

IZA DP No. 10374

**Rising Sea Levels and Sinking Property Values:  
The Effects of Hurricane Sandy on New York's  
Housing Market**

Francesc Ortega  
Süleyman Taspınar

November 2016

# **Rising Sea Levels and Sinking Property Values: The Effects of Hurricane Sandy on New York's Housing Market**

**Francesc Ortega**

*Queens College, CUNY  
and IZA*

**Süleyman Taşpınar**

*Queens College, CUNY*

Discussion Paper No. 10374  
November 2016

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0  
Fax: +49-228-3894-180  
E-mail: [iza@iza.org](mailto:iza@iza.org)

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

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

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

## ABSTRACT

### **Rising Sea Levels and Sinking Property Values: The Effects of Hurricane Sandy on New York's Housing Market\***

Are coastal cities adjusting to rising sea levels? This paper argues that large-scale events have the potential to ignite the process. We examine the effects of hurricane Sandy on the New York City housing market. We assemble a large plot-level dataset with rich geographic data on housing sales in New York City for the period 2003-2015, along with information on which building structures were damaged by the hurricane, and to what degree. Our difference-in-difference estimates provide robust evidence of a negative impact on the price trajectories of houses that were directly affected by Sandy. Interestingly, this is also the case for houses that were not damaged but face high risk of coastal flooding. Our results suggest that Sandy has increased the perceived risk of living in those neighborhoods. We also show that the negative effects on housing prices appear to be highly persistent.

JEL Classification: H56, K42, R33

Keywords: climate change, real estate, cities, Hurricane Sandy

Corresponding author:

Francesc Ortega  
Department of Economics  
Queens College, CUNY  
300A Powdermaker Hall  
65-30 Kissena Blvd.  
Queens, New York 11367  
USA  
E-mail: [fortega@qc.cuny.edu](mailto:fortega@qc.cuny.edu)

---

\* We benefitted from insightful comments by Alberto Abadie, Ghazala Azmat, Meta Brown, Natalia Bailey, Melissa Checker, Marc Conte, Giacomo di Giorgi, Daniel Hamermesh, Andrew Haughwout, Jennifer Hunt, Wilbert van der Klaaw, John Landon-Lane, Donghoon Lee, Marco Manacorda, Roberto Pancrazi, Giacomo Ponzetto, Roland Rathelot, Thijs van Rens, Nuria Rodriguez-Planas, Albert Saiz, Joseph Tracy, Andrea Tesei, Dean Savage, Marcos Vera-Hernandez and Marija Vukotic. We also thank seminar participants at Rutgers, Queen Mary, Warwick, the New York Fed and Fordham.

# 1 Introduction

Currently, the rate of global mean sea level rise is about 3 centimeters per decade (Stocker et al. (2013)) and likely to accelerate in the coming decades. Almost unanimously, the scientific community predicts that this will lead to rising sea levels along virtually all ocean coastlines and to an increase in the occurrence of extreme weather events and flooding. The cumulative rise in sea levels will pose important economic challenges in many regions around the world. Arguably, dense urban areas on the shore will face the largest economic costs because of the infrastructure and housing stock that cannot be easily relocated.<sup>1</sup>

Fortunately, the factors behind rising sea levels are well understood (warming oceans, loss of ice by glaciers and the thinning of the ice sheets) and scientists have produced accurate projections of the resulting increases in the risk of large-scale flooding. These projections offer the possibility to undertake gradual adjustments that could lead to large cost savings, either by adopting resiliency measures or engaging in “graceful depopulation”.<sup>2</sup> Unfortunately, there are plenty of impediments to the adoption of a gradual approach, ranging from psychological biases to coordination problems, social ties, lack of accurate information, misguided policies, credit constraints, and the expectation of financial assistance by the government in case of disaster.<sup>3</sup>

The goal of this paper is to analyze whether an extreme weather event is able to set in motion the changes in expectations required to ignite the process of adjustment to the available information. We focus on the impact of hurricane Sandy on the demand for housing in New York City. Hurricane Sandy, which hit New York on October 29, 2012, was the largest Atlantic hurricane on record, and the second costliest in U.S. history (behind hurricane Katrina), with damages amounting to over \$19 billion. Hurricane Sandy flooded 17% of the city and nearly 90,000 buildings mainly located on coastal areas. Following the tradition of Oates (1969), we hypothesize that housing values in these areas may have been permanently depressed, reflecting an increase in the perceived risk of living by the ocean

---

<sup>1</sup>According to Climate Central, nearly 5 million people in the United States currently live at locations that are likely to be flooded by the end of the century. The challenges are even more severe for China, with several large and fast-growing coastal urban areas, such as Shanghai, Tianjin or Shantou. Other examples of large cities in coastal areas are Mumbai, Miami, and Osaka (Hanson et al. (2011)).

<sup>2</sup>This expression is borrowed from Philip Stoddard, the mayor of South Miami. The exact quotation is “(...) *there is no keeping the water out. So ultimately this area has to depopulate. What I want to work toward is a slow and graceful depopulation, rather than a sudden and catastrophic one.*” (excerpt from “The Siege of Miami” by Elizabeth Kolbert. The New Yorker, December 21-28, 2015).

<sup>3</sup>From a global perspective, adjustments to the rise in sea levels over the long run may also be constrained by restrictions to international migration. As argued by Desmet et al. (2015), the geographic world distribution of productivity and income in the future will be largely shaped by the evolution of international migration restrictions.

and a shift in the demand for housing away from these neighborhoods.

To address these questions we assemble a large plot-level dataset with rich geographic data, as in [Hornbeck and Keniston \(2014\)](#) and [Ambrus et al. \(2016\)](#). The data contain sale prices (and a number of characteristics) for all sales in New York City in the period 2003-2015, along with information on which building structures were damaged by hurricane Sandy and to what degree (provided by FEMA). Methodologically, we present difference-in-difference estimates of the effect of Sandy on housing prices, along with some more flexible specifications. In essence, identification of these effects is based on the change in housing values in (narrowly defined) neighborhoods affected by hurricane Sandy relative to unaffected neighborhoods.<sup>4</sup> Importantly, we distinguish between the direct effects from flooding and related damage, and the indirect effect on the prices of properties that were not damaged by Sandy. To probe the (strong) assumptions required by difference-in-difference estimation we also report estimates for a more flexible specification that allows for endogenous breakpoints in the trajectories of housing values.

We find robust evidence of a negative and persistent impact on the price trajectories of houses that were affected by Sandy. Typically, properties that suffered more damage experienced a larger drop in value. But, interestingly, housing values fell even for properties that were not damaged but face high risk of coastal flooding. As we discuss in detail later on, this observation supports the interpretation of an increase in the perceived risk of living in areas at risk of flooding. As we discuss in more detail later, this finding is reminiscent of the results in [Abadie and Dermisi \(2008\)](#) on the effects of the 9/11 terrorist attacks on commercial locations near landmark buildings in Chicago, which were not directly damaged during the attacks.

We provide additional confirmation for our main findings by estimating property-fixed effects models and flexible specifications that allow for endogenous breakpoints in price trends. Additionally, the latter set of estimates traces the response of property prices over time at a quarterly frequency. These estimates provide several important results. First, the price trajectories of properties affected by Sandy were parallel to those of properties in unaffected areas until the arrival of the hurricane. Second, it is precisely in the quarter immediately after Sandy that property prices in the affected areas deviated sharply from housing prices elsewhere. Finally, we also document that in the years following the hurricane, housing prices in the affected areas partially recovered but have remained persistently

---

<sup>4</sup>By its nature, property sales data are not ideal to estimate property-fixed effects models. The reason is that only a minority of properties are sold more than once within our sample period. We discuss this issue at length later in the paper.

depressed. This finding is reminiscent of the results in [Ambrus et al. \(2016\)](#), which documents extreme persistence of negative shocks to the demand for housing at the neighborhood level.

[Section 2](#) presents a brief review of the related literature. [Section 3](#) describes the main data sources and the merging process. [Section 4](#) presents descriptive statistics on the main variables. [Section 5](#), [Section 6](#) and [Section 7](#) present the main results. [Section 8](#) concludes. All Tables and Figures are gathered at the end of the paper.

## 2 Literature

Since the pioneering work of [Oates \(1969\)](#), economists have been aware of the connection between local characteristics, such as local property tax rates, and housing (land) values. Methodologically, our work is more closely related to the literature on the determinants of housing prices within a city. This literature has focused on how some neighborhood characteristics capitalize in the price of housing. Chiefly among these, several authors have focused on the role of school quality (e.g. [Black \(1999\)](#), [Fack and Grenet \(2010\)](#) or [Schwartz et al. \(2014\)](#)). But the impacts of other neighborhood characteristics have also been analyzed, ranging from crime ([Thaler \(1978\)](#) and [Schwartz et al. \(2003\)](#)) to immigrant density ([Saiz and Wachter \(2011\)](#)).

Our study is also related to recent work that analyzes the effects of large geographically concentrated shocks on housing values using plot-level data. [Hornbeck and Keniston \(2014\)](#) study the aftermath of the 1872 Great Boston Fire using a longitudinal dataset of (assessed) housing values linked it to the exact burned area. They argue that the large increases in property values that followed the fire were due to the (well used) opportunity to redevelop the zone, breaking away from existing inefficient inertia. Their identification is based on the differential trends in housing values between units in the burned and unburned areas, allowing also the impacts to vary by distance to the burned area. In their recent study, [Ambrus et al. \(2016\)](#) exploit data for a cholera outbreak in a neighborhood in London in the 19th century. These authors also build a panel for housing prices at the tax lot level over a long period of time (1853-2015), and match it to housing maps and to the the number of deaths by house. They find that housing prices fell significantly in the affected area, with a permanent reduction in values of about 37%. They argue that the negative income shock led to a worsening of the composition of the neighborhood in terms of income, which locked in due to externalities.

Another strand of literature that is related to our work investigates the economic effects

of climate change. Several studies have analyzed the effects of hurricane Katrina. [McIntosh \(2008\)](#) examined the effects of Katrina-related migration of evacuees on the Houston metropolitan area labor market. Using monthly CPS data, she found that wages and unemployment among Houston locals were significantly negatively affected by the migration of Katrina evacuees. [Deryugina et al. \(2014\)](#) use data on individual tax returns to analyze the long-term economic effects of Katrina on the population of New Orleans prior to the hurricane using propensity score matching. These authors find evidence of persistent geographical displacement, but only transitory effects on income and employment.<sup>5</sup> [Groen et al. \(2015\)](#) provide estimates of the effects of hurricanes Katrina and Rita on employment and earnings. Using individual panel data these authors also find evidence of a temporary reduction in income, followed by prolonged increase in earnings due to the increased labor demand in sectors related to rebuilding. Besides the empirical studies noted above, there have also been important theoretical advances. [Gröger and Zylberberg \(2016\)](#) analyze coping mechanisms through labor in-migration of rural economies after catastrophic natural disaster. Specifically, they focus on the impact of typhoon Ketsana in Vietnam during 2009. These authors find that internal remittances from long-distance labor migrants is the most important coping mechanism. [Zhang \(2016\)](#) provides evidence that being located on the floodplain has a negative impact on the price of house sales using data for the metropolitan area of Fargo-Moorhead. This study also reports that, following a major flood in 2009, home prices fell temporarily but quickly recovered, but his study does not employ property-level damage estimates. [Desmet and Rossi-Hansberg \(2015\)](#) develop a dynamic spatial theoretical model of trade to analyze the effects of climate change. In this model innovation and diffusion are endogenous outcomes, and carbon emissions are a by-product of production. [Desmet et al. \(2015\)](#) go on to extend the previous framework by endogenizing migration, and use the model to simulate the effects of a rise in sea level.

Finally, our paper is also related to the empirical work on estimating the economic effects of terrorism. [Abadie and Gardeazabal \(2003\)](#) analyze the economic consequences of terrorism in the Basque Country (Spain) using synthetic control analysis, as well as a temporary truce as a natural experiment. Their findings reveal a large cost of terrorism. The work by [Abadie and Dermisi \(2008\)](#) is more closely related to ours in terms of methodology. The goal of their paper is to estimate the effect of the 9/11 attacks on the demand for downtown office space in Chicago. Even though unaffected by the attacks, the authors

---

<sup>5</sup>[Deryugina \(2015\)](#) studies the role of government transfer programs, such as unemployment insurance, and shows that the relief they provide is at least as large as that coming from emergency aid.

argue that there may have been an increase in the perception of risk and insecurity. Using panel data on vacancy rates for office buildings in Chicago, the authors produce difference-in-difference estimates where identification is based on the comparison in the trajectories of vacancy rates in or around landmark buildings relative to more distant buildings. They find that vacancy rates increased as a result of the terrorist attacks.

## 3 Data

### 3.1 Housing prices

Our main outcome variable is the sale price of a housing unit. Our data on housing prices is based on the universe of transactions taken place in New York City between years 2003 and 2015, collected and made available by the NYC Department of Finance.<sup>6</sup> By their nature, transactions-based datasets are very sparse because most housing units only appear once in our dataset.<sup>7</sup>

The original data are organized as a collection of files by borough and year. For each of the borough-year datasets we create an identifier for each transaction by borough-block-lot-apartment number, which we refer to as BBLA.<sup>8</sup> Besides sale price, the dataset also contains some additional information: building class, tax class, and the exact date the sale took place. We merge the data for all years and boroughs and drop commercial and industrial properties (on the basis of the tax class) from the sample. Thus all our observations refer to housing units – apartments or houses – that are considered residential property. Building class category also provides useful information since it allows us to classify housing units into: houses (one, two, or three-family) and apartments (Coops or Condos).

We do some minimal trimming of the data. Specifically, we eliminate units with a sale price below \$10,000 or above \$15,000,000.<sup>9</sup> We also drop housing units that are sold 10 or more times during the 12-year period covered by our dataset. The resulting dataset contains 758,664 year-BBLA observations. The majority of housing units (identified by their BBLA) appear only once (75%), 19% appear exactly twice and 4% appear three

---

<sup>6</sup>These are the data used by [Schwartz et al. \(2014\)](#).

<sup>7</sup>In fact, the majority of units do not appear in any given year because they were not sold in that year.

<sup>8</sup>For housing units in apartment buildings (Coops or Condos) or multi-family houses, we naturally use their apartment number. For the majority of one-family houses, the BBL (borough, block and lot) identify each house (and we set the apartment number to one). However, there are a few instances of several one-family houses that share the exact same BBL. This is the case, for instance, for houses that are part of a cooperative, such as the Breezy Point Cooperative in Breezy Point, Queens.

<sup>9</sup>In the raw data we had 1,065,325 observations. These two sample restrictions reduce the sample size by roughly 250,000 observations. Many of these consist of garages or other small constructions inside a lot.

times. About 32% of sales involved housing units in Queens, 26% in Manhattan, 25% in Brooklyn, 9% in Staten Island, and 8% in the Bronx. The most common building classes in terms of sales were cooperative apartments (23%), one-family homes (21%), two-family homes (18%), condo apartments (18%), and three-family homes (4%). [Figure 1](#) reports the count of annual transactions by building class and borough. Sales of apartments (Coops or Condos) take place mostly in Manhattan, whereas Queens and Brooklyn account for most of the sales of 1-family and 2-family houses, respectively.

Turning now to sale prices, [Table 1](#) reports median and 10-90 percentiles for each borough.<sup>10</sup> As expected, median prices are the highest in Manhattan (\$773,870), followed by Brooklyn (\$484,319), Queens (\$391,400), Staten Island (\$375,000) and the Bronx (\$350,000). However, because housing units in more expensive boroughs tend to be smaller, these figures understate the true differences in the cost of housing.<sup>11</sup> [Figure 3](#) displays median sale prices by year for Coop apartments and [Figure 4](#) reports median sale prices for 1-family houses (for all boroughs except Manhattan). We note that up until 2007 median prices were equal in Brooklyn and Queens but, since then, prices in Brooklyn have risen faster than in Queens. It is also interesting that the housing bust of 2007-2009 was much more pronounced for 1-family houses than for apartments.<sup>12</sup>

Later in the paper we will also make use of data on administratively assessed housing values (for 1-to-3 family houses), produced also by the New York City Department of Finance. These data are available for each tax lot and year, giving rise to a balanced panel. In order to predict market values for all units and years, a spatial hedonic regression model is estimated on the transactions data.<sup>13</sup> We will describe that data in more detail later on.

## 3.2 FEMA data

To measure the damage caused by hurricane Sandy in New York we rely on two datasets from FEMA. The first is the building point-damage determination estimates. These data combine aerial imagery with damage caused by inundation and assessed in the field by FEMA's Modeling Task Force.<sup>14</sup> This dataset contains damage estimates for each of the

---

<sup>10</sup>These data pool all years and have not been adjusted for inflation.

<sup>11</sup>The fraction of sales pertaining to coops and condos in each borough are as follows: 94% in Manhattan, 28% in the Bronx, 34% in Brooklyn, 32% in Queens and 5% in Staten Island.

<sup>12</sup>[Figure 2](#) reports median prices by borough and building type.

<sup>13</sup>For more details, see [Quintos \(2014\)](#).

<sup>14</sup>The Modeling Task Force is a group of experts specialized in impact assessments for earthquakes, hurricanes, and other natural disasters. This task force plays an important role in developing best estimates of the impacts before, during and after the events. Specifically, during Hurricane Sandy the Modeling Task Force coordinated with the U.S. Geological Survey to deploy surge sensors and field teams to obtain surge

almost 320,000 buildings in the Sandy inundation zone and includes over 15,000 points outside that zone for which aerial imagery damage determinations were made.<sup>15</sup>

In this dataset each building point is identified by its longitude and latitude. In addition the dataset contain a number of variables that quantify the damage suffered by each property due to Sandy. The most important variable for our purposes (*DMG\_COMBO*), which provides the most comprehensive assessment of the damaged caused by hurricane Sandy. According to FEMA, this is the preferred measure of damage for several reasons. First, in inundation events with short duration, like Sandy, it is important to complement aerial imagery with observed inundation depths for each structure. Second, these damage-determination data provide more comprehensive estimates than simply the households that applied for FEMA Individual Assistance since households that did not apply for assistance are also included here. In addition, all building occupancy types (residential, commercial, industrial, and so on) are also included.<sup>16</sup>

The combined damaged variable takes four values: affected (1), minor damage (2), major damage (3) or destroyed (4).<sup>17</sup> **Table 2** reports the classification of damage levels across buildings in the overall Sandy inundation zone and in the subset that is located within New York City. Column 1 reports damage levels for the buildings in the whole Sandy inundation area (close to 319,000 observations), with almost 7 percent of all buildings having suffered major damage. Column 2 reports on the points of the Sandy inundation zone located in New York City. Over 13 percent of all buildings in this area suffered major damage. Columns 3 through 7 report the damage distributions for each of the five boroughs. Focusing on the category of major damage, Staten Island and Queens were the boroughs that were hit the hardest, with 26 and 17 percent of the buildings having suffered major damage, followed by Brooklyn (8 percent). The Bronx and Manhattan were the boroughs for which major

---

assessments.

<sup>15</sup>Where available, the aerial imagery overrules the inundation-based damage assessment. In particular, “destroyed” determinations were only based on imagery.

<sup>16</sup>It is worth noting that the damage determination data contain only buildings that have been affected by Sandy to some degree. Some of these were initiated by individual applications for FEMA assistance (followed by field verifications) and others were included on the basis of aerial imagery of affected buildings. One may be concerned that this may potentially introduce some selection bias. To address this point we will also use an alternative, possibly more exogenous measure of damage based purely on the storm surge.

<sup>17</sup>For example, a building is declared to have suffered *major damage* if aerial imagery showed that more than 20 percent of the roof diaphragm was destroyed and some exterior walls collapsed. In terms of the inundation assessment, a classification of major damage requires a field verified flood depth greater than 5 feet. Our understanding is that when either of these conditions is met the property is considered to have suffered *major damage*. In comparison, a property is considered *destroyed* only if aerial imagery revealed that the majority of the exterior walls collapsed. For further details on the exact definition of the FEMA damage classification, visit <http://www.arcgis.com/home/item.html?id=307dd522499d4a44a33d7296a5da5ea0>.

damage was much less prevalent (2.40 and 0.31 percent, respectively).<sup>18</sup>

The second FEMA dataset that we rely on is the data on hurricane Sandy’s storm surge.<sup>19</sup> These data provide the geographic boundary of the area that got flooded during hurricane Sandy, mainly across the coastlines of New York and New Jersey, at a very high geographic resolution.<sup>20</sup> In addition the data report the level of flooding at each point (coded in a variable named DEPTH). These data are a crucial input into the point-damage determination estimates presented earlier, inducing high correlation between the two measures, as we illustrate later. Our interest in the storm surge data is that it allows us to build measures of the affectation due to Sandy that are not affected by idiosyncratic differences across properties in the level of preparedness for the storm.

### 3.3 Hurricane Evacuation Zones

Besides data from FEMA, we also rely heavily on the hurricane evacuation zones defined by New York City’s Emergency Management department and represent varying threat levels of coastal flooding. Specifically, the city is subdivided in 6 evacuation zones, with zone 1 being the one with the highest risk and zone 6 the one with the lowest risk.<sup>21</sup> We obtained the longitude-latitude points that define each of the evacuation zones and merge them into our dataset.

In our analysis we focus on evacuation zones 1 and 2, and define the set of buildings in these zones as HEZ12.<sup>22</sup> Table 12 in the Appendix reports the distribution of tax lots (BBLs) across hurricane evacuation zones in New York City. About 4.25 percent of the tax lots are located in HEZ1. Combined, HEZ12 accounts for 7.76 percent of the city’s tax lots. Combining the six evacuation zones accounts for 47 percent of the tax lots in the city. By boroughs, the three with the highest share of tax lots in HEZ12 are Staten Island (10.8%), Brooklyn (9.4%) and Queens (6.75%). The lowest shares are in Manhattan (5.6%) and the Bronx (3.3%).

---

<sup>18</sup>We note that these percents do not refer to all buildings in the city or borough but, rather, only to the buildings that were part of the inundation zone.

<sup>19</sup>According to NOAA (the National Oceanic and Atmospheric Administration), *storm surge* is the abnormal rise in of water generated by a storm, over and above the predicted astronomical tide. The storm surge needs to be distinguished from the *storm tide*, which is defined as the water level rise during a storm due to the combination of storm surge and the astronomical tide. Hurricane Sandy hit New York City at high tide, leading to a larger storm tide than if it had hit at a different time of the day.

<sup>20</sup>The raw storm surge data contain 350,154 latitude-longitude observations covering the 5 boroughs of the city.

<sup>21</sup>The map with the hurricane evacuation zones can be found at [http://www.nyc.gov/html/oem/downloads/pdf/hurricane\\_map\\_english.pdf](http://www.nyc.gov/html/oem/downloads/pdf/hurricane_map_english.pdf). The evacuation zones data that we obtained are accurate as of the 2010 Census, two years prior to hurricane Sandy.

<sup>22</sup>Analogously, we define HEZ1 as the set of properties in hurricane evacuation zone 1.

### 3.4 Merging Process

Combining the previous datasets in a way that was conducive to econometric analysis presented an important challenge. Each dataset used a different system of geographic coordinates. The housing dataset identified observations by exact address and tax lot identifiers; FEMA data employ spherical latitude and longitude; and the hurricane evacuation zones (HEZ) were geocoded using the cartesian approximation for New York State. The difficulty was to map all these data into a geographical structure that was convenient for our analysis.

Our strategy consisted in mapping the FEMA and HEZ datasets into tax lots, which could then be merged with the housing data. To do so we used an additional dataset as cross-walk. This dataset is called PLUTO and is a compilation of variables maintained by different New York City agencies that contains a wealth of information.<sup>23</sup> Crucial for our purposes, it contains shape files for the city’s tax lots (parcels). Thus we were able to map each of the points in the HEZ and FEMA datasets into the corresponding polygons of the tax lots.<sup>24</sup> We refer to this dataset as FEMA-HEZ, which contains the hurricane evacuation zone and the extent of damage for all tax lots in New York City.<sup>25</sup> As we describe in detail in the Appendix, the accuracy of this merge was extremely high. Furthermore, PLUTO identifies each parcel polygon by its center-point coordinates (based on the New York State plane approximation) along with its borough, block and lot (or BBL), which allows us to match the FEMA-HEZ data with the housing dataset by BBL.<sup>26</sup> The final transaction sales dataset has 578,500 BBL-Apartment-Year observations, where the apartment number has been set equal to one for 1-family houses.

## 4 Descriptive statistics

We begin by verifying that the final dataset retains the key features of the original data in terms of differences across boroughs in average housing prices and average damage inflicted

---

<sup>23</sup>PLUTO contains information on over 857,000 tax lots, corresponding to three types of data: tax lot characteristics, building characteristics, and district-level data. In PLUTO all apartments belonging to the same Coop will display the exact same information (e.g. year built) because they belong to the same tax lot (BBL). Unlike other city datasets, in PLUTO all Condo apartments in the same building appear under a common tax lot and thus are treated symmetrically to coop apartments.

<sup>24</sup>These steps were implemented using function *spTransform* in the RGDAL package in *R* (Bivand et al. (2015)).

<sup>25</sup>We note that the variables in this dataset (storm surge, damage determination points, and hurricane evacuation zones) do not vary over time.

<sup>26</sup>This way we were able to avoid string matching based on the address field, which is a highly inaccurate and tedious process. Note also that the housing dataset does have time-varying variables, such as the sale price.

by Sandy. [Table 3](#) reports these data. First, Manhattan remains as the borough with the highest median sale prices (634,523 dollars), followed by Brooklyn (491,553 dollars), Queens (400,680 dollars), the Bronx (374,869 dollars), and Staten Island (391,867 dollars). In comparison the median sale price (across all years and boroughs) for New York City is 430,000 dollars. The ranking and levels of prices are consistent with what we reported in [Table 1](#).<sup>27</sup> Next, we focus on the share of properties in each borough that are located in hurricane evacuation zones 1 or 2 in our combined, transactions-based dataset. Staten Island is the borough with the highest share of properties at the highest risk of evacuation (12.32 percent), followed by Brooklyn (9.16 percent), Queens (6.10 percent), and at a large distance behind, Manhattan (3.39 percent) and the Bronx (1.93 percent).<sup>28</sup> These figures match fairly well with those obtained from the FEMA-PLUTO data (reported in [Table 12](#) in the Appendix), with the main deviation referring to Manhattan (3.4% versus 5.6 %) due again to the missing condo apartments in our combined dataset. Finally, we turn to average damage levels caused by hurricane Sandy. In our combined dataset, Brooklyn and Queens are the boroughs that suffered the most damage. Respectively, 0.84 and 0.76 percent of all properties in these boroughs suffered major damage or were destroyed, compared to a city-wide average of 0.50 percent. Although the figures are not directly comparable because of the different denominators (citywide versus inundation zone), the ranking is consistent with the high levels of damage in these boroughs reported in the original FEMA data ([Table 2](#)). There is, however, one important difference between the two tables. Namely, Staten Island was the borough that suffered the most damage yet it is entirely missing from our final dataset. The reason for this is that none of the damaged properties in Staten Island have been sold within the period 2003-2015. Consequently, Staten Island will not play any role in the identification of our coefficients of interest.<sup>29</sup>

Next we provide descriptive statistics on our estimation sample ([Table 4](#)), which contains 578,500 (lot-apartment-year) observations.<sup>30</sup> Besides the sale price, we build indicators for being located in hurricane zones. Among these, we focus on (*HEZ12*), the indicator for

---

<sup>27</sup>The only significant deviation is the reduction in the median price for Manhattan, which is due to the fact that our combined dataset excludes Condo apartments that could not be merged due to a recoding in the PLUTO dataset.

<sup>28</sup>For New York City as a whole, 6.70 percent of observations in our final dataset are located in hurricane evacuation zones 1 and 2.

<sup>29</sup>This is unfortunate given that the largest variation in the data is probably found in the properties located in this borough. Nonetheless, Staten Island properties do appear in our estimation sample when we define the treatment variables on the basis of storm surge, or when we use the data on assessed market values.

<sup>30</sup>Recall that 1-family houses have been assigned apartment number 1.

being in hurricane evacuation zone 1 or 2, which takes a value of one for 6.7 percent of observations.<sup>31</sup> We view set *HEZ12* as containing all units that are subject to high risk of coastal flooding in New York City in the event of a hurricane. Because this classification is publicly known and well publicized by city agencies, it is the most relevant partition in terms of perceived risk of flooding. In our view all properties located in *HEZ12* were affected by hurricane Sandy in one way or another. Some were flooded and directly damaged. Others were spared but were able to observe first-hand the damaged caused by the hurricane to their neighborhood, which may have affected their preferences for living in that location.

Next, we define indicators for the level of damage suffered during FEMA.<sup>32</sup> These indicators effectively partition the set of units located in evacuation zones 1 and 2. Specifically, we define *Dam0* as the indicator for units in *HEZ12* that were not damaged by hurricane Sandy according to FEMA’s estimates. *Dam1* is an indicator function for units in *HEZ12* that suffered at most minor damage during Sandy while *Dam2* is the indicator for the units in *HEZ12* that suffered major damage or were destroyed. The relative frequency of these categories in our dataset is 2.5, 3.7, and 0.5 percent, respectively. That is, only a very small share of the properties affected by Sandy suffered major damage (and even fewer were destroyed). We also build alternative measures of damage that are purely based on FEMA’s storm surge data.<sup>33</sup> Again we define three indicators that partition the set *HEZ12*. *Sur0* is an indicator for those units in *HEZ12* that were not flooded. In turn *Sur1* and *Sur2* are indicators for being in *HEZ12* and having registered flooding below or above 5.48 feet, respectively.<sup>34</sup>

**Table 4** also reports the distribution of observations over building classes, where we distinguish between 1-, 2-, and 3-family homes, Coop apartments, Condos, and apartment buildings devoted to rental. By far the three most important categories are 1-family homes (29 percent), 2-family homes (25 percent), and Coop apartments (33 percent), which combined amount to 87 percent of all observations in our dataset. The table ends with some important control variables: (gross) square footage, which is only meaningful for houses, year built, and the last two years when a property was altered, according to city records.<sup>35</sup>

---

<sup>31</sup>In comparison, 3.7 percent are in hurricane evacuation zone 1 and 33 percent are located in one of the six evacuation zones.

<sup>32</sup>These indicators are based on the variable *DMG\_COMBO* in the FEMA building-point-damage determination estimates (described in **Table 2**).

<sup>33</sup>These variables are based on the variable *Depth* in the FEMA dataset for hurricane Sandy’s storm surge.

<sup>34</sup>This cutoff is the 90th percentile of the *depth* variable, conditional on positive values and was chosen so that the distribution of the surge-based indicators roughly resembles that of the damage-based indicators.

<sup>35</sup>For apartment buildings square footage refers to the building rather than to the apartment. Gross square footage is the sum of the surface of all construction in a lot, including basements, higher floors, and

In the analysis we combined year built and year altered into a single variable that replaces the year built by the most recent year of alteration. We then build categories for this variable and include them as dummy variables in our regression models.

## 5 Main Results

### 5.1 Specification

Consider an observation  $(i, z, t)$ , where  $i$  refers to an individual house or apartment (uniquely defined by borough-block-lot-apartment),  $z$  to the neighborhood (e.g. borough, zip code or block), and  $t$  to the quarter or year.

In essence we want to compare the price trajectories of three sets of housing units, before and after hurricane Sandy, to each other and to housing units that were not affected by Sandy in any way. On the basis of the previous definitions, we view all units in hurricane evacuation zones 1 and 2 as having been affected by Sandy with different levels of intensity. Some units were directly affected because they were damaged to some degree by Sandy (identified by indicators  $Dam1$  and  $Dam2$ ), while other units may have been only indirectly affected through an increase in the perceived risk of living in those units ( $Dam0$ ), similar to the effects emphasized in [Abadie and Dermisi \(2008\)](#).<sup>36</sup> In contrast we view housing units located outside of hurricane evacuation zones 1 and 2 as not having been affected at all by Sandy.<sup>37</sup>

Estimation of these effects lends itself nicely to a difference-in-difference estimator:

$$\begin{aligned} \ln p_{izt} = & \alpha_z + \alpha_t + \gamma_0 Dam0_i + \gamma_1 Dam1_i + \gamma_2 Dam2_i \\ & + Post_t \times (\beta_0 Dam0_i + \beta_1 Dam1_i + \beta_2 Dam2_i) \\ & + \gamma' X_{i,z} + \varepsilon_{izt}, \end{aligned} \tag{1}$$

where  $\alpha_z$  denotes neighborhood fixed-effects that will absorb all time-invariant differences in prices across neighborhoods and  $\alpha_t$  will denote quarter-year dummy variables.<sup>38</sup> Next, indicator variables  $Dam0_i$ ,  $Dam1_i$ , and  $Dam2_i$  denote the level of damage caused by Sandy, additional structures.

<sup>36</sup>The interpretation is analogous for the measures based on the depth of storm surge.

<sup>37</sup>In the language of experiments, HEZ12 is the “treatment group” and the rest of the city would be the “control group”. A very small number of properties located outside of HEZ12 experienced damaged. We excluded them from the regressions in order to have a ‘cleaner’ control group. This exclusion restriction resulted in a sample reduction of about 6,000 sales. It is likely that most of these units were damaged by fallen trees.

<sup>38</sup> $X_{i,z}$  denotes property-specific controls, such as year built or last altered.

as defined earlier. The coefficients accompanying these dummy variables will capture pre-Sandy differences in housing prices between the three groups of housing units in the hurricane evacuation zone.<sup>39</sup> Importantly, we expect the coefficients associated to these indicators to be close to zero, which would indicate that there were no systematic price differences between observationally equivalent housing units prior to Sandy. This is more likely to be the case when we employ geographically narrow definitions of neighborhoods.<sup>40</sup> The most important coefficients for our purposes are the interaction terms between the post-Sandy indicator and the indicators for the level of damage.<sup>41</sup> It is natural to expect these coefficients to be negative (or zero) and to be ordered in the following manner:  $\beta_2 \leq \beta_1 \leq \beta_0 \leq 0$ . Regarding the stochastic specification of the model, we believe it is plausible to assume that the error terms are uncorrelated across neighborhoods, but likely to be correlated across individual units within a neighborhood and over time. Consequently, we will report standard errors clustered at the neighborhood level.<sup>42</sup>

An important limitation of our dataset is that we do not have information on the post-Sandy reforms at the property level. Suppose that houses that suffered moderate levels of damage (identified by indicator *Dam1*) were repaired by the owners and brought back to their pre-Sandy condition. In this case we would expect  $\beta_1$  to equal  $\beta_0$ . Thus, to the extent, that the decisions to repair damaged properties vary systematically as a function of the level of damage, the point estimates for the interaction terms may not be ordered monotonically. We note that these considerations may be less relevant for properties that were not directly damaged by Sandy (identified by indicator *Dam0*), since there is no reason to expect their owners to have carried out post-Sandy adaptation investments at differential rates from other owners. If so,  $\beta_1$  and  $\beta_2$  may underestimate the true effect of Sandy on property prices, whereas this may not be the case for  $\beta_0$ .

## 5.2 A word on identification

To understand where the identification of our estimates is stemming from, it is helpful to think about how one would go about constructing the estimates manually. Intuitively, these

---

<sup>39</sup>Note also that the excluded category in this specification are all housing units in New York City that were outside of the hurricane evacuation zones 1 and 2. We will provide sensitivity analysis along this dimension.

<sup>40</sup>The interpretation of the coefficients for the regression models including the surge-based indicators is analogous.

<sup>41</sup>Hurricane Sandy hit New York City on October 29, 2012. Hence, the first quarter that may display an effect is the first quarter of 2013. We define the *Post* indicator as taking a value of one if the house was sold in November 2012 or later.

<sup>42</sup>In our preferred specifications we cluster standard errors by city block.

are the steps we would follow if we simply considered two periods (before and after). First, we would calculate the average sale price prior to hurricane Sandy for each neighborhood. Second, we would calculate the average sale price in each neighborhood in the period after Sandy. Finally, we would compare the change in the average price between neighborhoods that were affected by Sandy (e.g. located in *HEZ12*) to the average price increase in unaffected neighborhoods.<sup>43</sup>

Clearly, our method is immune to time-invariant differences across neighborhoods. However, we need to assume that the properties sold before and after Sandy in a given neighborhood are comparable. In other words, our specification does not include property fixed-effects. This assumption is more plausible when one uses a narrow definition of neighborhood, such as a city block. The reason is that individual properties within a block tend to be fairly homogeneous in terms of construction type and quality. Nevertheless, we are still concerned that Sandy might have induced selection into the sample of properties being sold. For example, if only the better properties in the affected areas are sold after Sandy (conditional on the observed levels of damage), our difference-in-difference estimates will be upward biased, understating the true price reductions due to the hurricane.<sup>44</sup>

Naturally, these concerns would be avoided altogether if we observed the same properties before and after Sandy since in that case we would be able to compute within-unit price changes that would not be affected by compositional changes in the set of properties sold in the market. [Section 6](#) will address the issue of selection by estimating property-fixed effect models and by directly examining the effects of Sandy on the composition of sales based on observable characteristics.

### 5.3 Results based on point-damage estimates

The main source of damage in the aftermath of hurricane Sandy was caused by Sandy's unusually powerful storm surge and the ensuing coastal flooding. As a result, the main predictors of the level of damage were proximity to the shore, building elevation, construction type, and building age.<sup>45</sup> As a result, older 1-family houses suffered the most severe structural damage. Specifically, these buildings accounted for only 18 percent of the buildings in Sandy's inundation zone. However, they accounted for 73 percent of all damaged buildings.

---

<sup>43</sup>This identification strategy is similar to that used in Abadie and Dermisi (2008) who compare the changes in vacancy rates before and after the 9/11 attacks for office buildings on the basis of their distance to landmark buildings in downtown Chicago.

<sup>44</sup>We will refer to this scenario as *positive selection*. In contrast, if only the worst properties are sold after Sandy, our sample will be affected by *negative selection* and our estimates will be biased downward.

<sup>45</sup>Older buildings were typically subject to less demanding construction codes.

In our regressions we will take into account the year of construction (and later alterations) for each building, as well as building type and distance to the shore and other bodies of water, which are used to define the hurricane evacuation zones.

We will present two sets of estimates that differ on the measures used to gauge the impact of Sandy. The first set of estimates is based on FEMA’s building-point damage estimates, while the second is based purely on the level of flooding due to Sandy’s storm surge (also collected by FEMA). The latter is perhaps a simpler measure, but the former is more comprehensive as it also incorporates data from aerial imagery taken immediately after the hurricane and is explicitly favored by FEMA as the most comprehensive measure of the damage caused by hurricane Sandy.<sup>46</sup>

**Table 5** reports the estimates. In all specifications the dependent variable is the log of the sale price and quarter-year dummies are included (but not displayed), along with categories for year built or last altered. Columns 1-3 report estimates based on the whole sample (571,940 observations), which includes both houses and apartments.<sup>47</sup> As we move from column 1 to 3 we employ increasingly narrower definitions of neighborhood, moving from borough (5), to zip code (182), and city block (24,029). Let us first focus on the point estimates for indicator variables  $Dam0$ ,  $Dam1$  and  $Dam2$ . Note that these estimates gradually become closer to one another. Indeed in column 3 we cannot reject the null hypothesis that they are all equal to each other and to zero. Next, we turn to our main coefficients of interest, the interaction terms between the damage indicators and the post-Sandy indicator. We find negative point estimates across all columns. In column 3, these point estimates are all large, significantly different from zero and roughly increasing (in absolute value) with the intensity of the damage caused by Sandy, ranging between 10 and 21 log points. These estimates suggest that hurricane Sandy affected negatively the price trajectories of the affected housing units.

Let us now turn to columns 4-6, which restrict to the subsample of 1-family and 2-family houses (165,642 observations) and include the log of gross square footage as a control. In this more homogeneous sample (column 4), the estimates are very similar to those in column 3. In terms of magnitude, the estimates in column 4 – our preferred specification – imply an 11 log point price penalty for housing units that were located in HEZ12 that were not directly damaged by Sandy, or only lightly. Houses that suffered major damage lost approximately

---

<sup>46</sup>The two sets of measures are strongly correlated. The pairwise correlation coefficients for  $(Dam0, Sur0)$ ,  $(Dam1, Sur1)$ , and  $(Dam2, Sur2)$  are, respectively, 0.71, 0.82, and 0.46.

<sup>47</sup>The sample excludes damaged properties located outside of hurricane evacuation zones 1 and 2, which reduces the sample from 578,500 to 571,940 observations.

one quarter of their value. Compared to the previous column, in column 5 we restrict to housing units located within one of the six hurricane evacuation zones. The latter effectively changes the relevant comparison group from housing units outside the evacuation zones to housing units in evacuation zones with a lower (but non-zero) degree of threat of coastal flooding.<sup>48</sup> As expected, in this case the point estimates for the interaction terms diminish slightly, ranging from  $-0.20$  to  $-0.07$ , but nevertheless remain significant both statistically and economically. In column 6 we include again apartments in the sample, but still focus on units in one of the six evacuation zones, which does not significantly affect the point estimates.

We note that in the majority of the columns in the Table, the point estimate for the coefficients associated to  $post \times Dam0$  and  $post \times Dam1$  are statistically undistinguishable from each other, even though the latter properties suffered more damage. Our interpretation is that owners of moderately damaged properties invested resources in repairing them before putting them up for sale.<sup>49</sup>

#### 5.4 Results based on storm surge

As a robustness check, next we repeat the analysis using an alternative measure that is purely based on the storm surge data. Specifically, we measure the intensity of the effects of Sandy on each individual property on the basis of the feet of flooding at that specific point. In the same fashion as before, we partition the set of properties in HEZ12 into three groups using indicator variables  $Sur0$ ,  $Sur1$  and  $Sur2$ . Properties in the first group did not experience any flooding and properties in the latter group experienced a storm surge of more than 5.48 feet.

**Table 6** reports our results. The interpretation of the coefficients in these regressions is identical to that used for the damage-based regressions. The general pattern is similar to that of the previous table, although the changes in post-Sandy price trajectories are slightly smaller. Our preferred estimates are reported in column 4, which is based on the subsample of 1-family and 2-family houses and controls for square footage. These estimates suggest price penalties of 9, 12 and 17 log points, only slightly smaller than those obtained with the damaged-based estimates (11, 11 and 24 log points, respectively).<sup>50</sup> The lower point estimates obtained here can also be driven by measurement error. Some properties that

---

<sup>48</sup>Recall that square footage is only available for houses, but not for apartments.

<sup>49</sup>These renovations may have been financed with flood insurance payouts, FEMA aid, or by the owners' themselves.

<sup>50</sup>This time the point estimates are ordered monotonically, as we had anticipated earlier.

experienced only low levels of flooding could nonetheless have been damaged severely. If, say, this damage took the form of structural damage to the roof because of the high winds, the surge-based measures will not reflect it even though the value of the property is likely to have been affected by this event.

## 6 Selection

The main limitation of the estimates in the previous section arises from the fact that we are not able to account for all potentially relevant dimensions of property-level heterogeneity. As a result, we are concerned that properties sold before and after Sandy in the affected neighborhoods may differ systematically. For example, it could be the case that only the relatively better properties were sold after Sandy, inducing positive selection into sales.

We address the selection concerns in three ways. First, we estimate models including property-fixed effects (on the subsample of repeat sales and on a new dataset that contains assessed market values for all properties) and secondly, by analyzing whether Sandy affected the probability of being sold on the basis of observable characteristics, which will provide suggestive evidence regarding the degree of selection in unobservables.

Let us first discuss our property-fixed effects estimates. First, we estimate our model on the subsample of housing units that were sold repeatedly within our sample period 2003-2015. Naturally, this reduces our sample size substantially and introduces concerns of sample selection. On the other hand, it allows us to estimate the more demanding fixed-effects specification. The second set of estimates is based on a new dataset that measures housing values on the basis of administrative assessments of market values. These data provide a balanced panel for all houses (tax class one) and, once again, can be used to estimate models with property fixed-effects. The potential downside of these data is that they are the result of an imputation process conducted by the NYC Department of Finance, which may complicate inference.<sup>51</sup>

Our property fixed-effects specification is the following:

$$\ln p_{it} = \alpha_i + \alpha_t + Post_t \times (\beta_0 Dam0_i + \beta_1 Dam1_i + \beta_2 Dam2_i) + \varepsilon_{it}, \quad (2)$$

where  $\alpha_i$  are property fixed effects that remove time-invariant differences across properties, and  $\alpha_t$  are quarter-year dummies. We also include interaction terms for post-Sandy years

---

<sup>51</sup>The estimates of market values for residential properties are based on algorithms that match each property to recent comparable sales. Thus, ultimately, this dataset is also based on transactions data. The data on assessed market values is produced and made public by the NYC Department of Finance on an annual basis.

with indicator variables for the intensity of the damage it caused. Note that the specification does not include property-level controls. This is because the control variables we included earlier are essentially time-invariant (within our sample period) and therefore absorbed by the property fixed-effects. As for the stochastic specification of the model, we present standard errors clustered at the city block level.<sup>52</sup> That is, we are assuming that price shocks across city blocks are uncorrelated, but allow for arbitrary correlations across individual properties within the block and over time.

## 6.1 Repeat Sales

We create an indicator for whether a particular property (defined by borough-block-lot-apartment) was sold more than once in the period 2003-2015. We find that 43% of the properties were sold more than once, with the remaining 57% being sold just once during our sample period.<sup>53</sup>

We now are able to estimate the specification in [Equation \(2\)](#). The results are collected in [Table 7](#). The first column considers all units affected somehow by Sandy bundled together (*HEZ12*). We find an 11 log point price reduction in the period after hurricane Sandy. Column 2 disaggregates *HEZ12* by the level of damage caused by Sandy. The coefficients for the interaction terms imply price reductions of 14, 8 and 17 log points for properties that were undamaged, lightly damaged, or severely damaged, respectively. These coefficients are fairly similar to the ones reported in the previous section. Column 3 employs surge-induced flooding as a measure of how Sandy affected specific properties. The coefficients are very similar to those in column 2 for properties that were either not flooded or only moderately (i.e. less than 5.5 feet). Instead the coefficient for houses affected by the largest levels of flooding (identified by *Sur2*) is now  $-0.07$ , substantially lower (in absolute value) than the corresponding coefficient in column 2.<sup>54</sup> Columns 4 through 6 re-estimate these models but restricting now to the subsample of 1-family and 2-family houses. The point estimates are a bit lower and standard errors are higher, but the general pattern stays pretty much the unchanged. The only exception is that the coefficient associated to the interaction term for *Sur2* (in column 6) is no longer significantly different from zero, which may reflect the small

---

<sup>52</sup>While we can also cluster at the property level (Borough-Block-Lot), this would require assuming that price shocks are uncorrelated across individual properties within a neighborhood, which is implausible. As it turns out, clustering standard errors by block, as opposed to by property, is a more conservative choice that gives rise to larger standard errors.

<sup>53</sup>28.22% of all properties were sold exactly twice, 9.82% were sold exactly three times, and 5.35% were sold four times or more.

<sup>54</sup>A possible interpretation is that the damage caused by flooding levels above 5.5 feet may not be all that different from the damage caused by lower levels of flooding.

number of observations in this category. All in all these estimates are quite similar to the main estimates presented in the previous section, suggesting that selection on unobservables among the properties sold after Sandy has not been very pronounced.

## 6.2 Assessed Market Values

The property assessment roll database contains about one million observations per year. We append fiscal years 1999-2015 and restrict to tax class 1 (1-to-3 unit houses), which we proceed to match with the PLUTO-HEZ-FEMA dataset. The final dataset contains over 11 million property-year observations, corresponding to 658,000 properties. Out of these, 44 percent correspond to Queens and 29 percent to Brooklyn. For the remainder observations, Staten Island accounts for 17 percent, the Bronx for 9 percent, and Manhattan for only 1 percent.<sup>55</sup> [Table 8](#) describes the merged dataset. Several things stand out. First, the new dataset is a balanced panel and the sample size is almost 20 times larger than the transactions-based dataset, even though it excludes apartments. In our data 8.1 percent of the observations are located in hurricane evacuation zones 1 and 2, compared to 6.7 percent in the transactions-based dataset. The distribution of units across the three measures of damage and storm surge is similar to the one described earlier. The average assessed market value is 513,186 dollars, somewhat lower than the average of 639,217 dollars in the transactions data. This discrepancy is mostly due to the fact that Manhattan is a much smaller share of the assessed values dataset. In conclusion, despite its much larger size, the combined dataset based on assessed market values is highly consistent with our main (transactions-based) dataset.

We now turn to the property fixed-effects estimates based on these data. [Table 9](#) presents the estimates. Columns 1 and 2 report estimates of a model for the general treatment associated to being located within hurricane evacuation zones 1 and 2. The difference between the two columns is that the first one refers to a specification with fixed-effects at the block level, whereas column 2 includes property fixed-effects. In both cases the estimated coefficient is very similar, uncovering a price reduction due to Sandy of about 20 log points.<sup>56</sup>

Next, we disaggregate the effects of Sandy using our three damaged-based treatments,

---

<sup>55</sup>The small weight of Manhattan is due to the fact that we are restricting now to 1-to-3-family homes, leaving apartment buildings out of the sample.

<sup>56</sup>The fact that the point estimates are so similar in the two specifications (in columns 1 and 2) suggests that block fixed-effects are effective in incorporating a great deal of (time-invariant) unobserved heterogeneity. After all, it is often the case that most properties within a block were built in a similar fashion and within a brief period of time.

reported in columns 3-5. Column 3 includes all properties in our sample, that is, all 1-to-3-family houses in New York City. We find large and statistically significant effects, amounting to reductions in value of 19, 22 and 23 log points, respectively, for properties that suffered no damage, minor damage, and major damage. Column 4 restricts the estimation to the properties located in the six evacuation zones, which decreases the size of the point estimates only slightly (by 1-2 log points). Column 5 reports the same specification as column 1 but in this case we report standard errors that are clustered at the block level. Standard errors increase substantially, but the estimates remain highly significant. Columns 6-8 report estimates based on the depth of flooding suffered by each property. The estimates are very similar to those obtained with the damage-based measures. This time there’s a more marked difference between the effects on the prices of houses that were not flooded (that experienced a price drop of about 13 log points) and those of houses that were flooded (which lost about one quarter of their value).

These estimates are qualitatively similar to those presented earlier estimated on our transactions data using block fixed-effects and property fixed-effects, though with one notable difference in terms of magnitudes.<sup>57</sup> The gaps in the post-Sandy price trajectories in the affected zones are now very similar regardless of the degree of damaged suffered by the properties. We suspect this is due to the imputation method used by the NYC Department of Finance. As far as we know, the growth rate in observed sale prices in a neighborhood is used to impute market values for properties that were not transacted in that year. This procedure will introduce a greater degree of co-movement among the three sets of properties if the match for ‘comparable’ properties does not take into account the degree of damaged actually incurred by the properties during Sandy. As a result of this, these data may not be very useful to separately identify the disaggregated ‘treatment’ effects, but we think it is still useful to identify the less demanding single treatment consisting in being located within *HEZ12*. The estimates reported here suggest that housing values may have fallen by approximately 20% in the areas at risk of hurricane-induced flooding, and that selection into sales may not have been too severe.

### 6.3 Sale probability

As noted above, our baseline estimates were based on models that included neighborhood (city block) fixed-effects, but not property fixed-effects and, as a result, may suffer from

---

<sup>57</sup>The most comparable estimates are those presented in column 4 of [Table 6](#) for the subsample of 1-family and 2-family houses, which were 9, 12 and 18 log points.

selection bias if Sandy induced a change in the composition (in terms of unobservable characteristics) of the properties sold. Next, we examine whether Sandy led to a change in the composition of sales on the basis of the *observed* level of damage. While not conclusive, we believe that evidence of selection on observables would suggest that selection on unobservables may also be present. In this analysis we are able to estimate property fixed-effects models as well, which is also helpful in inferring the degree of self-selection based on unobservable characteristics.

Specifically, we merged the assessed values dataset, which contains all 1-to-3 residences (houses) in the city, with the transactions dataset, which contains information on all sales (including apartments) over our sample period. We then created an indicator for each house-year that takes a value of one if the house was sold in that year, and zero otherwise.

**Figure 5** plots the fraction of properties sold in the city in each year. Let us begin with the left panel. The solid line refers to the city as a whole. We clearly observe the end of the housing boom and the subsequent bust. At the peak, six percent of all properties in the city were sold in year 2004. At the trough, in 2011, only about 2.5 percent of all properties were sold. The dashed line reports the sales in the HEZ12 area. Up until 2011, the two lines are remarkably similar. But from 2012 onward their behavior diverges, suggesting that hurricane Sandy had an effect on the rate of sales. In 2012, Sandy's year, and 2013 the rate of sales was significantly lower in the area impacted by the hurricane. But in 2014 and 2015 the rate of sales was higher in this area, suggesting a compensation for the depressed sales activity of the previous two years. In fact, the post-Sandy average rate of sales is practically the same for the two sets of properties.<sup>58</sup>

We now ask whether the composition of the sales in the period after Sandy has varied, in terms of observed damage levels, relative to the pre-Sandy period. Turning now to the right panel in **Figure 5**, we observe that the post-Sandy recovery in sales has been more dramatic for the undamaged (Dam0) and the severely damaged properties (Dam2) in the HEZ12 area, than for the lightly damaged ones (Dam1).<sup>59</sup> This pattern does not reveal a marked change in selection. Rather, Sandy slowed down the rate of sales during 2013 and we observe an offsetting surge in sales in 2014 and 2015.

We explore this issue further using regression analysis. Specifically, we now estimate the following linear-probability-model specifications where the dependent variable takes a

---

<sup>58</sup>As we shall see, a consequence of this is that the difference-in-difference estimates presented later will not be very informative.

<sup>59</sup>We note that in the boom years prior to 2007 the rate of sales were higher for these two groups. Thus what we find now is a reversion toward that pattern.

value of one if property  $i$  was sold in year  $t$ :

$$Sold_{it} = \alpha_i + \alpha_t + \beta Post_t \times HEZ12_i + \varepsilon_{it} \quad (3)$$

$$Sold_{it} = \alpha_i + \alpha_t + Post_t \times (\beta_0 Dam0_i + \beta_1 Dam1_i + \beta_2 Dam2_i) + \varepsilon_{it}, \quad (4)$$

where we note that it is feasible to include property fixed-effects  $\alpha_i$  in the specification. In specifications where this is the case, the estimates will not be affected by selection bias into sales (due to unobservable characteristics), contrary to what may happen in specifications that only include block fixed-effects ( $\alpha_z$ ).

The results are presented in [Table 10](#). Columns 1 and 2 report estimates of specifications with block fixed effects, while columns 3-6 present estimates based on property fixed-effects models. Starting with column 1, we find a small positive coefficient associated to the interaction between the post-Sandy indicator and the HEZ12 indicators. This is in line with the offsetting effects noted in [Figure 5](#), resulting in a small positive net increase in sales in the after-Sandy period. When disaggregating by the level of damage, despite their low precision, the point estimates suggest increases in post-Sandy sales for properties that were undamaged or only lightly damaged, but a reduction for severely damaged properties. This pattern suggests that, if anything, the composition of sales in HEZ12 in the post-Sandy period shifted toward less damaged properties (positive selection). Columns 3 and 4 include property fixed-effects. The point estimates are barely affected and confirm the findings of columns 1 and 2. In turn, columns 5 and 6 report the property fixed-effects estimates but clustering standard errors at the property level, which is a less conservative approach. If we are willing to make this assumption, the estimates now provide (apparently) stronger evidence of positive selection in sales in the after-Sandy period.

In sum, our analysis for the probability of being sold suggests that Sandy had a noticeable effect on the timing of sales, but we are able to rule out large compositional effects. If anything, our estimates suggest a small degree of positive selection into sales in the after-Sandy period, which would imply that our baseline results provide lower bounds on the price reductions caused by Sandy. These conclusions are also largely consistent with the above sale price estimates based on property-fixed effects models.

## 7 Endogenous breakpoints and persistence

So far our results suggest that the price trajectories of housing units that were, directly or indirectly, affected by hurricane Sandy have exhibited downward deviations that are consistent with the timing of the hurricane. It is possible though that the change in trends

had started prior to the hurricane, which would cast doubts on a causal interpretation of our estimates. In order to examine this issue, we propose the estimation of a flexible specification that allows the data to endogenously determine the breakpoint in the price trajectories. Additionally, the results in this section will allow us to gauge whether the effects of the hurricane on prices were temporary or appear to be persistent in time. The answer to this question will be instrumental to narrowing down the interpretations that are consistent with our empirical findings.

Specifically, we consider two specifications:

$$\ln p_{izt} = \alpha_z + \alpha_t + \beta^t HEZ12_i + \gamma X_{iz} + \varepsilon_{izt} \quad (5)$$

$$\ln p_{izt} = \alpha_z + \alpha_t + \beta_0^t Dam0_i + \beta_1^t Dam1_i + \beta_2^t Dam2_i + \gamma' X_{iz} + \varepsilon_{izt}, \quad (6)$$

where the dependent variable is the log of the price of housing unit  $i$  in neighborhood  $z$  and quarter (or year)  $t$ . The right-hand-side contains neighborhood fixed-effects, year fixed-effects, and unit-specific controls (square footage and year built or last altered). [Equation \(5\)](#) includes an indicator for being located in the HEZ12, with a coefficient that is allowed to vary period by period.<sup>60</sup> Similarly, [Equation \(6\)](#) allows for time-varying coefficients for each of the treatment variables (indicators  $Dam0$ ,  $Dam1$  and  $Ddam2$ ).

The estimates will provide evidence in favor of the causal effect of hurricane Sandy if we cannot reject the hypotheses  $\beta_t = 0$  for  $t = 2003, \dots, 2012$ , but we find negative values for  $\beta_{2013}$ ,  $\beta_{2014}$  and  $\beta_{2015}$ . In words, assigning causality of the effects to hurricane Sandy will be supported if we find statistically significant evidence of a reduction in price for the affected units starting in year 2013 or, even more conclusively, in the first quarter of that year.<sup>61</sup>

## 7.1 Estimates

[Table 11](#) presents the estimates. The estimation sample is our transactions-based dataset, restricted to 1-family and 2-family houses. Column 1 provides estimates for the model in [Equation \(5\)](#). The coefficients associated to the interaction terms are all small and not significantly different from zero up until year 2012. In contrast, the point estimates for years 2013-2015 are large (in absolute value) and significant at the usual 5 percent significance levels. These point estimates suggest an initial 18 log point price reduction in year 2013, with a partial recovery amounting to an 8 log point reduction in year 2014, and stabilizing

<sup>60</sup>Specifically, we include a series of interactions between time dummies and an indicator for HEZ12.

<sup>61</sup>Recall that hurricane Sandy reached New York on October 29, 2012.

at an 11 log point drop in value in year 2015. This pattern is illustrated in [Figure 6](#), which plots the point estimates and the 95% confidence interval, and its quarterly version ([Figure 7](#)). The Figures clearly show that the price reduction for units in HEZ12 took place precisely in the first quarter of 2013, providing strong confirmation in favor of identifying hurricane Sandy as the event responsible for the structural break. Furthermore, we learn that prices initially over-reacted and fell by close to 20 log points in the year immediately after Sandy.<sup>62</sup> In the following years sale prices partially recovered but even in 2015 the Sandy price penalty appears to be around 10 log points.

To gain a deeper understanding, we turn to the estimation of [Equation \(6\)](#), reported in columns 2 through 4 of [Table 11](#). In all three columns we notice the same pattern: small and non-significant estimates for all years with the exception of the large and statistically significant estimates for the post-Sandy years 2013-2015. The estimates suggest a similar pattern for units located in HEZ12 that were undamaged (identified by *Dam0*) or lightly damaged (*Dam1*): a large initial drop of about 17 log points, followed by a partial recovery so that in 2015 the price penalty is around 10 log points. In contrast the price penalty for severely damaged properties (*Dam2*) is large and fairly constant over time at about 25 log points throughout the post-Sandy period in our sample (2013-2015).

## 8 Conclusion

Our analysis has provided robust evidence that hurricane Sandy led to an important, and persistent, reduction in prices in the affected neighborhoods. Properties that suffered more damage typically experienced larger drops in value. However, even those that were not directly damaged have suffered a persistent decline in their price.<sup>63</sup> We have also provided evidence suggesting that these findings do not seem to be driven by changes in the selection of properties being transacted in the post-Sandy period.

What are the mechanisms that may lie behind these findings? One possibility is that the price reductions that we have uncovered were driven by the 2012 and 2013 reforms to the National Flood Insurance program, which resulted in increases in flood insurance costs for many households located on the floodplain of New York City. While this may certainly have affected housing values (as suggested by [Checker \(2016\)](#) and [Dixon et al. \(2013\)](#)), the

---

<sup>62</sup>In the quarterly version of the regression we find that sale prices in HEZ12 fell by close to 40 log points in the quarter immediately after Sandy, although the standard errors associated to the point estimate are very large.

<sup>63</sup>In a similar vein, [Ambrus et al. \(2016\)](#) report extremely persistent effects of a localized negative shock on housing values.

timing of the main releases of relevant information regarding these reforms is at odds with our findings.<sup>64</sup>

In our view, the most likely interpretation for our findings is that hurricane Sandy led to an increase in the perceived risk associated to living by the ocean. This is similar to the explanation put forward by [Abadie and Dermisi \(2008\)](#) to account for the reduction in the demand for downtown office space in Chicago following the attacks of 9/11. This finding is in line with the views in [Hornbeck and Keniston \(2014\)](#). These authors argued that natural disasters sometimes provide a silver lining, by setting in motion the changes in expectations required to induce profound institutional reforms, or ignite the process of adjustment to the wealth of information projecting substantial increases in sea levels in New York City.

Naturally, further research should focus on measuring whether households and businesses are indeed relocating to other areas of the city or, on the contrary, adopting the resiliency measures required to stay in place and withstand future episodes of massive flooding.

---

<sup>64</sup>The *Biggert-Waters Flood Insurance Reform Act* was passed in 2012, prior to Sandy. These Act proposed to eliminate subsidies to flood insurance rates and phased out a number of exemptions. However, vigorous public opposition led to the 2013 *Homeowner Flood Insurance Affordability Act*, which implemented a gradual phasing out of the subsidies.

## References

- Abadie, Alberto and Javier Gardeazabal, “The Economic Costs of Conflict: A Case Study of the Basque Country,” *American Economic Review*, March 2003, *93* (1), 113–132.
- and Sofia Dermisi, “Is terrorism eroding agglomeration economies in Central Business Districts? Lessons from the office real estate market in downtown Chicago,” *Journal of Urban Economics*, September 2008, *64* (2), 451–463.
- Ambrus, Attila, Erica Field, and Robert Gonzalez, “Loss in the Time of Cholera: Long-run Impact of a Disease Epidemic on the Urban Landscape,” Mimeo Duke University May 2016.
- Bivand, Roger, Tim Keitt, and Barry Rowlingson, *rgdal: Bindings for the Geospatial Data Abstraction Library* 2015. R package version 0.9-3.
- Black, Sandra E., “Do Better Schools Matter? Parental Valuation of Elementary Education,” *The Quarterly Journal of Economics*, 1999, *114* (2), 577–599.
- Checker, Melissa, “Stop FEMA Now: Nature 2.0 and the Sacrificed Citizen,” *Geoforum*, 2016, *Special issue: Nature 2.0: Social Media, Online Activism and the Cyberpolitics of Environmental Conservation*.
- Deryugina, Tatyana, “Disaster Insurance Versus Social Insurance in the Aftermath of Hurricanes,” Technical Report 2015.
- , Laura Kawano, and Steven Levitt, “The Economic Impact of Hurricane Katrina on Its Victims: Evidence from Individual Tax Returns,” Technical Report 2014.
- Desmet, Klaus and Esteban Rossi-Hansberg, “On the spatial economic impact of global warming,” *Journal of Urban Economics*, 2015, *88* (C), 16–37.
- , Dvid Krisztin Nagy, and Esteban Rossi-Hansberg, “The Geography of Development: Evaluating Migration Restrictions and Coastal Flooding,” NBER Working Papers 21087, National Bureau of Economic Research, Inc April 2015.
- Dixon, Lloyd, Noreen Clancy, Bruce Bender, Aaron Kofner, David Manheim, and Laura Zakaras, “Flood Insurance in New York City Following Hurricane Sandy,” Technical Report, RAND Corporation, Santa Monica, CA 2013.
- Fack, Gabrielle and Julien Grenet, “When do better schools raise housing prices? Evidence from Paris public and private schools,” *Journal of Public Economics*, February 2010, *94* (1-2), 59–77.
- Groen, Jeffrey A., Mark J. Kutzbach, and Anne E. Polivka, “Storms and Jobs: The Effect of Hurricanes on Individuals’ Employment and Earnings over the Long Term,” Working Papers 15-21, Center for Economic Studies, U.S. Census Bureau January 2015.
- Gröger, André and Yanos Zylberberg, “Internal Labor Migration as a Shock Coping Strategy: Evidence from a Typhoon,” *American Economic Journal: Applied Economics*, April 2016, *8* (2), 123–53.
- Hanson, Susan, Robert Nicholls, N. Ranger, S. Hallegatte, J. Corfee-Morlot, C. Herweijer, and J. Chateau, “A global ranking of port cities with high exposure to climate extremes,” *Climatic Change*, January 2011, *104* (1), 89–111.
- Hornbeck, Richard and Daniel Keniston, “Creative Destruction: Barriers to Urban Growth and the Great Boston Fire of 1872,” NBER Working Papers 20467, National Bureau of Economic Research, Inc September 2014.

- McIntosh, Molly Fifer, “Measuring the Labor Market Impacts of Hurricane Katrina Migration: Evidence from Houston, Texas,” *The American Economic Review*, 2008, 98 (2), 54–57.
- Oates, Wallace E, “The Effects of Property Taxes and Local Public Spending on Property Values: An Empirical Study of Tax Capitalization and the Tiebout Hypothesis,” *Journal of Political Economy*, Nov./Dec. 1969, 77 (6), 957–71.
- Quintos, Carmela, “Spatial Weight Matrices and Their Use As Baseline Values and Location-Adjustment Factors in Property Assessment Models,” *Journal of Property Tax Assessment and Administration*, 2014, 11 (4), 53–64.
- Saiz, Albert and Susan Wachter, “Immigration and the Neighborhood,” *American Economic Journal: Economic Policy*, May 2011, 3 (2), 169–88.
- Schwartz, Amy Ellen, Ioan Voicu, and Keren Mertens Horn, “Do choice schools break the link between public schools and property values? Evidence from house prices in New York City,” *Regional Science and Urban Economics*, 2014, 49 (C), 1–10.
- , Scott Susin, and Ioan Voicu, “Has falling crime driven New York City’s real estate boom?,” *Journal of Housing Research*, 2003, 14 (1).
- Stocker, Thomas F, Dahe Qin, Gian-Kasper Plattner, M Tignor, Simon K Allen, Judith Boschung, Alexander Nauels, Yu Xia, Vincent Bex, and Pauline M Midgley, “Climate Change 2013: The Physical Science Basis,” Technical Report 2013. 1535 pp.
- Thaler, Richard, “A note on the value of crime control: Evidence from the property market,” *Journal of Urban Economics*, January 1978, 5 (1), 137–145.
- Zhang, Lei, “Flood hazards impact on neighborhood house prices: A spatial quantile regression analysis,” *Regional Science and Urban Economics*, 2016, 60, 12 – 19.

Table 1: Sale Price by Borough. Select percentiles

Borough	10p	50p	90p	BBL-Apt.-Year Obs.
1 Manhattan	300,000	773,870	2,950,000	196,429
2 Bronx	95,000	350,000	650,000	59,629
3 Brooklyn	182,520	484,319	999,999	195,336
4 Queens	150,000	391,400	750,000	241,432
5 Staten Island	188,000	375,000	626,223	65,838
NYC	172,380	464,500	1,300,000	758,664

Source: Own calculations based on property sales data from the NYC Department of Finance, 2003-2015. Only residential properties, but including all building types (Cooperative apartments, Condos, 1-family houses, 2-family houses, and so on).

Table 2: FEMA damage determination estimates

Sample	All	NYC	Manhattan	Bronx	Brooklyn	Queens	SI
% Affected	50.01	39.10	39.88	69.87	43.50	32.20	35.74
% Minor	43.10	46.95	59.80	27.73	48.74	50.38	37.46
% Major	6.90	13.56	0.31	2.40	7.75	17.40	26.15
% Destroyed	0.26	0.39	0.00	0.00	0.02	0.01	0.66
Obs.	318,735	67,302	2,254	1,958	29,916	21,420	11,576

Notes: Own calculations based on FEMA’s building point damage determination estimates data. Specifically, we use the variable *DMG\_COMB* that is based on a combination of visible aerial imagery and field-verified inundation observation damage. Sample “All” refers to all buildings in the Sandy Inundation area (318,735) as well as points where visible aerial imagery damage determinations were made outside the inundation zone. Sample “NYC” refers to the subset of buildings that are in one of New York City’s five boroughs (Manhattan, Bronx, Brooklyn, Queens and Staten Island). Each column adds up to 100 as it reports the distribution over damage levels for each of the samples.

Table 3: Summary statistics by borough. Sales-FEMA dataset

Borough	Obs.	Sale price (median)	HEZ12 (pct.)	Major dam- age (pct.)	Major flooding (pct.)
1 Manhattan	100,183	634,523	3.41	0	0.23
2 Bronx	52,388	374,869	1.96	0.01	0.01
3 Brooklyn	152,779	491,553	9.24	0.84	0.65
4 Queens	214,492	400,680	6.12	0.76	0.45
5 Staten Is.	58,658	391,867	12.41	0	3.00
NYC	578,500	461,944	6.74	0.50	0.68

Notes: Pct. denotes percent. HEZ12 is an indicator for being located in hurricane evacuation zones 1 or 2. Column 4 reports percent of units that suffered major damage or were destroyed. Column 5 reports percent of units that suffered more than 5.48 feet of flooding.

Table 4: Summary statistics. Sales-FEMA dataset

Variable	Obs	Mean	Std. Dev.	Min	Max
borough	578,500	3.137	1.241	1	5
block	578,500	4771.176	3732.699	1	16350
lot	578,500	57.495	126.445	1	7515
year	578,500	2008.201	3.913	2003	2015
HEZ	578,500	1.369	2.173	0	6
HEZ1	578,500	.037	.189	0	1
HEZ12	578,500	.067	.251	0	1
HEZ16	578,500	.331	.471	0	1
DMG_COMBO	578,500	.09	.405	0	4
Dam0	578,500	.025	.157	0	1
Dam1	578,500	.037	.189	0	1
Dam2	578,500	.005	.07	0	1
Depth	578,500	.177	.873	0	14.004
Sur0	578,500	.018	.134	0	1
Sur1	578,500	.042	.201	0	1
Sur2	578,500	.007	.082	0	1
Sale price	578,500	639217.2	990963.1	10000	1.50e+07
bclass 1-fam	578,500	.289	.453	0	1
bclass 2-fam	578,500	.247	.431	0	1
bclass 3-fam	578,500	.065	.246	0	1
bclass Coop	578,500	.333	.471	0	1
bclass Condo	578,500	0	0	0	0
bclass Rentals	578,500	.065	.247	0	1
Price per sqf	385,139	318.305	6668.783	.032	1350000
Sqf. gross	385,139	3547.741	18433.41	1	3750565
Year built	578,500	1936.604	103.114	1800	2015
Year altered1	110,198	1992.525	12.242	1900	2014
Year altered2	11,045	2003.601	9.951	1921	2014

Table 5: Neighborhood Fixed-Effects Models. Sales-FEMA data. Damage indicators

Dep. Var. $\ln p$	1	2	3	4	5	6
Estimation	LSDV	LSDV	Within	Within	Within	Within
Dam0	0.05 [0.09]	0.04 [0.08]	-0.01 [0.11]	0.09 [0.05]	0.07 [0.05]	-0.02 [0.12]
Dam1	-0.04 [0.02]	0.06 [0.09]	0.01 [0.13]	0.11* [0.06]	0.1 [0.06]	0 [0.14]
Dam2	-0.33*** [0.02]	-0.17* [0.09]	-0.04 [0.13]	0.12 [0.07]	0.1 [0.07]	-0.04 [0.14]
Post $\times$ Dam0	-0.12* [0.05]	-0.13*** [0.04]	-0.13*** [0.02]	-0.11*** [0.02]	-0.07*** [0.02]	-0.09*** [0.02]
Post $\times$ Dam1	-0.07 [0.05]	-0.07** [0.04]	-0.10*** [0.02]	-0.11*** [0.02]	-0.07*** [0.02]	-0.06*** [0.02]
Post $\times$ Dam2	-0.18 [0.18]	-0.18 [0.12]	-0.21*** [0.05]	-0.24*** [0.05]	-0.20*** [0.05]	-0.17*** [0.05]
$\ln \text{sqf.}$				0.26*** [0.01]	0.28*** [0.01]	
Observations	571940	571940	571940	164838	62052	186126
R-squared	0.091	0.091	0.078	0.121	0.128	0.083
Nbh. FE	borough	zip code	block	block	block	block
N. clusters	5	182	24,029	18,032	6,482	8,829
Clustered s.e.	borough	zip code	block	block	block	block
Building types	all	all	all	fam1fam2	fam1fam2	all
Sample	all	all	all	all	HEZ16	HEZ16

Notes: Models estimated on the Sales-FEMA dataset. Dependent variable is the log of the sale price. Included, but not displayed in table: quarter-year dummies and year built or last altered categories. Post is defined as a sale occurring in November 2012 or later. *Dam0* is a dummy variable for whether the unit is located in HEZ12 but did not suffer any damage. *Dam1* indicates if a unit located in HEZ12 was affected but suffered at most minor damage. *Dam2* indicates if a unit located in HEZ12 suffered major damage or was considered destroyed. The sample excludes the damaged properties outside of HEZ12, which reduces the sample from 578,500 to 571,940 BBL-year observations. Standard errors clustered at the appropriate neighborhood level. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 6: Neighborhood Fixed-Effects Models (2): Surge-based indicators

Dep. var. $\ln p$	1	2	3	4	5	6
Estimation	LSDV	LSDV	Within	Within	Within	Within
Sur0	0.08 [0.05]	0.09 [0.08]	-0.01 [0.11]	0.09 [0.05]	0.07 [0.05]	-0.02 [0.12]
Sur1	-0.04** [0.01]	0.05 [0.08]	0.01 [0.12]	0.09 [0.06]	0.08 [0.06]	0 [0.12]
Sur2	-0.26*** [0.04]	-0.20** [0.09]	-0.05 [0.13]	0.08 [0.06]	0.07 [0.06]	-0.06 [0.13]
Post $\times$ Sur0	-0.05 [0.04]	-0.06 [0.04]	-0.07*** [0.02]	-0.09*** [0.02]	-0.05** [0.02]	-0.03 [0.02]
Post $\times$ Sur1	-0.08* [0.04]	-0.09** [0.03]	-0.12*** [0.02]	-0.12*** [0.02]	-0.08*** [0.02]	-0.08*** [0.02]
Post $\times$ Sur2	-0.20*** [0.04]	-0.21*** [0.06]	-0.21*** [0.04]	-0.17*** [0.04]	-0.13*** [0.04]	-0.17*** [0.04]
$\ln \text{sqf.}$				0.26*** [0.01]	0.28*** [0.01]	
Observations	571940	571940	571940	164838	62052	186126
R-squared	0.091	0.092	0.078	0.121	0.128	0.083
Nbh. FE	borough	zip code	block	block	block	block
N. clusters	5	182	24029	18032	6482	8829
Clustered s.e.	borough	zip code	block	block	block	block
Building types	all	all	all	fam1fam2	fam1fam2	all
Sample	all	all	all	all	HEZ16	HEZ16

Notes: Models estimated on the transactions-based final dataset. Dependent variable is the log of the sale price. Included, but not displayed in table: quarter-year dummies and year built or last altered categories. Post is defined as a sale occurring in November 2012 or later. *Sur0* is a dummy variable that takes a value of one if a unit is located in HEZ12 but was not flooded. *Sur1* is an indicator for whether a unit was located in HEZ12 and experience storm surge depth between 0 and 5.48 feet. *Sur2* indicates if a unit located in HEZ12 experienced storm surge above 5.48 feet (90th percentile conditional on positive storm surge). The sample excludes the damaged properties outside of HEZ12, which reduces the sample from 578,500 to 571,940 BBL-year observations. Standard errors clustered at the appropriate neighborhood level. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 7: Property FE. Repeat sales sample

Dep. Var. $\ln p$	1	2	3	4	5	6
Post $\times$ HEZ12	-0.11*** [0.02]			-0.08*** [0.02]		
Post $\times$ Dam0		-0.14*** [0.02]			-0.07*** [0.02]	
Post $\times$ Dam1		-0.08*** [0.03]			-0.07 [0.04]	
Post $\times$ Dam2		-0.17*** [0.06]			-0.18* [0.10]	
Post $\times$ Sur0			-0.13*** [0.03]			-0.10** [0.04]
Post $\times$ Sur1			-0.10*** [0.03]			-0.10*** [0.03]
Post $\times$ Sur2			-0.07** [0.04]			-0.02 [0.05]
Observations	248154	248154	248154	65764	65764	65764
Groups	106325	106325	106325	29955	29955	29955
R-squared	0.164	0.164	0.164	0.141	0.141	0.141
FE	BBL-Apt	BBL-Apt	BBL-Apt	BBL	BBL	BBL
Sample	All	All	All	fam1fam2	fam1fam2	fam1fam2

Notes: Models estimated on the Sales-FEMA final dataset. Dependent variable is the log of the sale price. Included, but not displayed in table: quarter-year dummies. The sample excludes the damaged properties outside of HEZ12, which reduces the sample from 578,500 to 571,940 BBL-year observations. Standard errors clustered at the city block level. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 8: Summary statistics. Combined dataset FEMA and assessed market values

Variable	Obs	Mean	Std. Dev.	Min	Max
borough	11,366,412	3.651	.892	1	5
block	11,366,412	5837.153	3677.889	3	16323
lot	11,366,412	61.151	117.41	1	6960
year	11,366,412	2007.099	4.816	1999	2015
DMG_COMBO	11,366,412	.1	.425	0	4
Depth	11,366,412	.208	.916	0	14.167
HEZ1	11,366,412	.044	.205	0	1
HEZ12	11,366,412	.081	.272	0	1
HEZ16	11,366,412	.368	.482	0	1
Dam0	11,366,412	.033	.179	0	1
Dam1	11,366,412	.042	.2	0	1
Dam2	11,366,412	.005	.073	0	1
Sur0	11,366,412	.021	.143	0	1
Sur1	11,366,412	.053	.224	0	1
Sur2	11,366,412	.007	.081	0	1
FULLVAL	11,366,412	513186	448451.3	10000	1.50e+07

Notes: HEZ1 is an indicator for being located in hurricane evacuation zone 1. Likewise, HEZ12 is an indicator for being located in hurricane evacuation zones 1 or 2, and HEZ16 for evacuation zones 1 through 6. FULLVAL is the assessed market value. These data contain only buildings in tax class one (houses).

Table 9: Property FE. Combined FEMA data with assessed market values

Dep. Var. $\ln p$	1	2	3	4	5	6	7	8
Post $\times$ HEZ12	-0.20*** [0.004]	-0.21*** [0.001]						
Post $\times$ Dam0			-0.19*** [0.002]	-0.17*** [0.002]	-0.19*** [0.007]			
Post $\times$ Dam1			-0.22*** [0.001]	-0.20*** [0.001]	-0.22*** [0.004]			
Post $\times$ Dam2			-0.23*** [0.003]	-0.21*** [0.003]	-0.23*** [0.008]			
Post $\times$ Sur0						-0.13*** [0.002]	-0.12*** [0.002]	-0.13*** [0.009]
Post $\times$ Sur1						-0.23*** [0.001]	-0.21*** [0.001]	-0.23*** [0.004]
Post $\times$ Sur2						-0.25*** [0.004]	-0.24*** [0.004]	-0.25*** [0.009]
Obs. (million)	11.2	11.2	11.2	4.0	11.2	11.2	4.0	11.2
R-squared	0.620	0.804	0.804	0.784	0.804	0.804	0.785	0.804
Fixed-effects	BB	BBL						
Number of Blocks	23,320							
Number of BBL		658,687	658,687	239,825	658,687	658,687	239,825	658,687
Sample	all	all	all	HEZ16	all	all	HEZ16	all
Clustered s.e.	BB	BBL	BBL	BBL	Block	BBL	BBL	Block

Notes: Models estimated on the dataset using assessed market values. All models include BBL (borough-block-lot) fixed effects. Dependent variable is the log of the assessed market value. Dam0 dummy variable if unit located in HEZ12 but did not suffer any damage. Dam1 dummy variable if unit located in HEZ12, was affected but suffered at most minor damage. Dam2 dummy variable if unit located in HEZ12 and suffered major damage or was destroyed. Sur0 dummy variable if unit located in HEZ12 but was not flooded. Sur1 dummy variable if unit located in HEZ12 and the storm surge depth was between 0 and 5.5 feet. Sur2 dummy variable if unit located in HEZ12 and storm surge above 5.5 feet (90th percentile conditional on positive storm surge). Sample excludes the damaged properties outside of HEZ12. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 10: Probability of Being Sold (LPM)

Dep. Var. Sold	1	2	3	4	5	6
Post × HEZ12	0.0015*		0.0014		0.0014**	
	[0.0009]		[0.001]		[0.0006]	
Post × Dam0		0.0025		0.0017		0.0017**
		[0.0017]		[0.002]		[0.0009]
Post × Dam1		0.0013		0.0016*		0.0016**
		[0.0009]		[0.001]		[0.0008]
Post × Dam2		-0.0029		-0.0028		-0.0028
		[0.0031]		[0.003]		[0.0024]
Observations	8,498,857	8,498,857	8,498,857	8,498,857	8,498,857	8,498,857
R-squared	0.005	0.005	0.005	0.005	0.005	0.005
FE	Block	Block	Property	Property	Property	Property
Cluster s.e.	Block	Block	Block	Block	Property	Property

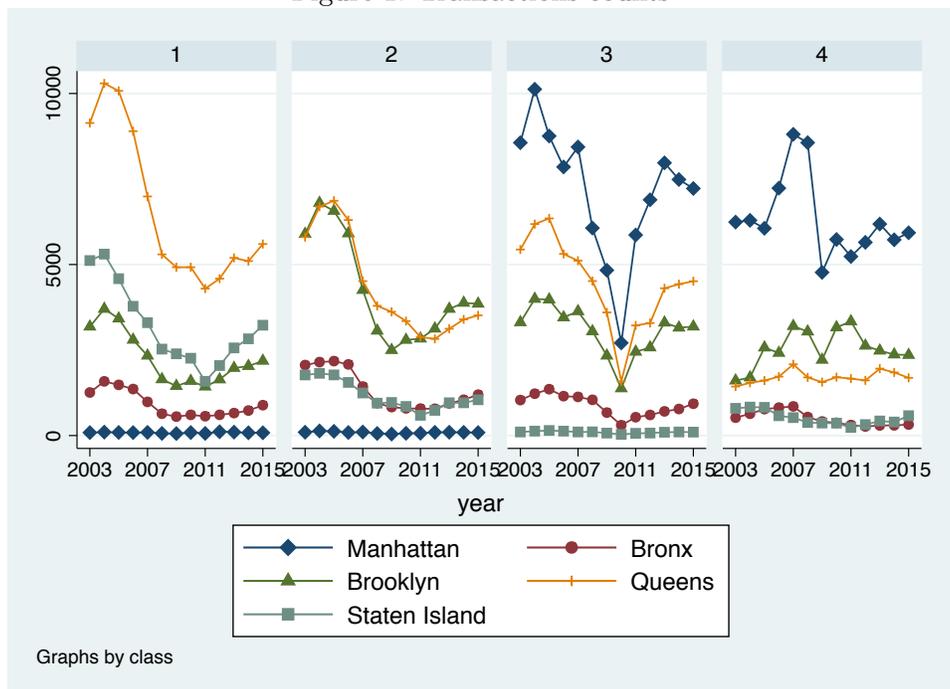
Notes: Merged dataset sales and assessed market values, years 2003-2015. Only tax-class-one buildings (houses) included. Condos have been excluded. Year dummies always included. All blocks in the city included (approx. 23,000). Number of tax lots (properties) included is 665,930. Year dummies included. Standard errors clustered at the block level. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table 11: Time-varying coefficients by year

Model	Model1	Model2	Model2	Model2
Treatment	HEZ12	Dam0	Dam1	Dam2
T × 2004	0.01 [0.02]	0.02 [0.02]	0.01 [0.03]	-0.04 [0.07]
T × 2005	0.02 [0.02]	0 [0.02]	0.04 [0.03]	-0.03 [0.07]
T × 2006	-0.02 [0.02]	-0.02 [0.02]	-0.03 [0.04]	0.07 [0.07]
T × 2007	0 [0.02]	-0.04 [0.03]	0.02 [0.03]	0.08 [0.06]
T × 2008	-0.04 [0.03]	-0.04 [0.03]	-0.05 [0.05]	-0.01 [0.06]
T × 2009	-0.02 [0.02]	-0.01 [0.03]	-0.03 [0.04]	-0.01 [0.07]
T × 2010	0.03 [0.02]	0.06** [0.03]	0.01 [0.04]	-0.13 [0.09]
T × 2011	0.02 [0.03]	0.01 [0.03]	0.04 [0.04]	-0.08 [0.10]
T × 2012	-0.02 [0.03]	-0.01 [0.03]	-0.01 [0.04]	-0.15* [0.09]
T × 2013	-0.18*** [0.03]	-0.17*** [0.03]	-0.18*** [0.04]	-0.26*** [0.09]
T × 2014	-0.08*** [0.02]	-0.07** [0.03]	-0.07** [0.03]	-0.24*** [0.08]
T × 2015	-0.11*** [0.02]	-0.12*** [0.03]	-0.08** [0.04]	-0.25** [0.10]
Observations	164838	164838		
Number of BB	18032	18032		
R-squared	0.114	0.115		
Fixed-effects	Block	Block		

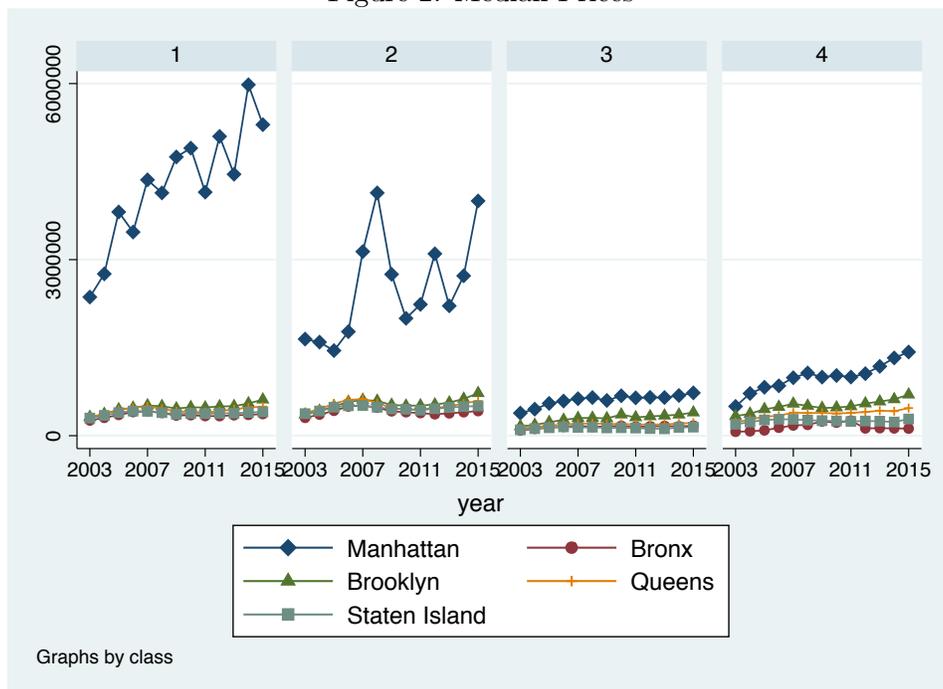
Notes: T denotes dummy for the corresponding treatment, HEZ12 or a Damage level indicator. Models 1 and 2 include block FE, year dummies, year built or last altered, and square footage. The sample includes only 1-family and 2-family houses. Standard errors clustered by block in both models. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Figure 1: Transactions counts



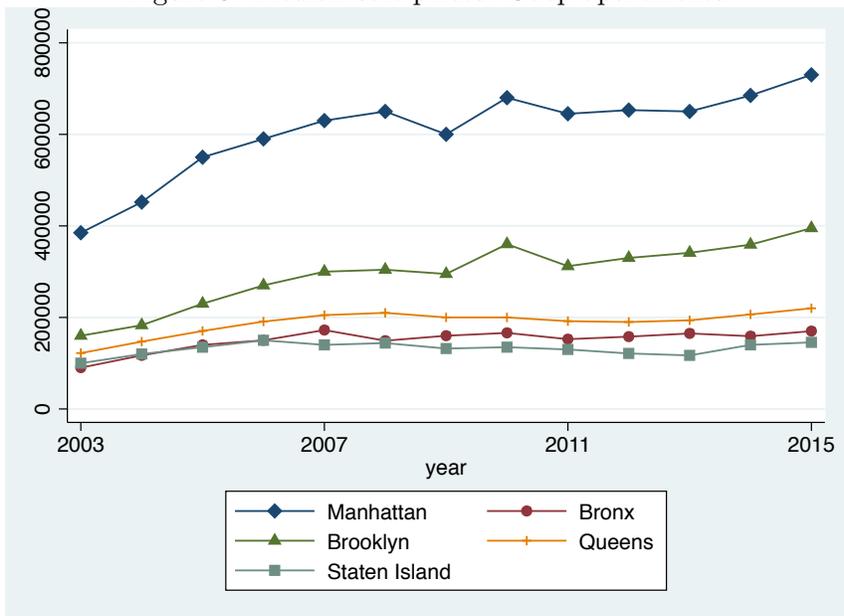
Notes: Transactions-based data from the NYC Department of Finance, 2003-2015. Type 1 refers to 1-family houses, type-2 refers to 2-family houses, type-3 refers to apartments in Cooperative buildings, and type-4 refers to Condominium apartments.

Figure 2: Median Prices



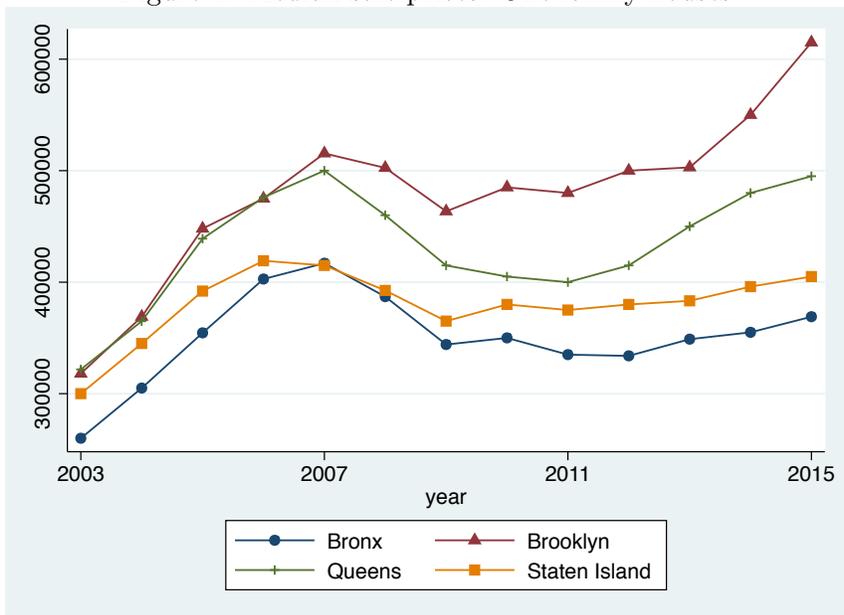
Notes: Transactions-based data from the NYC Department of Finance, 2003-2015. Type 1 refers to 1-family houses, type-2 refers to 2-family houses, type-3 refers to apartments in Cooperative buildings, and type-4 refers to Condominium apartments.

Figure 3: Median sale prices. Coop apartments



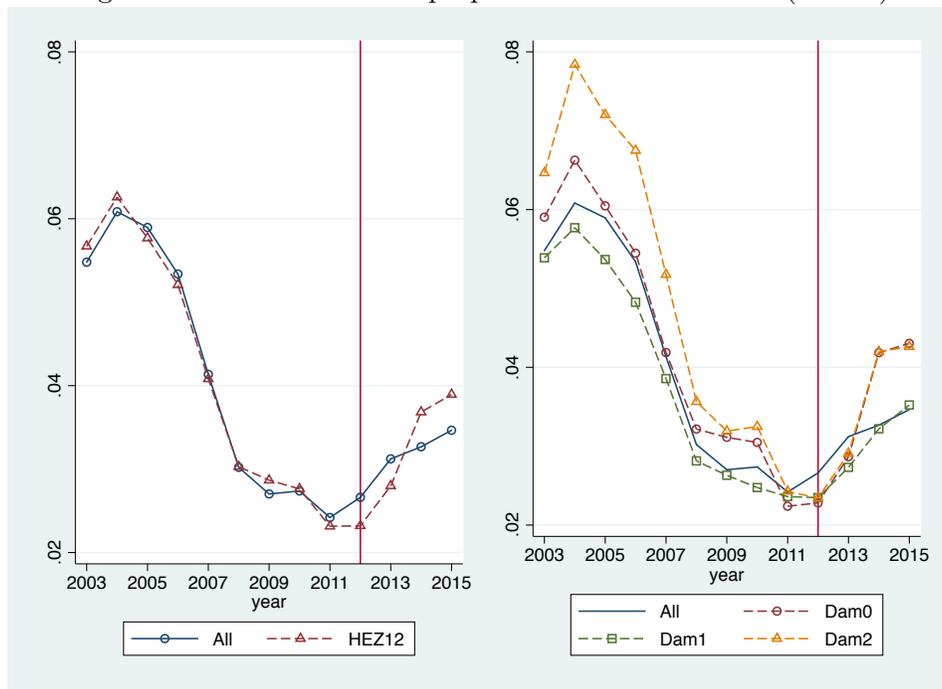
Notes: Transactions-based data from the NYC Department of Finance, 2003-2015.

Figure 4: Median sale prices. One-family houses



Notes: Transactions-based data from the NYC Department of Finance, 2003-2015. The borough of Manhattan has been excluded. Median sale prices for one-family houses in Manhattan are one order of magnitude higher than for the other boroughs.

Figure 5: Data. Fraction of properties sold. Tax class 1 (houses)



Notes: Fraction of properties in each category sold in any given year. Tax class 1 (houses only).

Figure 6: By-year point estimates  $D_{HEZ12}$

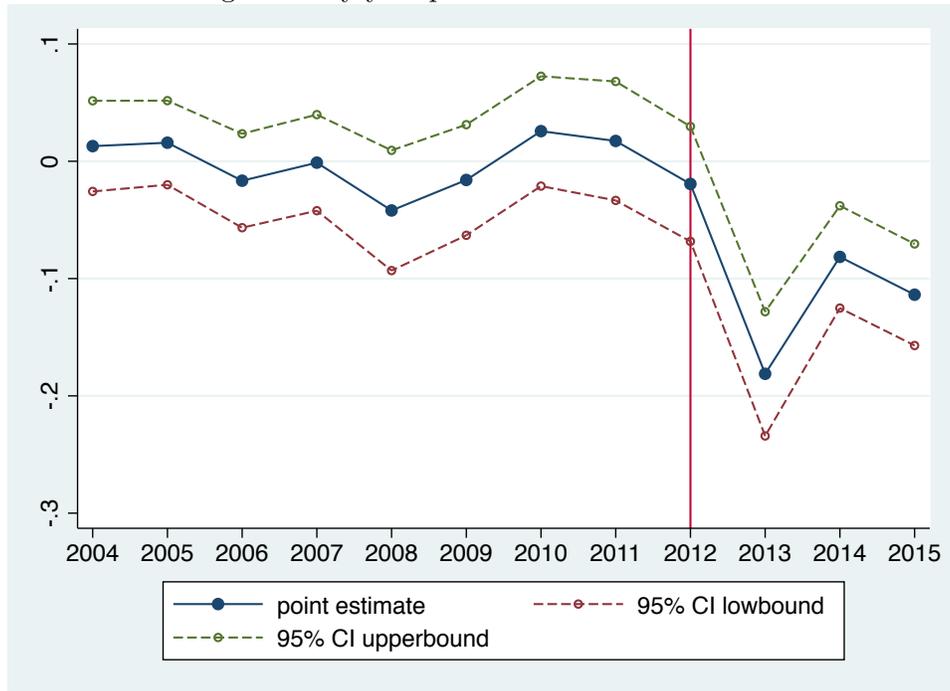
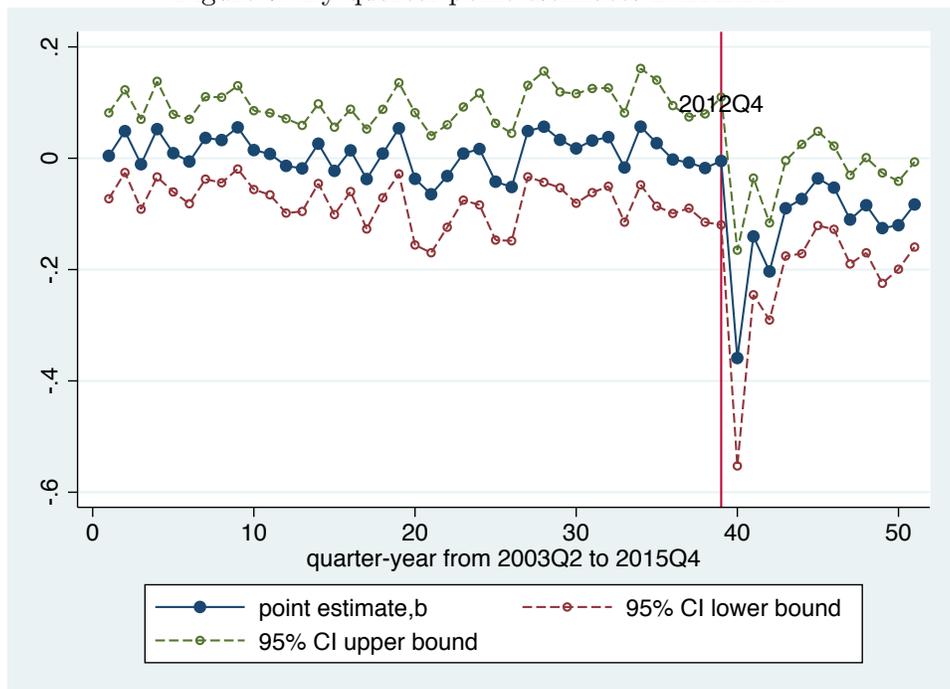


Figure 7: By-quarter point estimates  $D_{HEZ12}$



# Appendix

## A Details on Merging of Datasets

We describe in more detail the merging process. We also report on a number of checks aimed at evaluating the quality of the merge.

1. *FEMA Damage-Point Estimates and PLUTO*. In the FEMA data, each observation is characterized by its longitude and latitude in spherical coordinates. In total we had more than 55,000 individual points corresponding to New York City (and 319,000 for the overall Sandy inundation zone). We first mapped spherical coordinates to Cartesian XY (New York state) coordinates. Next we mapped these into New York City tax lots using the shape files provided by PLUTO. In the resulting dataset each observation is identified by its BBL (and its longitude and latitude).

Then we proceeded to check the quality of the merge between the FEMA and PLUTO datasets. About 99.7% of the cases in the FEMA data mapped into a NYC tax lot. Next, we randomly sampled 50 cases and manually checked that their spherical coordinates landed in the correct tax lot.<sup>65</sup> The matches were correct in 98% of the cases (49 out of the 50).<sup>66</sup> In short the mapping from FEMA to tax lots in PLUTO was extremely accurate.

2. *Multiplicity of FEMA cases within a BBL*. In the FEMA data, each observation is uniquely defined by an administrative ID, which is not useful for our purposes, and a latitude-longitude (Cartesian) pair. However, not all of these observations are uniquely matched to a single BBL.<sup>67</sup> Specifically, 14% of all FEMA cases correspond to multiple determinations points within the same tax lot.<sup>68</sup> We adopt the simplest option: we average damage values across all cases within the same BBL.
3. *FEMA-PLUTO and HEZ*. We checked the quality of this match in a similar manner as previously. Again the success rate was very high: only 0.4% of the cases (fewer than 200) in the FEMA-PLUTO data were not matched to a tax lot. We again randomly sampled 25 cases from the FEMA-PLUTO-HEZ dataset. We checked the spherical coordinates for each of those points using again the NYC City Map to locate the resulting tax lot. Then we used the NYC Hurricane Evacuation Zone Map to check the evacuation zone assigned to that point ([http://www.nyc.gov/html/oem/downloads/pdf/hurricane\\_map.english.pdf](http://www.nyc.gov/html/oem/downloads/pdf/hurricane_map.english.pdf)). The success rate was 100%.
4. *FEMA Storm Surge and PLUTO*. The raw storm surge data contains 350,154 observations covering the 5 boroughs of the city. Each observation refers to a longitude-latitude pair and the data has high geographic resolution. Hence, not surprisingly,

---

<sup>65</sup>To do this we used the NYC City Map (<http://maps.nyc.gov/doitt/nycitymap>).

<sup>66</sup>In the unsuccessful match the procedure identified the neighboring lot.

<sup>67</sup>The 55,534 observations correspond to 47,879 unique BBLs.

<sup>68</sup>Specifically, 3.80% of the observations appear exactly twice in a BBL, 1.08% appear exactly three times in a BBL, 0.66% appear four times, and 8% appear 5 or more times. The most extreme case is a BBL for which we have 1,911 observations, which corresponds to the Breezy Point Cooperative in Queens that contains many one-family houses.

many points map into the same BBL and therefore there are many duplicates (about 2,000 on average but ranging from 1 to 30,089). Since our unit of analysis is based on BBLs in the final dataset, we now collapse by BBL. The resulting data contains 7,675 observations. We then proceed to merge with PLUTO and obtain a perfect match (except for one observation). Some of those BBLs are among the small number that cannot be assigned to a hurricane evacuation zone (including the non-evacuation zone). In the end 6,449 BBLs can be matched with the PLUTO-HEZ dataset. We view this list of BBLs as the complete list of BBLs that were located in the Sandy surge area.

**The PLUTO-HEZ-FEMA Data.** This dataset encompasses all the data that is time-invariant: the inclusion or not of each tax lot in a hurricane evacuation zones and the level of damage (if any) suffered during Sandy. The unit of observation is the BBL.<sup>69</sup> We then merge these data with the property sales dataset, where the unit of observation is the BBL-Apartment and year.

The merger proceeds in several steps. First, we begin with the PLUTO-HEZ dataset, which contains 857,000 tax lots. However, in 27,000 cases the hurricane evacuation zone is missing. We drop these observations so that the resulting dataset has about 830,000 tax lots. Second, we merge with the PLUTO-FEMA dataset, which contains roughly 48,000 cases (tax lots). This dataset contains all the tax lots (buildings) affected by Sandy. The vast majority (98%) of the tax lots in PLUTO-FEMA are successfully matched to the (much larger) set of tax lots in the PLUTO-HEZ dataset. The crucial step now is that we assign a zero value for the damage variable to all tax lots that were in the PLUTO-HEZ dataset but were not in the PLUTO-FEMA dataset. That is, we rely on the fact that the FEMA dataset contained all buildings affected by Sandy and that any building not included in the dataset was not damaged. The combined PLUTO-HEZ-FEMA dataset contains over 830,000 tax lots. We present some descriptive statistics in [Table 13](#).

---

<sup>69</sup>Recall that in the FEMA dataset we collapsed all cases by BBL so that instances of multiple cases with the same BBL got averaged into a single value.

## B Tables and Figures

Table 12: **Hurricane Evacuation Zones. Percent distribution.**

<b>HEZ</b>	NYC	Manhattan	Bronx	Brooklyn	Queens	SI
1	4.25	2.38	2.12	3.78	5.33	4.68
2	3.51	3.17	1.15	5.6	1.42	6.15
3	5.08	5.12	2.03	10.47	2.44	2.14
4	5.01	7.28	3.73	8.09	2.55	4.71
5	9.66	12.61	12.11	12.81	7.12	6.47
6	8.73	13.67	8.48	9.27	9.05	5.17
X	60.65	54.69	66.35	48.8	68.85	63.85
NA	3.07	1.09	4.04	1.19	3.25	6.83
Total	857,115	42,810	89,596	277,121	324,128	123,460

Source: The table reports the number of tax lots (BBLs) in each evacuation zone. Each column adds up to 100. Own calculations based on PLUTO-HEZ dataset. Zone 1 is the highest risk and zone 6 the lowest risk. Zone X refers to buildings outside the evacuation zones. NA refers to tax lots for which the HEZ was not available.

Table 13: Summary statistics PLUTO-HEZ-FEMA dataset.

Variable	Obs	Mean	Std. Dev.	Min	Max
block	830800	5279.852	3665.687	1	16350
lot	830800	120.406	684.162	1	7533
BBL	830800	3.50e+09	1.03e+09	1.00e+09	5.08e+09
xcoord	830800	1006261	32253.97	913225	1067279
ycoord	830800	191612	30493.76	120920	272275
borough_code	830800	3.449	1.018	1	5
<b>HEZ</b>	830800	1.519	2.217	0	6
<b>DMG_COMBO</b>	830800	.095	.416	0	4

Table 14: Building-class distribution

Building Class	Housing dataset	Final dataset
1-family house %	23	29
2-family house %	19	25
3-family house %	5	6.5
Coop. Apt. %	27	33
Condominium %	20	0
Rentals %	5	6.5
Sum	99	100

Table 15: Damage by building class within HEZ12. Sales-FEMA dataset.

Damage	None	Affected or Minor	Major or Destroyed
1-family house	7,149	6,701	1,069
2-family house	3,541	6,582	786
Coop apt.	1,896	7,147	459
<b>All HEZ12</b>	13,481	22,740	2,727