ([Appendix A](#)), which resulted in a dataset containing 96,092 transactions in CT and 211,735 in MA between 1998 and 2020, with a comprehensive set of the corresponding house, lot, and geographic attributes. These transactions correspond to 67,789 properties in CT and 148,693 in MA, with 7,435 and 13,601 falling in high-risk SFHA zones (A, AE, AO, or VE), respectively.

Major efforts to acquire and improve the data are summarized in [Appendix Table A2](#). [Appendix Figure A1](#) shows a concept map of the data refinement process.

Viewshed Analysis

Previous hedonic studies have found that views of particular geographic features are significant determinants of housing prices (Geoghegan, Wainger, and Bockstael 1997; Benson et al. 1998; Lake, Lovett, and Day 2000; Bateman et al. 2002; Paterson and Boyle 2002; Bin et al. 2008; Schmitz 2008). In coastal housing markets where flood hazards and ocean views are highly correlated, it is impossible to estimate a consistent hedonic coefficient of flood zone without acquiring a proper measure of ocean view. View attributes in commonly available real estate datasets are not accurate. Schmitz (2008) found that viewshed analysis results, even with only bare earth elevation, are much more accurate in positive

²⁰The eventually used school zone indicator will be an intersect of these three school district layers.

water view counts than what is reported in the multiple listing service data. Via viewshed analysis (i.e., a GIS function) based on LiDAR data (i.e., containing high-resolution structural elevation information), Bin et al. (2008) constructed a “viewscape” measure representing the degree of ocean view angle in a one-mile distance for 1,075 coastal single-family residents in North Carolina. We improved the viewshed calculation procedure reported in Bin et al. (2008) to make it scalable and applied it to more than 387,000 properties in coastal CT and MA. The process reported below is designed in ArcGIS and automated in Python 3 with Arcpy functions. We finished the viewshed analysis for MA (>332,000 properties) within 20 days.

The viewshed of a viewpoint is the part of a surface (i.e., represented by an elevation model) that is visible to the viewpoint. Viewshed analysis calculates viewsheds. The precision of viewshed analysis depends on the quality of the elevation surface and the specification of the viewpoint.

As the best available digital elevation data to generate the elevation surface, LiDAR datasets generated in 2012 (after Hurricane Sandy) were acquired for coastal CT and MA from NOAA.²¹ To measure the elevation of a targeted geographic feature, the LiDAR method illuminates the target with pulsed laser light beams, receives the reflected beams with a sensor, and calculates the elevation by analyzing the time span between the emission and the reception. Hence, the original LiDAR data are densely distributed points with elevation information (representing the structure elevation, or bare earth elevation if no structure exists). We aggregated these points with a mean function to generate elevation surfaces with a resolution of 5×5 ft.

To generate the proper viewpoints, address points from ZTRAX property data were relocated to the building footprint centroid via a process detailed in [Appendix D](#). After the checks and revisions, the viewpoints were correctly assigned to the corresponding building for more than 95% of the addresses falling in the LiDAR spatial domain. For the remaining

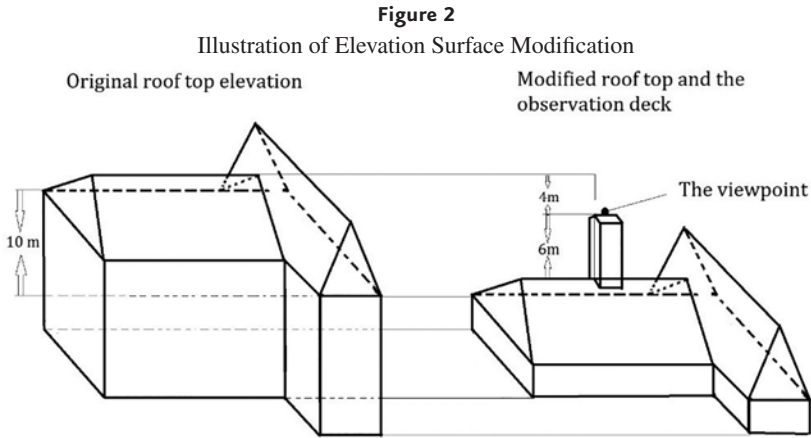
address points, we had insufficient information to guarantee that the address-building-matching is correct. Because there appears to be no systematic spatial pattern to these, they were purged from the analysis.

The proper specification of a viewpoint also requires modifying the elevation surface in the corresponding building footprint. We created an observation deck on top of the building footprint, which is a 6-m-tall (19.68 ft. tall) cuboid whose top surface is a 5×5 ft. square (i.e., one cell of the elevation surface) and has the same centroid as the building footprint. To mimic the average view from the highest floor of the building, we lowered the whole building footprint elevation (i.e., roof elevation) by 10 m (about 32.8 ft.).²² As a result, the viewpoint is 4 m lower than the original roof elevation, while the surrounding part of the roof is lowered by 10 m (as shown in Figure 2) to ensure that the view from the viewpoint is not blocked by the roof structures. This specification intends to mimic the viewpoint of a person standing on the highest floor of a house. This modification was only applied for a property when conducting viewshed analysis for that property.

We calculated the within-one-mile viewshed of each property in the LiDAR data domain. The elevation surface and the resulting viewsheds are shown in [Appendix Figure D4](#), which shows that the viewsheds resulting from the viewshed analysis properly reflect the structures on the elevation surface, including buildings, vegetation, and high grounds. These viewsheds were intersected with the ocean polygon (including Long Island Sound) so that the ocean view of each property can be recovered. Four variables were generated to deliver the information in the ocean view: ocean view area (in square feet), ocean view angle (in degrees), number of slices of ocean view, and ocean view distance (in feet).

²²To do this, one has to modify the part of elevation surface that falls in the building footprint when calculating viewshed. Because the clip function in ArcGIS does not clip through a raster cell, lowering the surface exactly within the building footprint will leave remaining high-elevation cells from the edge of the building, which, in the viewshed analysis, will function like high pillars blocking the view. To avoid this problem, we make a buffer (5 ft.) of the building footprint and lower the elevation surface within it.

²¹LiDAR layers can be accessed at <https://coast.noaa.gov/dataviewer/#/lidar/search/>.



Note: This modified building structure is merged into the elevation surface to replace the original building structure, and the viewshed analysis for this building is conducted with the modified surface (i.e., other structures are not modified).

Variables

Finally, we construct a set of variables describing the property attributes that coincide with the transactions.²³ Based on the calibrated address coordinates and available GIS data, a set of geographic variables were constructed for each property.²⁴ The summary statistics of selected variables in the original sample are presented per county group in Table 1 (check [Appendix E](#) for summary statistic tables for all variables, including by-flood-zone and postmatching statistics).²⁵

²³This set of property variables include post-FIRM status, building square footage, lot square feet, ground elevation, building age, with pool, garage capacity (number of cars), number of stories, the total number of rooms, the total number of bedrooms, total calculated bath count, air conditioning, fireplace number.

²⁴This set of geographic variables include floodplain status, surge from Superstorm Sandy in feet, BFE in feet, ocean-front (house on coastline, only for CT), riverfront (house on river, only for CT), waterfront across street (access to the ocean only blocked by a street, only for CT), ocean view area in square feet (within a one-mile radius), ocean view total angle in degrees, number of ocean view slices (how many obstructions to full view), distance to the nearest highway exit, distance to the nearest highway (primary roads and secondary roads), distance to the nearest railroad, distance to the coastline, distance to the nearest airport, proportion of agricultural land within a half-mile radius, proportion of developed land within a half-mile radius circle, proportion of forest land within a half-mile radius, proportion of open space or wetland within a half-mile radius.

²⁵Distance variables are in log forms in the analyses that follow. They are generally measured in two ways: if the

3. Methodology

This study attempts to consistently estimate the hedonic parameters of coastal amenities and flood zone status. Although the flood zone designation is not likely endogenous,²⁶ the estimates of interest are likely to be subject to misspecification biases, attenuation biases, and confounding effects from unobservable factors. The data-refinement process largely mitigates the latter two problems, but a well-designed identification strategy is still desired to provide consistent coefficient estimates for the flood zone status and the coastal amenities. We implement a quasi-experimental design.

In the first step, we use a nearest-neighbor matching algorithm to improve covariate overlap and alleviate misspecification bias on the treatment effect (i.e., the flood zone effect). Matching is widely used as a nonparametric preprocessing method to drop control units (i.e., transactions of property outside the flood zone in this study) that are quite different from

variable indicates an amenity or disamenity accessible directly through the environment, it is measured in hundreds of feet; if it indicates a facility that involves driving, it is measured in miles. An example is that the distance to nearest highway is measured in hundreds of feet, while the distance to nearest highway exit is measured in miles.

²⁶Conditional on observables (including elevations and distance to the coast), the flood zone designation is not likely to be decided by the dependent variable (housing prices) or emerge from the price generation process.

Table 1
Summary Statistics Original Sample

Variable	Mean (SD)					
	CTCG1	CTCG2	CTCG3	MACG1 ^a	MACG2	MACG3
Housing price (\$)	625,980.30 (534,865.48)	346,222.02 (244,518.71)	368,298.016 (270,987.67)	301,488.34 (200,161.84)	575,629.36 (391,224.77)	474,796.33 (263,422.31)
Flood zone (SFHA)	0.115 (0.319)	0.094 (0.292)	0.126 (0.332)	0.075 (0.264)	0.111 (0.314)	0.098 (0.297)
Ln (ocean view area [sq. ft.])	0.409 (2.269)	0.876 (3.386)	1.301 (4.068)	0.909 (3.377)	0.549 (2.657)	0.900 (3.419)
Ln (ocean view angle)	0.056 (0.400)	0.141 (0.660)	0.203 (0.799)	0.130 (0.620)	0.074 (0.466)	0.136 (0.647)
Ocean viewshed slice no.	0.096 (0.891)	0.202 (1.477)	0.329 (1.721)	0.334 (2.576)	0.213 (2.648)	0.344 (2.686)
Ocean viewshed distance (ft.)	5,167.197 (722.325)	5,035.694 (1,041.586)	4,906.266 (1,271.953)	5,039.460 (1,008.281)	5,138.075 (777.746)	5,032.400 (1,032.181)
Coast front	0.004 (0.064)	0.015 (0.123)	0.019 (0.137)	—	—	—
River front	0.006 (0.075)	0.004 (0.062)	0.017 (0.129)	—	—	—
Waterfront across street	0.002 (0.041)	0.004 (0.063)	0.007 (0.080)	—	—	—
Building sq. ft. (living space)	1,985.692 (1,067.370)	1,701.586 (778.764)	1,725.769 (785.382)	1,669.738 (516.570)	1,759.958 (603.861)	1,807.735 (634.350)
Lot sq. ft.	18,750.658 (22,929.872)	20,558.661 (31,212.687)	30,499.819 (50,645.934)	16,955.178 (53,425.323)	23,313.372 (55,341.864)	15,473.128 (26,889.447)
Building age	58.932 (30.157)	59.062 (32.146)	53.989 (33.479)	66.903 (37.426)	41.227 (29.776)	65.608 (37.758)
Ground elevation	93.555 (68.233)	69.078 (69.921)	65.164 (58.312)	22.471 (13.260)	12.907 (8.975)	19.478 (13.539)
Total bedrooms	3.275 (0.869)	3.140 (0.845)	3.122 (0.839)	2.954 (0.855)	3.050 (0.836)	3.049 (0.961)
Ln (distance to coastline [hundreds ft.])	4.034 (0.990)	3.642 (1.286)	3.576 (1.332)	3.507 (1.070)	2.920 (1.023)	3.133 (1.186)
Ln (distance to nearest highway exit [miles])	0.618 (0.252)	0.770 (0.361)	0.899 (0.362)	0.865 (0.498)	1.458 (0.721)	1.116 (0.505)
Ln (distance to nearest highway [hundreds ft.])	2.582 (0.821)	2.511 (0.904)	2.600 (0.946)	3.346 (1.216)	3.400 (1.445)	2.705 (1.036)
Ln (distance to nearest airport [hundreds ft.]) ^b	5.312 (0.677)	5.017 (0.802)	5.053 (0.717)	4.294 (0.671)	4.875 (0.791)	4.797 (0.743)
<i>N</i>	42,314	32,250	19,042	15,541	45,223	130,389

Note: CTCG / MACG = county groups; SFHA = special flood hazard area.

^a The waterfront variables are not generated for Massachusetts.

^b The airports here include small airports to approximate the disamenities from being close to an airport.

the treated (i.e., transactions within the flood zone) units (e.g., Abbott and Klaiber 2013; Lawley and Towe 2014; Alix-Garcia, Sims, and Yañez-Pagans 2015; Muehlenbachs, Spiller, and Timmins 2015; Johnston and Moeltner 2019; Towe and Chen 2022). As a result, a matched sample is generated in which the overlap in the covariate distributions is improved from the original sample (Dehejia and Wahba 1999; Rubin 2006; Ho et al. 2007;

Stuart 2010; Imbens and Wooldridge 2009), and the average treatment effect on the treated (ATT) can be recovered with fewer concerns on misspecification bias (Rubin 1979; Ho et al. 2007; Imbens and Wooldridge 2009; Abbott and Klaiber 2013; Johnston and Moeltner 2019).

To mitigate confounding effects from unobservable spatial and temporal factors, unobservable socioeconomic factors, and major

regulation heterogeneities (i.e., buyers with a mortgage loan have to pay for insurance up front), the matching process requires exact matching on transaction year, mortgage status (i.e., with a mortgage or not), neighborhood category (i.e., high-income block or not),²⁷ county group, and coastal proximity band (i.e., in the viewshed analysis area or not). We dropped all observations within “cells” (defined by the full interaction of discrete variables in the exact matching) that did not contain at least two treated or at least four controls,²⁸ to ensure the matching quality (Abadie and Imbens 2006). We recognize that finer exact matching categorical variables (e.g., requiring an exact match on the school zone) will mitigate more unobservable effects but overly restricting the exact match can lead to poor matching quality due to the curse of dimensionality (Abbott and Klaiber 2013; Johnston and Moeltner 2019).²⁹

In each exact matching cell, the matching application searches for the nearest neighboring control of a treated observation based on the Mahalanobis distance of a comprehensive set of continuous property attributes.³⁰

²⁷The income is calculated with 2017 census data on block groups in the study area. As the block-median-income distribution shows a normal-type distribution pattern with a long tail above \$150,000 per year, we classify households in the block groups with median-annual-income higher than \$150,000 as being in a high-income neighborhood.

²⁸In later analyses, roughly 1% of the treated are dropped in this process.

²⁹Intuitively, more granular categories will divide the space into much smaller “cells,” and the possibility of finding a good match for a treated observation will drop fast because observations falling in the same cell decrease quickly. Although unobservables on more detailed levels (e.g., quarterly level, school zone level, and quarterly by school zone level) will not be considered by this matching process, regressions presented immediately below will address the confounding effects from them.

³⁰This set of continuous variables includes (natural) log of ocean view area, log of ocean view angle, log of ocean viewshed distance, relative ground elevation (deviation from the town means), log of distance to the coastline, log of distance to the nearest highway (primary roads and secondary roads), log of distance to the nearest highway exit, log of distance to the nearest railroad, log of distance to the nearest airport, ratio of developed land within a half-mile radius, ratio of forest land within a half-mile radius, ratio of open space or wetland within a half-mile radius, building square feet, lot square feet, building age, number of buildings, total calculated bathroom count, garage capacity (number of

In the second step, we ran a weighted regression (noted as matched regression hereafter) based on the matched sample, where the weights are generated by the matching process. These weights are used to maintain the balance between the number of treated and the number of controls in the regression. The regression is described by equation [1] (note that the weights are not written out). Y represents the outcome variable—the natural log of transaction prices— C denotes the set of k continuous covariates (including coastal amenities) in the matching, S denotes the set of l categorical property characters (including building condition, sewer service status, and heating type) that cannot be put into the exact match due to the curse of dimensionality, i is the property identifier, t is used to index periods (i.e., year-quarter) and denotes quarterly level fixed effects, m denotes month-in-a-year fixed effects to control for seasonality, and st denotes census tract by year fixed effects. The flood zone effect is represented by τ , and the ocean amenity effects are represented by β^k . [Appendix E](#) gives more details about this research design.

$$Y_{it} = \alpha + \tau \cdot SFHA_{it} + \sum_k \beta^k C_{it}^k + \sum_l \gamma^l S_{it}^l + m + t + st + \varepsilon_{it}. \quad [1]$$

Similar matched regression approaches are applied and recommended in many previous studies (e.g., Ho et al. 2007; Abadie and Imbens 2011; Alix-Garcia, Sims, and Yañez-Pagans 2015; Towe and Chen 2022). This hedonic estimation approach was conducted for each county group (as noted in Figure 1). To investigate the effects of the choice of neighbor numbers in the matching process, we estimated different models with different numbers of neighbors in matching as a robustness check ([Appendix I](#)). In the end, we aimed to present a simplified process (regarding data refinement and modeling) that is scalable to other coastal areas, which potentially means an ocean amenity vector C with a lower dimension (below k) or representing slightly

cars), fireplace number, total number of rooms, total number of bedrooms, number of stories.

Table 2
Panel Results: Full Models for Connecticut

Sample ^a	CTCG1	CTCG1	CTCG2	CTCG2	CTCG3	CTCG3
	OLS	Matched	OLS	Matched	OLS	Matched
Outcome	Ln(price)					
Flood zone (SFHA)	-0.0142 (0.0152)	-0.0158 (0.0100)	-0.0208 (0.0137)	-0.0382*** (0.00907)	0.00925 (0.0196)	-0.00439 (0.0200)
Ln (ocean view area)	0.00184 (0.00279)	-0.000198 (0.00194)	0.000500 (0.00268)	0.00431 (0.00350)	0.00615* (0.00261)	0.00996*** (0.00264)
Ln (ocean view angle)	0.0241 (0.0228)	0.0462* (0.0208)	0.0336*** (0.00813)	0.0297** (0.00885)	0.0233+ (0.0132)	0.0118 (0.0121)
Ocean viewshed slice no.	-0.00947* (0.00441)	-0.00841 (0.00600)	-0.00623** (0.00189)	-0.00500* (0.00200)	0.00151 (0.00365)	0.000228 (0.00393)
Ln (ocean view distance)	-0.0437*** (0.0110)	-0.0383** (0.0130)	-0.0222* (0.0108)	-0.0153 (0.00920)	0.00195 (0.0128)	0.0119 (0.0130)
Coast front	0.136+ (0.0740)	0.120 (0.0839)	0.454*** (0.0436)	0.427*** (0.0321)	0.459*** (0.0345)	0.432*** (0.0414)
River front	0.124* (0.0525)	0.118* (0.0479)	0.310*** (0.0449)	0.233** (0.0694)	0.302*** (0.0373)	0.293*** (0.0438)
Waterfront across street	0.191** (0.0679)	0.170* (0.0653)	0.282*** (0.0328)	0.240*** (0.0419)	0.254*** (0.0427)	0.250*** (0.0532)
Ln (distance to coastline)	-0.0946*** (0.0202)	-0.107*** (0.0258)	-0.0222* (0.0100)	-0.00586 (0.0140)	-0.0582*** (0.0161)	-0.0660*** (0.0211)
Month fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year-quarter fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Census tract by year	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.902	0.901	0.813	0.849	0.775	0.789
Observations in SFHA	4,867	4,867	3,024	3,024	2,394	2,394
Observations outside SFHA	37,447	3,041	29,226	2,262	16,648	1,553
Observations	42,314	7,908	32,250	5,286	19,042	3,947

Note: Standard errors are in parentheses. Robust standard errors clustered by census tract are reported. CTCG = county groups; SFHA = special flood hazard area.

^a The matching process for “matched” required exact match on transaction year. In addition, this one-on-one-with-replacement matching requires exact matching on coastal proximity bands, neighborhood category, and mortgage status.

+ $p < 0.1$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

different variables (categorical instead of numerical measures).

4. Results

Before delving into the hedonic coefficient estimates of coastal amenities and flood zone, we present the by-flood-zone prices of the coastal housing market in CT and MA. The summary statistics of housing prices by flood zones are shown in [Appendix Table C1](#), and the trends of average prices by SFHA status are shown in [Appendix Figure D1](#). The flood zones are always correlated with higher housing prices, which reflects the fact that coastal flood zones are usually associated with higher levels of various water-related amenities compared with non-flood-zone areas. This

highlights the necessity to tease out amenity effects when identifying the flood zone price effect.

Parameter Estimates of the Full Model

The effects of the flood zones and coastal amenities on housing prices are estimated with the primary strategy presented above as well as alternative strategies for comparison (i.e., simple hedonic regressions with the unmatched sample). Because waterfront variables are not available for MA, the results displayed in Table 2 are only for CT. While the baseline ordinary least squares (OLS) estimates do not give any significant flood zone effects, the matched regression pushes the flood zone effects for all the county groups uniformly to the negative side and shows that

the discount is about 3.8% (significance level 0.001) for CTCG2 (New Haven County). These results suggest that the estimation strategies combined with proper coastal amenities effectively separate the water-related amenity effects from the flood zone effects. The results suggest significant price premiums for many coastal amenities across the board, while the significance and magnitude for each view and waterfront variable are slightly different for different locations. [Appendix H](#) displays the full set of coefficient estimates.

The covariate balance of the prematch and postmatch sample is described via Q-Q plots and standardized distances in [Appendix Figure E1](#). For simplicity, we group the county groups in one figure. These results show that the covariate discrepancies are narrowed from the prematch sample to the postmatch sample, but sizable differences remain in relative ground elevation (the rule of thumb is 0.25 for standardized distances as suggested in Imbens and Wooldridge 2009).³¹

Simplified Models with Ocean View Representing Coastal Amenities

The complexity of the ocean amenity generation process here might seem onerous for large-scale applications, especially since the waterfront variables are generated with a case-by-case inspection. So we adopted a simplified specification to test whether including only the ocean view variables can capture the waterfront premiums, as these amenities are highly correlated. The interpretation of view coefficients from this simplified model should be different from the full model as waterfront effects are absorbed into the view effects. However, if the explanatory power (*R*-squared) of the simplified model is close to the full model and resulting policy variable estimates (the flood zone effects) are of little qualitative difference, this simplified model is still meaningful and yields accurate policy

implications. Table 3 shows the simplified model estimates for CTCG2 and a convenient model that does not include the view variables nor the waterfront variables (i.e., ocean amenities are approximated by coastal proxy and elevation). The coefficient estimates for CTCG2 demonstrate that the simplified model has a similar *R*-squared (0.835 vs. 0.849) and still yields a significant flood zone discount (though less sizable) compared with the full model. In contrast to the convenient model, it is clear that including only the ocean view variables achieves the task of disentangling the effects of coastal hazards and coastal amenities.

The hedonic prices of ocean view can be compared with Benson et al. (1998). Since Benson et al. (1998) used qualitative measures on ocean view, we adopt qualitative variable specifications for the comparison. Benson et al. classify ocean view into four categories: category 1 represents unobstructed full ocean view, category 2 represents good ocean view with some obstructions, category 3 represents ocean view with significant obstructions, and category 4 represents poor ocean view. Because Benson et al. (1998) do not have waterfront variables, it is reasonable to assume that category 1 view absorbs oceanfront effects. We classify our ocean amenities into four ocean view categories (see [Appendix F](#) for detailed definition and distributions) in an effort to mimic the four-category classification in Benson et al. (1998) and estimate the qualitative models based on the matched sample.³² As shown in the last three columns of Table 3, the qualitative models offer similar explanatory power and flood zone effects, compared with the simplified specification and the full specification. The coefficient estimates in the qualitative models are surprisingly close to Benson et al. (1998) (e.g., their ocean view category 1 coefficient spans from 0.46 to 0.49 from 1990 to 1993, falling in the ballpark of our estimates of 0.421 to 0.525), although there are differences in study periods and regions. These estimates are practically important because categorical view variables are more likely to be used in real-life

³¹We changed the specification for elevation to check whether it matters for the matching and for the regression coefficients; the final results are robust to these changes. After controlling all the other coastal amenities and disamenities, elevation coefficients are not at all statistically significant, meaning that it's not a strong factor of housing prices (conditional on the controls).

³²To be consistent with Benson et al. (1998), the waterfront variables are not included in these qualitative models.

Table 3
Panel Results: Simplified and Qualitative View Models for Connecticut

Sample	CTCG2 Matched ^a	CTCG2 Matched	CTCG1 Matched	CTCG2 Matched	CTCG3 Matched
Model	Simplified (No Waterfront)	Convenient	Qualitative View	Qualitative View	Qualitative View
Outcome	Ln(price)				
Flood zone (SFHA)	-0.0283* (0.0110)	-0.0140 (0.0145)	-0.0161 (0.00975)	-0.0304* (0.0114)	0.000424 (0.0209)
Ln (ocean view area)	-0.00816* (0.00373)				
Ln (ocean view angle)	0.0974*** (0.00950)				
Ocean viewshed slice no.	-0.00640** (0.00200)				
Ln (ocean view distance)	-0.0502*** (0.0137)				
Ocean view category 1			0.421*** (0.0472)	0.525*** (0.0323)	0.437*** (0.0773)
Ocean view category 2			0.365*** (0.0366)	0.409*** (0.0277)	0.393*** (0.0504)
Ocean view category 3			0.246*** (0.0497)	0.263*** (0.0429)	0.223*** (0.0448)
Ocean view category 4			0.0735*** (0.0159)	0.0773** (0.0263)	0.0922*** (0.0208)
Ln (distance to coastline)	-0.0397* (0.0194)	-0.112*** (0.0245)	-0.138*** (0.0184)	-0.0473* (0.0196)	-0.129*** (0.0184)
Month fixed effects	Yes	Yes	Yes	Yes	Yes
Year-quarter fixed effects	Yes	Yes	Yes	Yes	Yes
Census tract by year	Yes	Yes	Yes	Yes	Yes
R-squared	0.836	0.805	0.900	0.833	0.769
Observations in SFHA	3,024	3,024	4,867	3,024	2,394
Observations outside SFHA	2,262	2,342	3,041	2,262	1,553
Observations	5,286	5,366	7,908	5,286	3,947

Note: Standard errors are in parentheses. Robust standard errors clustered by census tract are reported. CTCG = county groups; SFHA = special flood hazard area.

^a The matching process required exact match on transaction year. In addition, this one-on-one-with-replacement matching also requires exact matching on coastal proximity bands (three bands defined coastal proximity percentiles), neighborhood category, and mortgage status. Note that the matching process in the convenient models do not involve ocean view variables.

⁺ $p < 0.1$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

communications compared with numerical view measures.

Results here suggest that, while properly measured ocean amenities are crucial in coastal hedonic studies on hazard indicators, the model may not have to involve an exhaustive set of variables (i.e., including merely the ocean view variables might be enough), and the qualitative specification on amenity variables could function adequately in separating the pricing effect of the hazard variables. Although these implications seem insignificant at first glance, they are crucial in expanding the study to a much larger spatial extent: the

labor-intensive nature of inspecting direct water access makes mass processing waterfront variables difficult, but the viewshed calculation could be automated and hence scalable.

Extending the Analysis to MA

Based on these results, we can extend the analysis to the coastal towns of MA and estimate a pricing effect of flood zone status that is close enough to the true effect. The simplified model and qualitative model for MA are presented in Table 4. We found a significant flood zone discount in MACG1 (Bristol

Table 4
Panel Results: Simplified and Qualitative View Models for Massachusetts

Sample ^a	MACG1 Matched	MACG2 Matched	MACG3 Matched	MACG1 Matched	MACG2 Matched	MACG3 Matched
Model	Simplified	Simplified	Simplified	Qualitative View	Qualitative View	Qualitative View
Outcome	Ln(price)					
Flood zone (SFHA)	-0.0720** (0.0194)	0.0455* (0.0178)	-0.0115 (0.0123)	-0.0733** (0.0201)	0.0428* (0.0183)	-0.0143 (0.0120)
Ln (ocean view area)	-0.00975* (0.00376)	0.0103** (0.00361)	-0.00440** (0.00132)			
Ln (ocean view angle)	0.0503* (0.0190)	0.0294 (0.0187)	0.0596*** (0.00923)			
Ocean viewshed slice no.	0.00142 (0.00120)	0.00203*** (0.000579)	0.000564 (0.000730)			
Ln (ocean view distance)	-0.0504** (0.0153)	-0.0354** (0.0120)	-0.0290*** (0.00707)			
Ocean view category 1				0.288** (0.0864)	0.444*** (0.0667)	0.373*** (0.0468)
Ocean view category 2				0.179* (0.0652)	0.398*** (0.0466)	0.260*** (0.0321)
Ocean view category 3				0.0715** (0.0219)	0.428*** (0.0830)	0.185*** (0.0323)
Ocean view category 4				0.0348* (0.0165)	0.197*** (0.0273)	0.0505*** (0.0127)
Ln (distance to coastline)	-0.170*** (0.0300)	-0.221*** (0.0171)	-0.106*** (0.0116)	-0.186*** (0.0250)	-0.227*** (0.0175)	-0.115*** (0.0118)
Month fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year-quarter fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Census tract by year	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.798	0.762	0.777	0.796	0.759	0.775
Observations in SFHA	1,169	5,020	12,730	1,169	5,020	12,730
Observations outside SFHA	852	3,700	8,815	852	3,700	8,815
Observations	2,021	8,720	21,545	2,021	8,720	21,545

Note: Standard errors are in parentheses. Robust standard errors clustered by census tract are reported. MACG = county groups; SFHA = special flood hazard area.

^a The matching process for "matched" required exact match on transaction year. In addition, this one-on-one-with-replacement matching also requires exact matching on coastal proximity bands, neighborhood category, and mortgage status.

⁺ $p < 0.1$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

County) and an insignificant flood zone discount in MACG3 (Suffolk, Plymouth, Norfolk, and Essex Counties). Interestingly, the flood zone effect on prices is significantly positive for MACG2 (Dukes, Nantucket, and Barnstable Counties).³³ This could be related to the nature of the communities and vast investment in coastal properties,³⁴ which lead to coastal property features (e.g., fortification,

luxurious design, and expensive construction materials) that cannot be covered in our data and confound the flood zone effects.³⁵

Hazard Discounts and Amenity Premiums

To give a clear view of the price differences associated with flood zone designation and coastal amenities, we transfer the hedonic estimates to price discounts or premiums

³³This result is robust, as shown in the check against different data process and different number of neighbors.

³⁴It is perceived that many of the properties are vacation homes, and the owners are paying more attention to coastal amenities rather than flood risk. One can tell this from the extra high ocean view coefficients in the second-to-last column in Table 4.

³⁵These luxurious features are not fully represented by the coastal proximity and ocean view variables in our data, but they are highly correlated with the flood zone variable and property prices.

Table 5
Average Estimates of Flood Zone Discounts and Coastal Amenity Premiums for Coastal Towns in Connecticut and Massachusetts

	CTCG1	CTCG2	CTCG3	MACG1	MACG2	MACG3
Counties	Fairfield	New Haven	New London Middlesex	Bristol	Dukes Nantucket Barnstable	Suffolk Plymouth Norfolk Essex
Flood zone (SFHA) ^a	(-1.60%) ^b [-0.0346, 0.0030]	-2.99% [-0.0514, -0.0080]	(0.04%) [-0.0397, 0.0423]	-7.07% [-0.1066, -0.0333]	(4.37%) [0.0070, 0.0818]	-1.42% [-0.0371, 0.0093]
Ocean view category 1	52.35% [0.3889, 0.6711]	69.05% [0.5868, 0.8009]	54.81% [0.3304, 0.8013]	33.38% [0.1260, 0.5799]	55.89% [0.3679, 0.7767]	45.21% [0.3248, 0.5916]
Ocean view category 2	44.05% [0.3408, 0.5476]	50.53% [0.4258, 0.5893]	48.14% [0.3421, 0.6352]	19.60% [0.0525, 0.3591]	48.88% [0.3589, 0.6312]	29.69% [0.2178, 0.3811]
Ocean view category 3	27.89% [0.1602, 0.4097]	30.08% [0.1959, 0.4149]	24.98% [0.1448, 0.3645]	7.41% [0.0290, 0.1212]	53.42% [0.3038, 0.8052]	20.32% [0.1294, 0.2819]
Ocean view category 4	7.63% [0.0432, 0.1103]	8.04% [0.0261, 0.1375]	9.66% [0.0528, 0.1422]	3.54% [0.0025, 0.0694]	21.77% [0.1543, 0.2847]	5.18% [0.0259, 0.0783]
Single-family residence price average (2017\$)	625,980	346,222	368,298	301,488	575,629	474,796

Note: CTCG / MACG = county groups; SFHA = special flood hazard area.

^a Mean price percentage differences (attribute level 1 vs. level 0, holding all other attributes constant and with level 0 price as the base) are reported as well as the 95% confidence intervals (in square brackets). To give a better idea of how much dollar values are associated, the average coastal property prices of the corresponding region is reported.

^b The flood zone discounts in parentheses suggest that these discounts are not supported by high-confidence statistical results due to power issues or alternative explanations, so any interpretation of these numbers should use caution.

in Table 5. We use the panel estimates with qualitative view variables (Table 3 for CT and Table 4 for MA) as the base of this transformation. The average price premium of coastal amenities represented by view category 1 ranges from 52.35% to 69.05% in CT and 33.38% to 55.89% in MA, and the premium for view category 2 is also quite sizable.

5. Robustness Check and Discussion

Robustness against Number of Neighbors in Matching

A common discussion on using matching is the robustness of the estimates against the choice of number of neighbors in the matching algorithm (e.g., Ferraro, McIntosh, and Ospina 2007; Lawley and Towe 2014). We varied the number of neighbors in the matching algorithm, generated different matched

samples, and show how these changes affect our final estimates. [Appendix I](#) shows the estimates with more matching neighbors ([Appendix Table I1](#) for CT and [Appendix Table I2](#) for MA), which are comparable with the main matched regression estimates in Tables 2 and 4. We find that the coefficient estimates are generally robust to the number of neighbors. The noticeable qualitative difference is that the coefficients of ln(ocean view angle) for CTCG3 and MACG2 are not statistically significant in Table 2, while they become significantly positive after including more neighbors (in [Appendix Tables I1 and I2](#), respectively). This phenomenon suggests that some coefficient estimates in the main results may be affected by the randomness involved when using only one neighbor, and including more neighbors mitigates such kind of negative impact (of using one neighbor) at the cost of the overall quality of the matching (i.e., leading to slightly bigger discrepancies in covariates).

Null Effect Estimates and Statistical Power Calculation

There are insignificant flood zone effect estimates for certain county groups across different specifications (for CTCG1, CTCG3, and MACG3 in Tables 2–4). To interpret these null results, it is necessary to know whether we have sufficient statistical power to detect a certain effect size that deviates from zero (Ferraro and Shukla 2020; Murfin and Spiegel 2020). [Appendix J](#) clarifies this issue by conducting power calculations on the matched samples in CTCG1, CTCG3, and MACG3.

In power analysis, the recommended sample size at a certain effect size should be compared with the treatment group with a smaller size. Therefore, when we only use one matched neighbor (the main specifications), the group sizes are decided by the control group. When we involve multiple matched neighbors (under “Robustness against Number of Neighbors in Matching” above and in [Appendix I](#)), the group sizes are generally decided by the treated group.

From [Appendix I](#), we learn that the insignificant flood zone effects stay insignificant across specifications with different numbers of matched neighbors, which, in combination with the power calculation in [Appendix J](#), allows us to derive general conclusions for the flood zone effects in CTCG1, CTCG3, and MACG3. The power analyses show that we do not have a statistical power of 80% to claim that the insignificant estimates for CTCG1 and CTCG3 are null effects (instead of Type II errors). For MACG3, we do have a statistical power of 80% to claim that the insignificant estimates are null effects, which suggests that the average flood zone discount in MACG3 is close to zero. These conclusions are meaningful because they give further evidence of the heterogeneity in flood zone discounts across county groups, as we have estimated significant flood zone discounts of about 3% for CTCG2, about 7% for MACG1, and significant flood zone premiums for MACG2.

Caveats and Remaining Concerns on Endogeneity

Because flood zone is decided exogenously to the housing market, previous hedonic literature does not consider the endogeneity issue when assessing its price effect. The endogeneity, however, does exist due to systematic sorting over unobserved individual tastes and tolerances. This endogeneity happens for flood zone and coastal amenities. A similar story in the hedonic analysis is pointed out by Bartik (1987, 83) that “a household with greater tastes for a characteristic will choose greater quantities of that characteristic.” The modified version of that story would be that households with a lower level of distaste of the flood risk (or a higher level of taste of ocean view) will choose to buy homes in the flood zone (or with better ocean views). If this is true, the estimated hedonic coefficients will underestimate the flood zone discounts and overestimate the coastal amenity premiums for the average resident.³⁶ Because this endogeneity is based on individual (or household) unobservable tastes, the techniques based on property-level variables do not eliminate it. More important, although this is in line with the risk-based sorting theory, it cannot be solved without a proper instrumental variable since the taste is unobserved (unlike race or income). Within the realm of this study, we acknowledge that this endogeneity issue is likely to exist and it potentially biases the estimates if the target is average estimates for the population.

6. Concluding Remarks

Building on a meticulous process of data collection and refinement, this study investigates the hedonic pricing of coastal amenities and hazard indicators. The results show that ocean view and being a waterfront property significantly increase coastal housing prices, and being in the flood zone significantly lowers coastal housing prices in New Haven

³⁶ Assuming that the average households in the matched control sample can represent the vast majority of the population in terms of taste.

County, CT, and Bristol County, MA. Based on a qualitative view analysis, we find that a good set of coastal amenities (represented by unobstructed wide ocean view) add 33.38%–69.05% to the coastal property prices on the coasts of CT and MA. The simplified models and qualitative models yield implications that are qualitatively identical to the full models, and the estimates are robust to different specifications in data processing and matching.

This study entails a scalable process of calculating ocean views. As shown in the results, with proper measures of the coastal amenities, the valuation for the flood risk indicator is less misleading, which is likely true for other coastal hazards. In the face of rising coastal hazard risks brought about by climate change, it is an urgent task to calculate and evaluate coastal amenities on a large scale to better understand residents' preference toward coastal hazards and related policies. We conducted this exact task for an important part of the U.S. coast but a bigger contribution might be providing the scalable ocean view procedure. We hope to see many more studies on coastal amenities and hazards, using the process detailed here or developing new tools in the same spirit (e.g., a reliable algorithm to calculate waterfront status).

Our results, in combination with lessons from previous studies (Benson et al. 1998; Bin, Kruse, and Landry 2008; Bin et al. 2008; Murfin and Spiegel 2020), yield important policy implications. While the ocean amenities, represented by ocean views, are considerably and consistently valued across different years and locations, the flood risk is not capitalized nearly as much in the housing market.³⁷ This article suggests the inconsistency in flood

zone capitalization across New England counties, and Murfin and Spiegel (2020) show that the sea level rise indicators are not universally capitalized. These findings might suggest that the signaling process of the rising flood risks is not strong enough for the coastal residents to perceive it as strongly as they perceive the ocean amenities, which would lead to the policy implication that stronger signaling (e.g., higher incentives and easier access) are needed. From the perspective of improving policy efficiency, it will be helpful to study the existing signaling programs across time and space to determine their costs and benefits before making a conclusive statement on how to signal.

Acknowledgments

This article benefits from the funding and support provided by Connecticut Institute for Resilience and Climate Adaptation. The authors received many helpful insights and comments from Nancy Bockstael. We thank Diane Ifkovic and other practitioners for their generous help in accessing data and practical knowledge. We also thank Stephen Swallow and Kathleen Segerson; seminar participants at the University of Rhode Island, Mississippi State University, and Virginia Tech; and conference participants at the Association of Environmental and Resource Economists annual conference, Agricultural and Applied Economics Association annual conference, and the ZTRAX conference for helpful questions and discussions. Property data are provided by ZTRAX. More information on accessing the data can be found at <http://www.zillow.com/ztrax>. The results and opinions are those of the authors and do not reflect the position of Zillow Group. All remaining errors are ours.

References

- Abadie, A., and G. W. Imbens. 2006. "Large Sample Properties of Matching Estimators for Average Treatment Effects." *Econometrica* 74 (1): 235–67.
- . 2011. "Bias-Corrected Matching Estimators for Average Treatment Effects." *Journal of Business & Economic Statistics* 29 (1): 1–11.

³⁷A rough comparison on whether the flood zone discount low is to check the National Flood Insurance Program (NFIP) costs. The average annual NFIP premium is \$1,395 (in 2018) for a property of \$250,000 and \$2,133 if we project the rate to the full value of an average property (about \$611,000) in CT. The net present value is \$32,608 with a constant annual payment for an infinite horizon at a 7% discount rate and \$28,324 for 30 years. If we consider a complete capitalization of the net present value) of the future insurance payments the resulted flood zone discount is at least more than 4.5%. If we consider that the NFIP rate is highly subsidized (much lower than it should be), a flood zone discount of 3% is considerably lower than the fair discount of the flood risk.

- Abbott, J. K., and H. A. Klaiber. 2013. "The Value of Water as an Urban Club Good: A Matching Approach to Community-Provided Lakes." *Journal of Environmental Economics and Management* 65 (2): 208–24.
- Alix-Garcia, J. M., K. R. E. Sims, and P. Yañez-Pagans. 2015. "Only One Tree from Each Seed? Environmental Effectiveness and Poverty Alleviation in Mexico's Payments for Ecosystem Services Program." *American Economic Journal: Economic Policy* 7 (4): 1–40.
- Bartik, T. J. 1987. "The Estimation of Demand Parameters in Hedonic Price Models." *Journal of Political Economy* 95 (1): 81–88.
- Bateman, I. J., A. P. Jones, A. A. Lovett, I. R. Lake, and B. H. Day. 2002. "Applying Geographical Information Systems (GIS) to Environmental and Resource Economics." *Environmental and Resource Economics* 22 (1): 219–69.
- Bee, C. A. 2013. "The Geographic Concentration of High-Income Households: 2007–2011." Washington, DC: U.S. Census Bureau. Available at <https://www.census.gov/library/publications/2013/acs/acsbr11-23.html>.
- Beltrán, A., D. Maddison, and R. J. R. Elliott. 2018. "Is Flood Risk Capitalised into Property Values?" *Ecological Economics* 146: 668–85.
- Benson, E. D., J. L. Hansen, A. L. Schwartz, and G. T. Smersh. 1998. "Pricing Residential Amenities: The Value of a View." *Journal of Real Estate Finance and Economics* 16 (1): 55–73.
- Bin, O., T. W. Crawford, J. B. Kruse, and C. E. Landry. 2008. "Viewscapes and Flood Hazard: Coastal Housing Market Response to Amenities and Risk." *Land Economics* 84 (3): 434–48. <https://doi.org/10.3368/le.84.3.434>.
- Bin, O., J. B. Kruse, and C. E. Landry. 2008. "Flood Hazards, Insurance Rates, and Amenities: Evidence from the Coastal Housing Market." *Journal of Risk and Insurance* 75 (1): 63–82.
- Dehejia, R. H., and S. Wahba. 1999. "Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs." *Journal of the American Statistical Association* 94 (448): 1053–62. <https://doi.org/10.1080/01621459.1999.10473858>.
- DiVincenti, C. 2006. "The Evaluation of the National Flood Insurance Program." Final Report. Available at [http://www.vclevees.com/doc/NFIP%20Final%20Report%20\(October%202006\).pdf](http://www.vclevees.com/doc/NFIP%20Final%20Report%20(October%202006).pdf).
- Ferraro, P. J., C. McIntosh, and M. Ospina. 2007. "The Effectiveness of the U.S. Endangered Species Act: An Econometric Analysis Using Matching Methods." *Journal of Environmental Economics and Management* 54 (3): 245–61.
- Ferraro, P. J., and P. Shukla. 2020. "Feature: Is a Replicability Crisis on the Horizon for Environmental and Resource Economics?" *Review of Environmental Economics and Policy* 14 (2): 339–51. <https://doi.org/10.1093/reep/reaa011>.
- Geoghegan, J., L. A. Wainger, and N. E. Bockstael. 1997. "Spatial Landscape Indices in a Hedonic Framework: An Ecological Economics Analysis Using GIS." *Ecological Economics* 23 (3): 251–64.
- Gori, A., N. Lin, D. Xi, and K. Emanuel. 2022. "Tropical Cyclone Climatology Change Greatly Exacerbates US Extreme Rainfall-Surge Hazard." *Nature Climate Change* 12:171–78. <https://doi.org/10.1038/s41558-021-01272-7>.
- Ho, D. E., K. Imai, G. King, and E. A. Stuart. 2007. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis* 15 (3): 199–236.
- Imbens, G. W., and J. M. Wooldridge. 2009. "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economic Literature* 47 (1): 5–86.
- Johnston, R. J., and K. Moeltner. 2019. "Special Flood Hazard Effects on Coastal and Interior Home Values: One Size Does Not Fit All." *Environmental and Resource Economics* 74 (1): 181–210.
- Kulp, S. A., and B. H. Strauss. 2019. "New Elevation Data Triple Estimates of Global Vulnerability to Sea-Level Rise and Coastal Flooding." *Nature Communications* 10 (1): 1–12.
- Lake, I. R., A. A. Lovett, and B. H. Day. 2006. "Improving Land Compensation Procedures via GIS and Hedonic Pricing." *Environment and Planning C: Government and Policy* 18 (6): 681–96.
- Lawley, C., and C. Towe. 2014. "Capitalized Costs of Habitat Conservation Easements." *American Journal of Agricultural Economics* 96 (3): 657–72.
- Muehlenbachs, L., E. Spiller, and C. Timmins. 2015. "The Housing Market Impacts of Shale Gas Development." *American Economic Review* 105 (12): 3633–59. <https://doi.org/10.1257/aer.20140079>.
- Murfin, J., and M. Spiegel. 2020. "Is the Risk of Sea Level Rise Capitalized in Residential Real Estate?" *Review of Financial Studies* 33 (3): 1217–55. <https://doi.org/10.1093/rfs/hhz134>.

- Paterson, R. W., and K. J. Boyle. 2002. "Out of Sight, Out of Mind? Using GIS to Incorporate Visibility in Hedonic Property Value Models." *Land Economics* 78 (3): 417–25. <https://doi.org/10.2307/3146899>.
- Rubin, D. B. 1979. "Using Multivariate Matched Sampling and Regression Adjustment to Control Bias in Observational Studies." *Journal of the American Statistical Association* 74 (366a): 318–28.
- . 2006. *Matched Sampling for Causal Effects*. Cambridge: Cambridge University Press.
- Stuart, E. A. 2010. "Matching Methods for Causal Inference: A Review and a Look Forward." *Statistical Science: A Review Journal of the Institute of Mathematical Statistics* 25 (1): 1.
- Sander, H. A., and S. M. Manson. 2007. "Heights and Locations of Artificial Structures in Viewshed Calculation: How Close Is Close Enough?" *Landscape and Urban Planning* 82 (4): 257–70.
- Schmitz, N. 2008. "Viewshed Analyses to Measure the Impact of Lake Views on Urban Residential Properties." *Appraisal Journal* 76 (3): 224.
- Towe, C., and Z. Chen. 2023. "The Impact of Recreational Homes on Agricultural Land Use." *Land Economics* 99 (1): 17–37. <https://doi.org/10.3368/le.090420-0140R1>.