NIST Technical Note 1996

Estimating Structural Seismic Retrofit Costs for Federal Buildings

Juan F. Fung
David T. Butry
Siamak Sattar
Steven L. McCabe

NIST

**National Institute of
Standards and Technology**
U.S. Department of Commerce

NIST Technical Note 1996

Estimating Structural Seismic Retrofit Costs for Federal Buildings

Juan F. Fung
David T. Butry
Applied Economics Office
Engineering Laboratory

Siamak Sattar
Steven L. McCabe
Earthquake Engineering Group
Engineering Laboratory

June 2018

Abstract

In NIST Technical Note (TN) 1973 [1], we develop a methodology for estimating seismic retrofit costs. The general approach is to build a predictive model of seismic retrofit costs that (1) captures the essential factors that drive costs, such as building construction type and square footage; and (2) is reproducible using data available to decision makers.

In this report, we apply the methodology to generate retrofit cost predictions for *typical* federal buildings. Typical federal buildings are meant to be statistically representative of the federal building inventory, but are *not actual federal buildings*. Our results suggest that retrofitting an entire typical building inventory for life safety will cost about $20 billion. The caveat is that our model only predicts *structural* mitigation costs. Other construction costs, as well as *non-structural* mitigation costs, can easily increase the total cost.

Presidential Executive Order 13717 (EO 13717), *Establishing a Federal Earthquake Risk Management Standard*, encourages federal agencies to "enhance resilience...[to] future earthquakes" by evaluating and retrofitting existing federal buildings based on current existing building codes. This report, together with NIST TN 1973 [1], provides a roadmap for building owners to quickly approximate expected seismic retrofit costs.

i

Disclaimer

The policy of the National Institute of Standards and Technology is to use metric units in all its published materials. Because this report is intended for the U.S. construction industry that uses U.S. customary units, it is more practical and less confusing to include U.S. customary units as well as metric units. Measurement values in this report are therefore stated in metric units first, followed by the corresponding values in U.S. customary units within parentheses.

Certain commercial entities, equipment, or materials may be identified in this document to describe an experimental procedure or concept adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the entities, materials, or equipment are necessarily the best available for the purpose.

Table of Contents

List of Tables

List of Figures

Glossary

FEMA: Federal Emergency Management Agency

FRPP: Federal Real Property Profile

g: The acceleration of gravity, 9.8 (m/s^2)

GLM: Generalized Linear Model

IBC: International Building Code

IEBC: International Existing Building Code

IO: Immediate Occupancy, as defined in FEMA 156 [2]

NEHRP: National Earthquake Hazards Reduction Program

NIST: National Institute of Standards and Technology

LS: Life Safety, as defined in FEMA 156 [2]

pga: Peak ground acceleration (with a 10 % probability of exceedance in 50 years), as a fraction of standard gravity (g)

RMSE: Root Mean Square Error

TN: Technical Note

USD: United States Dollars

USGS: United States Geological Survey

1. Introduction

Presidential Executive Order 13717 (EO 13717), *Establishing a Federal Earthquake Risk Management Standard*, addresses the seismic safety of federal buildings. EO 13717 asks "each executive department and agency...to enhance resilience by reducing risk to the lives of building occupants and improving continued performance of essential functions following future earthquakes." In particular, existing federal buildings should be evaluated and retrofitted based on current existing building codes.[1] However, while guidance on evaluation and retrofit practices is readily available (e.g., ASCE/SEI 41-13 [3], FEMA 547 [4], RP-8 [5], and RP-9 [6]), a standard approach to estimating retrofit costs does not exist. Moreover, the absence of easily obtainable estimates can make retrofits an impractcal option for decision makers.

In this report, we present cost estimates for the seismic retrofit of typical buildings owned and leased by the federal government. In particular, we apply a methodology that we develop in NIST Technical Note (TN) 1973 [1], "A Methodology for Estimating Seismic Retrofit Costs," to *predict* retrofit costs for typical buildings in the Federal Real Property Profile (FRPP) inventory of federal buildings.

Typical buildings are parametrized by observable building characteristics such as building construction type, age, and square footage, as well as seismicity and the desired performance objective (life safety or immediate occupancy) for the building. Our methodological approach is to fit a linear regression model of costs as a function of such observable building characteristics, and to use this model to generate predictions for typical federally owned and leased buildings.

An important caveat is that typical buildings are meant to be *statistically representative* of the federal building inventory, but are not *actual* federal buildings. This is largely a result of the data limitations in the FRPP data relative to the data requirements for retrofit cost prediction, as described in Sec. 2.2.

As such, **the cost estimates presented in this report should not be interpreted as guidance for budget and approporiations decisions**. Rather, our estimates serve as a baseline order of magnitude for retrofit costs under reasonable assumptions about the federal building inventory, subject to the limiations of the data. Federal agencies have better information about their buildings and thus should be able to obtain more accurate estimates themselves for planning and budgeting purposes.

The predicted average total cost to retrofit all buildings in a typical building inventory for a life safety performance objective, *LS*, is \$19.649 billion, with 95 % prediction interval (\$9.709 billion, \$33.781 billion).[2]

If, instead, we assume buildings deemed to be critical are retrofitted for an immediate occupancy performance objective, *IO*, and the rest for *LS*, as discussed in Sec. 2.4.1, then the predicted average total cost to retrofit the entire inventory is \$20.122 billion, with 95 % prediction interval (\$9.951 billion, \$34.601 billion).

---

[1]Typically the International Existing Building Code (IEBC).

[2]All values are given in 2016 USD, unless otherwise noted.

Table 1. Seismic retrofit options and associated cost estimates.

Option	*LS* retrofits	*IO* retrofits	Cost
A	117 813	0	\$19.649 billion
B	101 947	15 866	\$20.122 billion

The two options are presented in Table 1. More details are given in Sec. 3, including costs categorized by agency, region, performance objective, whether a building is owned or leased, and building historic status.

1.1   Seismic risk to federally owned and leased buildings

Significant earthquake risk—the potential for damage from earthquakes—threatens federal buildings and thus the functioning of the federal government. To better understand seismic risk to federal buildings, consider Fig. 1, which maps peak ground acceleration (*pga*) with a 10 % probability of exceedance in 50 years for the continental United States. Peak ground acceleration is a measure used to communicate the potential hazard from ground motion during an earthquake in relation to gravity.[3] In Fig. 1, red areas show places where there is a 10 % chance that within a 50 year period, ground motion acceleration from an earthquake would be severe.

Figure 2, on the other hand, maps the concentration of federally owned and leased buildings in the continental United States by their estimated seismicity.[4] Adopting the approach we develop in NIST TN 1973 [1], we define low seismicity ("L") as $pga < 0.1$; moderate seismicity ("M") as $pga \in [0.1, 0.2)$; high seismicity ("H") as $pga \in [0.2, 0.4)$; and anything greater than 0.4 as very high seismicity ("VH"). Estimated seismicity values are derived from the USGS data given in Fig. 1.

Note the overlap of high building density with high seismic hazard in many parts of the country, particularly in the west coast. Moreover, a fairly high concentration of buildings in the midwest and east coast is vulnerable to the New Madrid seismic zone, centered roughly where Missouri and Tennessee meet.

Table 2 provides a more detailed picture of seismic risk to federal buildings by agency. The table presents summaries of seismic hazard, based on average and maximum *pga*, as well as other statistics that characterize federal buildings in the FRPP. Given the large number of federally owned and leased buildings managed across 25 federal agencies (roughly 130 819 in fiscal year 2015), mitigation is a growing concern.[5]

---

[3]While this may not be the best measure of seismicity (as compared to horizontal spectral response, for example), it is a suitable proxy for seismic hazard risk and is roughly on the same scale as the measure of seismicity used in FEMA 156 [2] and FEMA 157 [7], as discussed in NIST TN 1973 [1].

[4]We focus on the continental United States due to availability of seismicity data; see [1].

[5]Federal building data in Fig. 2 and Table 2 is based on the 117 813 buildings from the Federal Real Property (FRPP) for Fiscal Year 2015 (FY15) that we were able to geocode, and does not include buildings reported by the Department of Defense, which reported an additional 142 306 buildings in FY 15. See https://www. realpropertyprofile.gov/FRPPMS for more information.
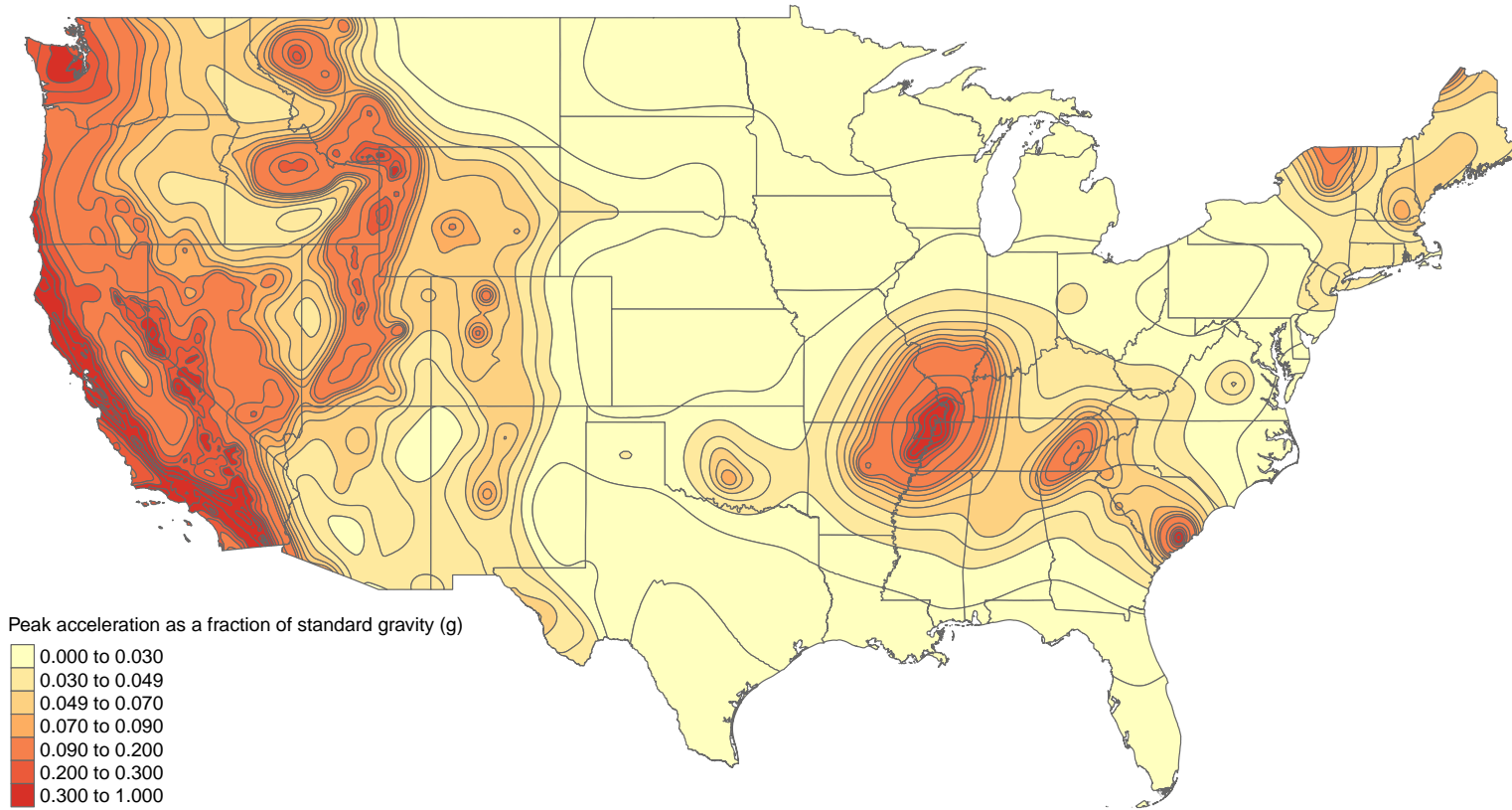
Fig. 1. Peak ground acceleration (*pga*) with a 10 % probability of exceedance in 50 years for the continental United States, based on the 2014 USGS National Seismic Hazard Map data.

Building seismicity
- L
- M
- H
- VH
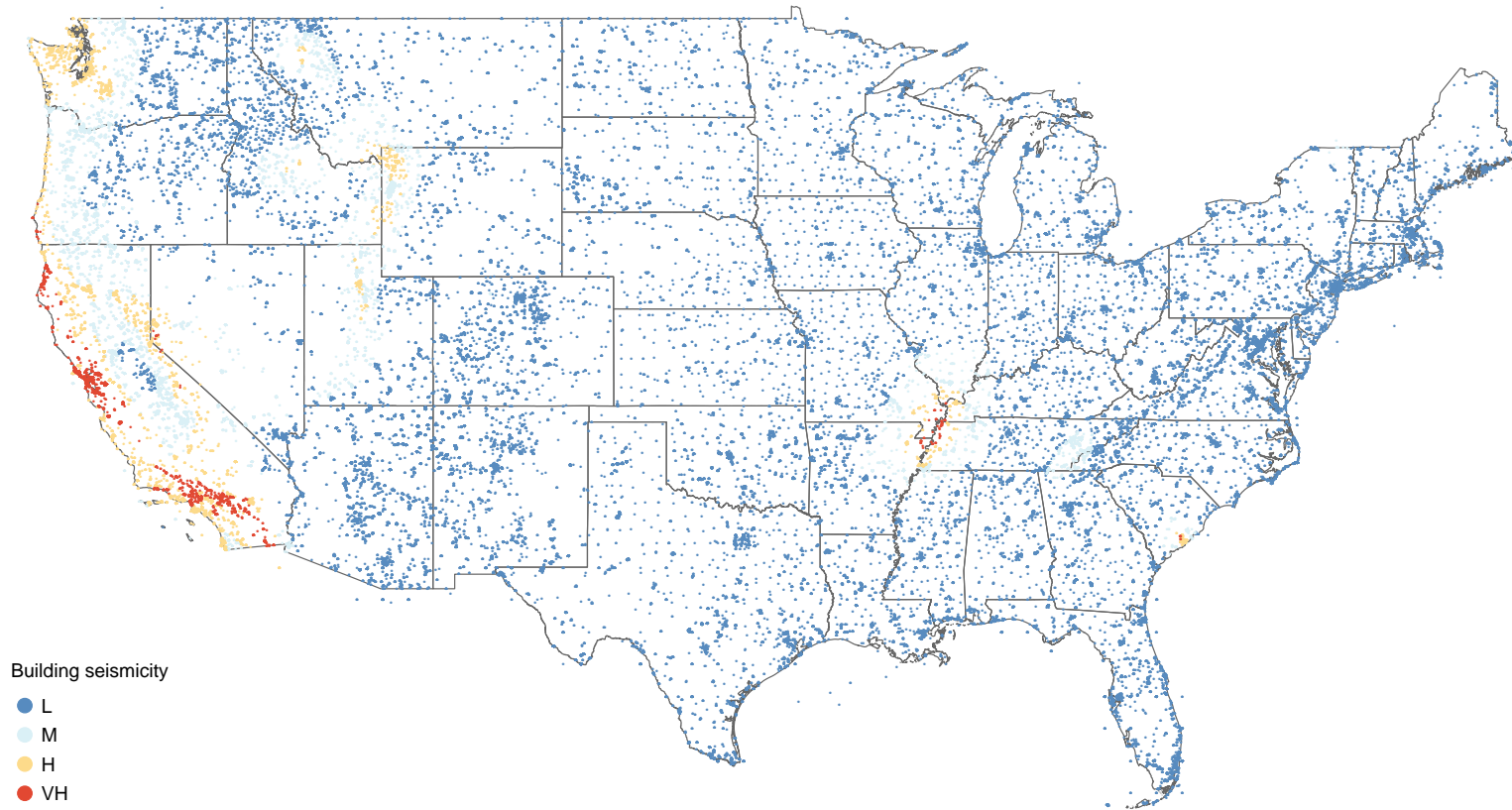
Fig. 2. Federal building seismicity dot-density map for the United States, based on FRPP building data for FY15 and the hazard map given in Fig. 1.

Table 2. Summary statistics for federal buildings by agency, based on FRPP building data for FY15.

Agency	Acronym	Number of buildings	Mean pga (g)	Max pga (g)	Area: 1000 sq ft (sq m)	Percent Owned	Percent Historic
Broadcasting Board of Governors	BBG	30	0.056	0.150	137 (13)	100 %	0 %
DC Court Services & Offender Supervision Agency	CSOSA	10	0.020	0.020	239 (22)	0 %	0 %
Department of Homeland Security	DHS	8331	0.091	0.800	47 351 (4399)	87.36 %	1.8 %
Department of Commerce	DOC	436	0.067	0.600	8353 (776)	79.59 %	2.52 %
Department of Energy	DOE	10 611	0.091	0.600	115 309 (10 712)	99.51 %	10.62 %
Department of the Interior	DOI	40 670	0.076	0.800	99 973 (9287)	99.28 %	20.18 %
Department of Justice	DOJ	3878	0.053	0.400	70 096 (6512)	99.23 %	12.09 %
Department of Labor	DOL	2343	0.067	0.600	25 175 (2339)	89.46 %	1.11 %
Department of State	DOS	140	0.026	0.300	1437 (133)	98.57 %	0 %
Department of Transportation	DOT	3468	0.067	0.800	14 866 (1381)	81.03 %	2.54 %
Environmental Protection Agency	EPA	172	0.069	0.300	4215 (392)	100 %	5.81 %
Federal Communications Commission	FCC	45	0.082	0.400	103 (10)	100 %	2.22 %
General Services Administration	GSA	8467	0.056	0.600	423 331 (39 327)	18.63 %	5.43 %
Department of Health and Human Services	HHS	2650	0.030	0.250	31 500 (2926)	96.68 %	6.91 %
John F. Kennedy Center for the Performing Arts	JFKC	1	0.020	0.020	1500 (139)	100 %	100 %
National Archives and Records Administration	NARA	27	0.047	0.300	4910 (456)	81.48 %	11.11 %
National Aeronautics And Space Administration	NASA	2675	0.083	0.800	47 089 (4375)	99.21 %	12.37 %
National Science Foundation	NSF	199	0.034	0.300	1207 (112)	92.46 %	0 %
Office of Personnel Management	OPM	4	0.025	0.030	75 (7)	0 %	0 %
Smithsonian Institution	SI	24	0.026	0.050	1594 (148)	0 %	0 %
Department of the Treasury	TRE	106	0.059	0.600	6371 (592)	10.38 %	4.72 %
Tennessee Valley Authority	TVA	2442	0.089	0.600	28 317 (2631)	98.65 %	0 %
Department of Agriculture	USDA	23 884	0.083	0.800	55 884 (5192)	86.39 %	9.29 %
United States Holocaust Memorial Council (Holocaust Museum)	USHMM	2	0.020	0.020	320 (30)	100 %	50 %
Department of Veterans Affairs	VA	7198	0.063	0.600	80 783 (7505)	77.41 %	27.48 %

1.2 An overview of the methodology and how to apply it

The methodology we introduce in NIST TN 1973 [1] builds on FEMA 156 [2] and 157 [7], *Typical Costs for Seismic Rehabilitation of Existing Buildings*, Volumes 1 and 2, respectively—hereafter FEMA 156/157. In particular, we use the data collected for the FEMA reports to build a predictive model of seismic retrofit costs.
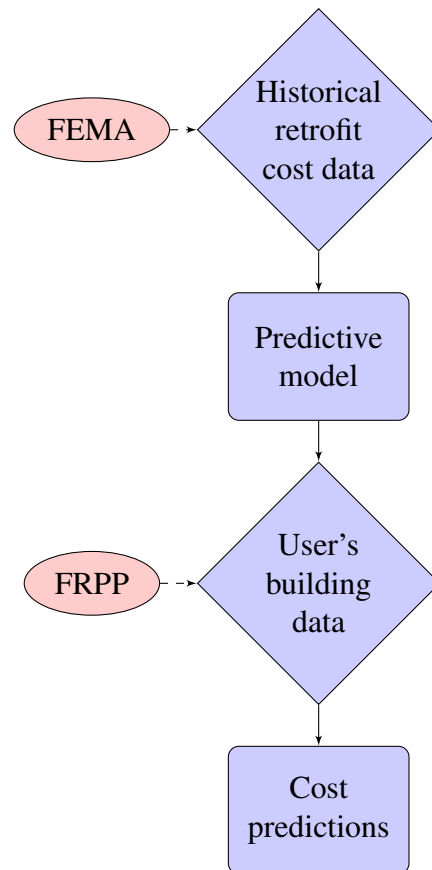
Fig. 3. Flow diagram illustrating the general methodology; data used in the present report given in red.

In this report, we apply the NIST TN 1973 [1] methodology to generate seismic retrofit cost predictions for federal buildings. Figure 3 illustrates the general approach and how we apply it:

- The main input is a reliable data set on *past* retrofit costs (in our case, the FEMA 156/157 data).

- The input data is used to "train" a predictive model.[6] In NIST TN 1973 [1], we discuss how a decision maker might choose a preditive model.

---

[6]By train a model, we mean that the input data is used to fit a linear regression model of retrofit costs that can be used to generate predictions when given new data.

- User's building data represents the building or buildings for which a user wants to predict retrofit costs (in our case, the FRPP data).

- Finally, the predictive model is used to generate cost predictions for the user's building data.

The details are given in NIST TN 1973 [1]. In Appendix A, we discuss how we modify the NIST TN 1973 [1] methodology for the present report. Although we develop the methodology in order to estimate retrofit costs for federal buildings, our approach is generalizable to any building inventory and can assist decision makers in prioritizing seismic mitigation strategies.

1.3 Guidance for decision makers

In order to help decision makers with seismic retrofit planning, a cost estimation methodology should provide inputs into the decision-making process.

Whether a retrofit project should be undertaken by a federal agency will depend on whether it is *cost effective*: is it cheaper to retrofit an existing building than to build a new one? To provide a measure of cost effectiveness, we compare retrofit costs to building replacement value in Sec. 3.6. In particular, we provide guidance on how retrofit cost predictions may be combined with estimates of building replacement value, and present cost estimates for buildings with a retrofit-replace ratio less than 1 (i.e., whether a retrofit is cost effective relative to replacing the building).

Of course, our model only predicts structural construction costs. As we discuss in Sec. 3.6, additional construction costs, as well as non-structural mitigation costs, can increase the total retrofit cost and must therefore be accounted for in decision making.[7]

---

[7]In Appendix C, we show how increasing retrofit costs can affect cost effectiveness.

2. Federal building data

In this section, we describe the data collected in the Federal Real Property Profile (FRPP) for FY15. The data is limited in the sense that building age, height, and building type, which are needed for the preditive model, are not collected as part of the FRPP. Section 2.2 discusses these limitations and Sec. 2.3 describes how we use this data to generate typical buildings that are representative of federal buildings.

2.1 The FRPP data

Data on federally owned and leased buildings is available to *federal employees only* from the General Services Administration (GSA) Federal Real Property Profile (FRPP).[8] The FRPP is a centralized database of the federal government's inventory of land, building, and structure assets. Each agency submits data on its assets annually. For this report, we obtained FRPP building data for Fiscal Year 2015 (FY15).

Table 2, in Sec. 1.1, lists average and maximum hazard, number of buildings, and square footage by agency. In addition, the table lists the percent of buildings that are owned by the reporting agency (versus those that are leased by the agency) and the percentage of buildings that are deemed historic.[9] For ease of presentation, the table is sorted by agency. The table is derived from 117 813 owned and leased buildings as reported by 25 federal agencies in fiscal year 2015.

With the exception of the DC Court Services & Offender Supervision Agency (CSOSA), the Office of Personnel Management (OPM), and the Smithsonian Institution (SI), all of whose entire building inventory is leased, and the General Services Administration (GSA), which owns about 18 % of the buildings in its inventory, agencies own between 80% to 100 % of the buildings in their inventory. Note that the agencies with the highest ownership rates also have the largest number of buildings and the largest total square footage.

Agencies with at least 10 % of their buildings deemed as historic (DOE, DOI, DOJ, NARA, NASA, VA) also own a large fraction of the buildings in their inventory. Note that the John F. Kennedy Center for the Performing Arts (JFKC) and the United States Holocaust memorial Council (USHMM) are unique in that with they have the largest percentage of historic buildings in their inventories (100 % and 50 %, respectively), but own a very small number of buildings (one and two, respectively) and the buildings are: (1) not used for typical government business; and (2) exposed to very low seismicity.

---

[8]To be precise, each agency must grant approval for access to their FRPP data. See https://www.realpropertyprofile.gov/FRPPMS/FRPP_Login for more information.

[9]Buildings deemed historic are those for the FRPP Historic Status indicator lists the building as either a National Historic Landmark (NHL), National Register Eligible (NRE), or National Register Listed (NRL). Otherwise, the building is deemed as not historic.

2.2 FRPP data limitations

In Appendix A, we discuss the predictors (building charcteristics) needed for the predictive model. While the FRPP data includes a great deal of information on each building, including an estimate of building replacement value, primary building use (e.g., office, school, warehouse), and building utilization rate, some of the characteristics needed for our model are not collected. In particular, the following key characteristics are not collected for the FRPP:

- Building age or year built;

- Number of stories or building height;

- Building construction type.

In addition, since the data is based on agencies self reporting, and some of questions are optional, characteristics such as location and building utilization rate have missing values.[10]

Nevertheless, we can obtain reasonable predictions for retrofit costs by using the data in the FRPP and making some assumptions about the data that we do not have. Federal agencies should be able to easily obtain more complete information on building characteristics and thus obtain more accurate predictions when applying our approach.

2.3 Generating typical buildings

Given the data limitations discussed in Sec. 2.2, we develop a method to generate "typical" federal buildings based on the FRPP data. We call these buildings "typical" because they are meant to be representative of the federal building inventory, but are not actual federal buildings. We should note that we do not advocate this approach as part of the methodology; rather, it is the way we circumvent the data limitations.

The approach we take is to use *proxies* for the building characteristics not collected for the FRPP (building age, height, and type). A typical building is generated as follows:

1. Consider a building, i, in the FRPP, with reported building characteristics, x_i (e.g., building area, location, and historic status).

2. Given the available information, x_i, about building i, find a suitable proxy for building age, height, and type, z_i.

3. Add typical building i', with characteristics x_i, z_i, to the typical building inventory.

4. Repeat for the next building in the FRPP.

---

[10]It is worth noting that primary building use (or "occupancy class"), one of the predictors in FEMA 156/157 that is missing from the publicly available version of the FEMA data, is an optional element collected for the FRPP.

The result is an inventory of typical buildings: a collection of buildings based on the FRPP, but with all the characteristics needed for the predictive model. As we discuss in Appendix B, we repeat this process 1000 times. This results in 1000 typical building inventories, each containing 117 813 buildings.

We combine three disparate sources to proxy for building age, height, and type, as shown in Table 3. Details are given in Appendix B.

Table 3. Data sources for proxies; see Appendix Eq. (5).

Characteristic	Depends on	Values
Age	Census Region	{Pre-1950, 1950-1970, Post-1970}[1]
Height	Census Region	{Low-Rise, Mid-Rise, High-Rise}[2]
Type |Age, Height	Hazus Region	See Appendix Fig. 4[3]

[1] Source: Census, American Community Survey (ACS) 1-year estimates for 2010.
[2] Source: Energy Information Administraion, Commerical Buildings Energy Consumption Survey (CBECS) for 1999.
[3] Source: FEMA, Hazus 2.1, General Building Stock (GBS), Tables 3A.2-3A.15 [11].

Note that the second column corresponds to *x*, the building characteristics that we observe for buildings in the FRPP. While data on Age is available at the state level, data on Height and Type are only available at the Census and Hazus Region levels, respectively.[11] We therefore use Age proxies at the Census Region level.[12]

In Sec. 3, we use the predictive model described in Sec. A to make retrofit cost predictions for each typical building inventory, which allows us to construct prediction intervals.

2.4 Other assumptions

The typical buildings generated in Sec. 2.3 include most of the desired building features for our predictive model. The only missing ingredients are (1) performance objective and (2) occupancy during construction. In practice, both of these are choice variables for a decision maker. We make some benchmark assumptions in order to obtain our predictions.

2.4.1 Performance objective

By default, we assume all buildings will be retrofitted for life safety, *LS*, which we call Option A in Table 1. We assume that some buildings already satisfy immediate occupancy, *IO*, (i.e., Veteran's Affairs hospitals), while others should be *IO* based on the International Building Code (IBC) definition of a Risk Category IV building.[13] Table 4 shows the

---

[11] Hazus groups the 50 states and the District of Columbia into three Hazus Regions (East, Midwest, West). See, for instance, Table 3C.1 in the Hazus Earthquake Model Technical Manual [11] for more information.
[12] It is worth noting that we are using *housing* age as a proxy for general building age. While imperfect, the Census data on housing age is the only source we could find that covers the entire United States.
[13] See, for instance, https://www2.iccsafe.org/cs/committeeArea/pdf_file/BU_12_113_12.pdf.

building use categories for which we assume FRPP buildings are already *IO* and therefore exclude from prediction.

Table 4. Number of buildings assumed to be IO in FRPP, by building use and agency.

Use	Acronym	Number of buildings
	HHS	37
	DHS	30
	VA	706
	DOE	7
Hospital	GSA	1
	NASA	4
	DOI	1
	DHS	283
	DOT	7077
	DOI	70
Navigation and Traffic Aids	VA	1
	DOE	8

Since the FRPP provides information on primary building use, we designate certain buildings as important for national security and thus flag them as "essential" based on a broad interpretation of the International Building Code definitions of Risk Category IV buildings. Such essential buildings may not be up to current codes and thus may require an immediate occupancy, *IO*, performance objective. The building use categories corresponding to essential buildings are given in Appendix Table 16.

Other buildings are flagged as potentially important for national security ("maybe essential"), and therefore suitable candidates for *IO*, as shown in Appendix Table 17. However, our assumption that such buildings are reasonable candidates for an *IO* retrofit is purely speculative, as such buildings do not satisfy the Risk Category IV definition.

Most of the results in Sec. 3 assume typical buildings are retrofitted for *LS*. In Sec. 3.3, however, we provide predictions assuming "essential" and "maybe essential" buildings are retrofitted for *IO*, which includes 31 732 out of 133 679 buildings in our inventory, or 23.737 %. This corresponds to Option B in Table 1.

2.4.2 Occupancy during construction

Occupants may be left in place during a retrofit, the most expensive option in terms of *structural* construction costs, or they may be vacated to another site, the least expensive

option in terms of structural costs. An intermediate option is to temporarily relocate occupants to another part of the building during a retrofit. The FRPP data provides information on building *utilization rates* for some buildings.[14] Where this information is available, we employ it as follows:

- If a building is *unutilized*, then we assume the building is *vacant* during construction;[15]

- If a building is *underutilized*, then we assume occupants are *temporarily relocated* to another location in the building.[16]

- Otherwise, we assume occupants are left *in place* during construction. This includes any buildings that may be designated as *utilized*, as well as those for which no utilization information is given.[17]

The assumption that occupants are left in place during construction may yield conservative estimates of retrofit costs. Buildings may have extra capacity for agencies to move occupants around. The default assumption of leaving occupants in place should provide an upper bound on construction costs and thus allow decision makers to assess rental rates for relocating occupants to other buildings. Of course, there are other costs associated with relocating occupants, even within the same building.

---

[14]In particular, it is only required to be reported for buildings with the following predominant uses: Offices, Hospitals, Family Housing, Dormitories and Barracks, Warehouses, and Laboratories. See the FRPP Data Dictionary for FY 2015 [12].

[15]An unutilized building is "an entire property or portion thereof, with or without improvements, not occupied for current program purposes for the accountable executive agency or occupied in caretaker status only." See [12].

[16]An underutilized building, according to [12], is "an entire property or portion thereof, with or without improvements, which is used only at irregular periods or intermittently by the accountable landholding agency for current program purposes of that agency, or which is used for current program purposes that can be satisfied with only a portion of the property."

[17]A building is utilized if it is neither unutilized or underutilized. See [12].

3. Cost estimates for typical federal buildings

In this section, we present seismic retrofit cost predictions for typical federal buildings. Our approach provides predictions for average retrofit cost by agency, as well as 95 % prediction intervals.

We generate $M = 1000$ typical building inventories, each of size $N =$117 813, using the procedure described in Sec. 2.3 and Appendix B. Note that the original FRPP inventory has $N' =$130 819 buildings. Our typical building inventories exclude buildings assumed to satisfy *IO*, as discussed in Sec. 2.4.1, as well as buildings for which no location information (and thus seismicity) is available.

The predicted average total cost to retrofit a typical federal building inventory for *LS*, Option A in Table 1, is $19.649 billion, with 95 % prediction interval ($9.709 billion, $33.781 billion). This works out to $166 782 per building, on average.

If buildings flagged as "essential" and "maybe essential" are retrofitted for *IO*, Option B in Sec. 2.4.1, then the predicted average total cost to retrofit an entire inventory is $20.122 billion, with 95 % prediction interval ($9.951 billion, $34.601 billion). These estimates assume 23.737 % of the buildings in a typical inventory are retrofitted for *IO*.

Caveats. It is worth reminding the reader that cost predictions are for *typical* buildings, as discussed in Sec. 2.3, rather than actual federal buildings. Moreover, as we discuss in NIST TN 1973 [1], these predictions are only as good as the data used to train our predictive models and subject to non-negible prediction error, especially when it comes to predicting costs for *IO*.

Moreover, structural retrofit cost estimates underestimate actual retrofit costs because they do not include:

- non-structural mitigation costs (e.g., securing bookshelves);

- other construction costs (e.g., repairing pre-existing damage).

Including such costs may double average cost estimates, as discussed in Appendix. A.2.[18] Finally, we assume all typical buildings require a retrofit. However, in practice it is unlikely that all federal buildings will require retrofitting. Federal agencies should have a better idea of how many of their buildings require retrofit.

3.1 Cost estimates by agency

We summarize estimates of average unit cost by agency, based on our typical building inventories. Table 5 presents the predicted average cost per square foot (square meter), as well 95 % prediction interval upper and lower bound estimates, by agency.[19] The table is sorted by average unit cost. On average, each agency would face between about $15 and $30 per square foot (or about $160 and $317 per square meter) for typical buildings.

---

[18]See Appendix Table 13 for summary statistics showing this pattern in the FEMA training data.

[19]These results do not include agencies with a small number of buildings, which are potentially identifiable. For example, the John F. Kennedy Center for Performing Arts (JFKC) only has one building.

Table 5. Predicted average unit cost by agency (with 95 % prediction interval) in $/sq ft (sq m) for Option A.

Agency	Lower bound	Mean cost	Upper bound
NASA	14.49 (1.35)	29.40 (316.47)	50.55 (4.70)
TVA	14.68 (1.36)	29.21 (314.42)	49.73 (4.62)
BBG	14.18 (1.32)	29.05 (312.70)	49.94 (4.64)
DOI	13.20 (1.23)	26.10 (280.95)	45.50 (4.23)
VA	12.77 (1.19)	25.91 (278.90)	44.38 (4.12)
DOS	13.07 (1.21)	25.80 (277.72)	44.22 (4.11)
FCC	12.17 (1.13)	24.76 (266.52)	42.58 (3.96)
DOT	12.01 (1.12)	24.14 (259.85)	41.53 (3.86)
DOE	12.03 (1.12)	23.62 (254.25)	41.15 (3.82)
USDA	11.75 (1.09)	22.91 (246.61)	40.21 (3.74)
DHS	11.30 (1.05)	22.90 (246.50)	39.35 (3.66)
DOJ	11.24 (1.04)	22.86 (246.07)	39.04 (3.63)
EPA	10.17 (0.94)	21.08 (226.91)	35.86 (3.33)
DOC	9.90 (0.92)	19.99 (215.18)	34.40 (3.20)
HHS	9.77 (0.91)	19.22 (206.89)	33.62 (3.12)
NSF	9.68 (0.90)	18.86 (203.01)	33.32 (3.10)
DOL	9.36 (0.87)	18.81 (202.48)	32.22 (2.99)
GSA	9.03 (0.84)	18.30 (196.99)	31.38 (2.92)
TRE	8.60 (0.80)	17.49 (188.27)	29.99 (2.79)
NARA	7.05 (0.65)	14.65 (157.70)	24.79 (2.30)

3.2 Cost estimates by region

Table 6 presents estimates of average unit retrofit cost by Census region. We do not refine these results further by agency in order to avoid unintended individual building identifications with respect to the FRPP.

Note that the South Census region faces the highest average unit cost. In fact, the South Census region also faces the highest estimated total cost, $11.72 billion. The disparity holds even when we consider cost per building. Given 37 176 buildings in the region, the average cost per building is $315 249. In contrast, the average cost per building in the West Census region, the second most expensive region, is $74 561, with an estimated total cost of $4.08 billion. One factor potentially driving this disparity may be the relative distributions of building type, age, and height in each region. Nevertheless, the results illustrate how costs can vary across the country.

14

Table 6. Predicted average unit cost by Census region (with 95 % prediction interval) in $/sq ft (sq m) for Option A.

Census region	Lower bound	Mean cost	Upper bound
South	16.17 (1.50)	32.95 (354.68)	56.12 (5.21)
West	10.98 (1.02)	20.75 (223.36)	37.26 (3.46)
Northeast	8.74 (0.81)	19.23 (207)	31.78 (2.95)
Midwest	8.67 (0.81)	17.60 (189.45)	29.90 (2.78)

3.3 Cost estimates by performance objective

Retrofit cost estimates presented thus far assume all typical buildings are retrofitted for life safety, *LS*. In this section, we present estimates assuming select buildings are retrofitted for immediate occupancy, *IO*, as discussed in Sec. 2.4.1. Table 7 presents estimates of average unit cost by performance objective.

Table 7. Predicted average unit cost by performance objective (with 95 % prediction interval) in $/sq ft (sq m) for Option B.

Performance	Lower bound	Mean cost	Upper bound
IO	13.20 (1.23)	26.22 (282.24)	45.45 (4.22)
LS	12.15 (1.13)	24.10 (259.42)	41.87 (3.89)

Note that only 15 866 buildings are flagged as "essential" and "maybe essential" and, therefore, candidates for *IO*. The average predicted cost per unit to retrofit for *IO* is only about 9 % larger than the average predicted cost per unit for *LS*. Moreover, the average cost per building for *IO* is $288 449. In contrast, in this scenario the average cost per building for *LS* is $152 484.

3.4 Cost estimates by building ownership

While an agency may not be entirely responsible for the retrofit costs of buildings it leases, cost estimates are nevertheless valuable both to the lessee agency and to the lessor, whether it is another agency or a non-government entity.

Table 8 presents predicted average cost per unit for owned and leased buildings. Note the much higher unit cost for owned buildings, which represent 87.988% of buildings in a typical inventory.

The average total cost per owned building is $151 851, while the cost per leased building is $273 294.

A possible explanation is that leased buildings are, on average, larger: 18 028 square feet (1675 square meters), compared to 7829 square feet (727 square meters) for owned buildings; that is, leased buildings are 230.262 % larger than owned buildings, on average.

15

Table 8. Predicted average cost per unit by building ownership (with 95 % prediction interval) in $/sq ft (sq m) for Option A.

Ownership	Lower bound	Mean cost	Upper bound
Owned	12.56 (1.17)	24.88 (267.81)	43.25 (4.02)
Leased	8.99 (0.84)	18.13 (195.16)	31.12 (2.89)

3.5 Cost estimates by historic status

We include an indicator for a building's historic status in Eq. (3) because retrofits for historic buildings are likely to be different than retrofits for non-historic buildings. In this section, we present cost estimates by building historic status.

Table 9 presents the predicted average cost per unit by historic status. Note that only 12.96 % of typical buildings are deemed historic.

Table 9. Predicted average unit cost by historic status (with 95 % prediction interval) in $/sq ft (sq m) for Option A.

Historic status	Lower bound	Mean cost	Upper bound
Historic	24.00 (2.23)	47.72 (513.67)	82.84 (7.70)
Not historic	10.35 (0.96)	20.53 (220.99)	35.64 (3.31)

The predicted average cost per building is $448 559 for historic buildings. For non-historic buildings, the average cost per building is $124 819 Thus, on average, a historic building costs 3.59 times as much as non-historic buildings to retrofit.

This is driven by the much greater cost per unit for historic buildings: $79.66 per square foot ($857.52 per square meter), compared with $34.27 per square foot ($368.88 per square meter) for non-historic buildings.

3.6 Cost effectiveness of retrofitting typical federal buildings

In this section, we evaluate the cost effectiveness of retrofitting typical federal buildings by comparing retrofit costs to self-reported estimates of a building's replacement value.[20] In our FRPP data, 14 904 buildings, or 12.7 %, are missing replacement value. Moreover, the distribution of replacement value is highly skewed. The bottom 1 % of reported values, or 1074 buildings, report a replacement value of no greater than $1012.72. The top 1 % of reported replacement values, or 1030 buildings, are at least $48.6 million. Such outliers

---

[20]The FRPP Data Dictionary provides the following guidance: "In determining the Replacement Value of Asset, agencies are to assume that they will replace the asset with a newly constructed asset of the same size at the same location at today's buildings standards and codes. Agencies are not to consider the current condition or need of an asset. The result is adjusted by area cost and inflation, as appropriate." Replacement value is not reported for leased properties [12].

have the potential to make our retrofit cost estimates seem more cost effective than building replacement. Table 10 provides some summary statistics for reported replacement value.

Table 10. Summary statistics: replacement value estimates reported in FRPP (in thousands of $).

Mean	1st quartile	Median	3rd quartile	s.d.
1431	69.5	260	788	4 271

Note that we exclude the top and bottom 1 % of reported values. The average reported replacement value in this range is $1.431 million.

For the remaining buildings, we compute the retrofit-replace ratio as follows. Since our model only predicts structural construction costs, we scale our estimate of average structural costs up by a constant $\alpha > 0$ in order to approximate *total* retrofit costs. Thus, we compute the retrofit-replace ratio as:

$$\text{retrofit-replace ratio} \equiv \frac{\text{structural retrofit cost} \times \alpha}{\text{replacement value}} \tag{1}$$

Based on the average ratio of total to structural costs as shown in Appendix Table 13, we set $\alpha = 2$.

Given Eq. (1), a decision maker may decide whether to pursue a retrofit based on its cost effectiveness:

- If the ratio of retrofit cost to replacement value is greater than 1, then retrofitting a building is not cost effective;

- If the ratio of retrofit cost to replacement value is less than 1, then retrofitting a building is potentially cost effective.

Table 11 presents summary statistics for the retrofit-replace ratio, computed as in Eq. (1) with $\alpha = 2$, when we assume all buildings are retrofitted for *LS*.[21]

Table 11. Summary statistics: retrofit-replace ratio for $\alpha = 2$.

Mean	1st quartile	Median	3rd quartile	s.d.	Percent cost effective
0.299	0.11	0.175	0.313	0.896	97.51 %

A total of 98 291 buildings have a retrofit-replace ratio less than 1, or 97.5% of buildings. The total cost to retrofit for *LS* only those buildings for which the retrofit-replace ratio

---

[21]In Appendix Table 18 and Fig. 5, we show that most buildings have a retrofit-replace ratio less than 1 even for large values of α.

is less than 1 is $9.28 billion, with 95 % prediction interval $4.623 billion, $16.01 billion. Recall that the total cost to retrofit *all* buildings for *LS* is $19.649 billion.[22]

Table 12 presents average unit cost by agency, as well as 95 % prediction interval estimates, for buildings with retrofit-replace ratio less than 1. Note that this excludes agencies with only a few buildings, as in Table 5.

Table 12. Predicted average unit cost by agency for buildings with retrofit-replace ratio less than 1 (with 95 % prediction interval) in $/sq ft (sq m) for Option A.

Agency	Lower bound	Mean cost	Upper bound
VA	14.02 (1.30)	28.45 (306.24)	48.71 (4.53)
NASA	14.08 (1.31)	28.44 (306.14)	48.94 (4.55)
TVA	13.94 (1.30)	27.79 (299.14)	47.32 (4.40)
DOI	13.12 (1.22)	25.93 (279.12)	45.23 (4.20)
BBG	12.46 (1.16)	25.88 (278.58)	44.33 (4.12)
FCC	12.15 (1.13)	24.68 (265.66)	42.46 (3.94)
DOT	12.24 (1.14)	24.59 (264.69)	42.33 (3.93)
GSA	11.88 (1.10)	24 (258.34)	41.18 (3.83)
DOE	11.93 (1.11)	23.41 (251.99)	40.81 (3.79)
USDA	11.93 (1.11)	23.15 (249.19)	40.78 (3.79)
DHS	11.22 (1.04)	22.75 (244.89)	39.11 (3.63)
DOJ	11.15 (1.04)	22.66 (243.92)	38.72 (3.60)
DOS	11.09 (1.03)	22.43 (241.44)	38.52 (3.58)
EPA	10.76 (1.00)	22.19 (238.86)	37.85 (3.52)
DOC	9.92 (0.92)	19.91 (214.32)	34.36 (3.19)
NSF	9.79 (0.91)	19.07 (205.27)	33.68 (3.13)
DOL	9.46 (0.88)	18.93 (203.77)	32.47 (3.02)
HHS	9.27 (0.86)	18.19 (195.80)	31.89 (2.96)
TRE	8.42 (0.78)	16.33 (175.78)	28.35 (2.63)
NARA	7.32 (0.68)	15.31 (164.80)	25.68 (2.39)

---

[22]These estimates exclude buildings with very high or very low reported replacement value. If we include all buildings, then 84.3 % of buildings have a retrofit-replace ratio less than 1, and the total cost is $15.339 billion.

4. Conclusion

In this report, we extend a methodology that we develop in NIST TN 1973 [1] for predicting seismic retrofit costs. We use a predictive model to estimate retrofit costs for typical federal buildings. Typical federal buildings are meant to be representative of the federal building inventory, but are not actual federal buildings.

We find that the predicted average total cost to retrofit typical federal buildings for life safety, *LS*, is about \$19.5 billion, or about \$167 000 per building. If we retrofit "essential" and "maybe essential" buildings, as defined in Sec. 2.4.1, for immediate occupancy, *IO*, and all other buildings for *LS*, the total cost is predicted to be about \$20 billion, on average.

To put these estimates in context, consider estimated losses from natural disasters in 2017, which included three devastating US hurricanes (Harvey, Irma, and Maria). Munich Re, an insurance company, estimates overall losses to be \$135 billion globally, with 50 % attributed to the United States.[23]

While these figures include a major earthquake in Mexico, losses in the United States have not been driven by earthquake damage since the 1994 Northridge earthquake. To evaluate potential losses from an earthquake, consider estimated Annualized Earthquake Loss (AEL), defined as the "estimated long-term value of earthquake losses to the general building stock in any single year," as estimated in FEMA P-366 [13]. For 2017, AEL is estimated to be \$6.2 billion for the United States.[24]

Finally, consider the value of construction put in place in the United States. The Value of Construction Put In Place survey, conducted by the US Census Bureau, "covers construction work done each month [in the US] on new structures or improvements to existing structures...[and] includes the cost of labor and materials, cost of architectural and engineering work, overhead costs, interest and taxes paid during construction, and contractor's profits." In March 2016, the total value of construction put in place was \$1.14 trillion. Private construction put in place totaled \$842.28 billion, while the total value of federal construction put in place was \$30.9 billion. Nonresidential private construction put in place had a total value of \$406.9 billion, while nonresidential federal construction was \$30.2 billion.[25]

Retrofitting federal buildings for seismic safety has the potential to reduce losses from natural disasters. Our estimate of the total structural construction cost to retrofit *all* buildings (excluding buildings owned by the Department of Defense) is on the same order as the annual value of federal construction. Given that it is unlikely that all buildings would be retrofitted simultaneously, or that all buildings require a retrofit, the true cost may be smaller.

---

[23]https://www.munichre.com/en/media-relations/publications/press-releases/2018/2018-01-04-press-release/index.html.

[24]Estimates in FEMA P-266 [13] are given in 2014 USD. We use the Bureau of Labor Statistics' Consumer Price Index (CPI) for all items and all urban consumers to adjust these values to 2016 USD; see https://fred.stlouisfed.org/series/CPIAUCSL/.

[25]All cited values are seasonally adjusted. See https://www.census.gov/construction/c30/c30index.html for more information and historical releases.

4.1 Caveats and future directions

Note that our predictions should not be interpreted as estimates for determining retrofit spending. Rather, they serve to illustrate order of magnitude for buidings that are generated under the assumptions described in Sec. 2.3. Building owners likely can access better information on their building inventories, and thus obtain more accurate estimates. Moreover, we assume all buildings are retrofitted, whereas building owners can evaluate whether buildings require a retrofit. In particular, if a building has already been retrofitted, then a building owner may use this information to train their own models, thereby increasing predictive accuracy.

As a reminder, we focus on predicting *structural* retrofit costs. We do not account for *non-structural* mitigation, or other construction costs associated with a retrofit. As illustrated in Appendix Table 13, total construction costs are twice as large as structural construction costs, on average, in our training data. In principle, a building owner can use our approach to estimate structural costs, and then use a rule of thumb (e.g., cost ×2) to estimate *total* costs. For a more accurate estimate, a building owner may want to analyze more current data that breaks down average construction costs. We leave this for future research.

A more important direction for future research, highlighted in NIST TN 1973 [1], is to estimate *all costs*, including *direct* costs (e.g, construction costs, financing costs, and costs to move occupants) and *indirect* costs (e.g., loss of productivity during construction, externalities to neighbors). The best approach may be to survey the construction literature in general and to obtain reasonable bounds for such costs. This is also left for future research.

References

[1] Fung J, Butry D, Sattar S, McCabe S (2017) A methodology for estimating seismic retrofit costs. National Institute of Stanards and Technology, Technical Report NIST 1973.

[2] FEMA (1994) Typical costs for seismic rehabilitation of existing buildings, volume 1: Summary. Federal Emergency Management Agency, Technical Report FEMA 156.

[3] ASCE/SEI (2014) Seismic evaluation and retrofit of existing buildings. Structural Engineering Institute of the American Society of Civil Engineers, Technical Report ASCE/SEI 41-13.

[4] FEMA (2006) Techniques for the seismic rehabilitation of existing buildings. Federal Emergency Management Agency, Technical Report FEMA 547.

[5] Holmes W, Bonowitz D, Cagley J, Poland C (2011) Standards of seismic safety for existing federally owned and leased buildings-icssc recommended practice 8. *Grant/Contract Reports (NISTGCR)-11-917-12* (RP 8).

[6] Hayes Jr JR, McCabe SL, Mahoney M (2017) Implementation guidelines for executive order 13717: Establishing a federal earthquake risk management standard. National Institute of Standards and Technology, Technical Report RP 9.

[7] FEMA (1995) Typical costs for seismic rehabilitation of existing buildings, volume 2: Supporting documentation. Federal Emergency Management Agency, Technical Report FEMA 157.

[8] Egbelakin T, Wilkinson S, Ingham J (2014) Economic impediments to successful seismic retrofitting decisions. *Structural Survey* 32(5):449–466.

[9] Egbelakin T, Wilkinson S, Potangaroa R, Ingham J (2011) Enhancing seismic risk mitigation decisions: A motivational approach. *Construction Management and Economics* 29(10):1003–1016.

[10] Cheng MY, Chen CS (2011) Optimal planning model for school buildings considering the tradeoff of seismic resistance and cost effectiveness: A taiwan case study. *Structural and Multidisciplinary Optimization* 43(6):863–879.

[11] FEMA (2012) *Hazus Earthquake Model Technical Manual*, 2nd Ed.

[12] GSA (2015) *Federal Real Property Council: 2015 Guidance for Real Property Inventory Reporting*, version 3 Ed.

[13] FEMA (2017) Hazus estimated annualized earthquake losses for the united states. Federal Emergency Management Agency, Technical Report FEMA P-366.

[14] Fung J, Butry D, Sattar S, McCabe S (2018) Cost estimates for the seismic retrofit of federally owned and leased buildings. *Proceedings of the 11th National Conference in Earthquake Engineering* (Earthquake Engineering Research Institute), p Forthcoming.

Appendix: Supplemental materials

The supplementary material included in this Appendix covers:

A. Details on the NIST TN 1973 [1] methodology and how we apply it in the present report;

B. How we generate typical buildings, including buildings assumed "essential" and "maybe essential;"

C. Supplementary results on retrofit-replace ratio estimates.

A. Overview of the methodology

In this section, we describe the methodology we develop in NIST TN 1973 [1], as well as how we apply it to the present report.

A.1 Background: NIST TN 1973

In attempting to predict retrofit costs, the most natural approach is to use information on *past* retrofit projects. However, such data is difficult to come by, especially data that includes details on building characteristics.

Fortunately, a reliable source for such cost estimates exists. FEMA 156 [2], published in 1994, and FEMA 157 [7], published in 1995, provide both a methodology for estimating seismic retrofit costs and an extensive database of retrofit cost estimates.[26]

Our predictive model is based on fitting a series of regression models on a *subset* of the FEMA data ("training" the model) and estimating the associated prediction error on the rest of the data ("testing" the model). Minimizing prediction error in the testing step depends on having a good training data set.

In NIST TN 1973 [1], we propose a model of retrofit costs as a function of the following features:

- Area: building area (in square feet);

- Age: building age (in years);

- Height: building height (in number of above and below ground stories, or in feet);

- Seismicity: peak ground acceleration with a 10 % probability of exceedance in 50 years;

- Performance Objective: life safety or immediate occupancy;

- Occupancy condition: what happens to occupants during construction (e.g., leave occupants in place during construction);

---

[26]In particular, the data can be found at https://www.fema.gov/media-library/assets/documents/30220.

- Building type: the model building type (e.g., wood frame);

- Historic: an indicator for whether the building is deemed historic.

The model we estimate in NIST TN 1973 [1] is:

$$\ln(C)_{s,p,b} = \alpha + \eta_s + \delta_p + \zeta_{sp} + \gamma_b + \beta_1 \ln(\text{Area}) + \beta_2 \ln(\text{Age})$$
$$+ \beta_3 \ln(\text{Stories}) + \beta_4(\text{Occup}) + \beta_5(\text{Historic}) + \varepsilon \tag{2}$$

where C is total construction cost per square foot and is assumed to follow a log-normal distribution (that is, $\ln(C)$ follows a normal distribution); η_s represents seismicity; δ_p represents the performance objective; ζ_{sp} represents the combined effect of seismicity and performance objective; and γ_b represents the building type.

Fig. 4, reproduced from FEMA 156 [2], provides a list of the building model types that appear in the training data, as well as their associated "building groups." Building model types are categorized into building groups in [2], presumably based on structural similarities, as we discuss in NIST TN 1973 [1]. In Eq. (2), we use building *group* rather than type.

BUILDING GROUP	MODEL	FEMA 178 BUILDING TYPES
1	URM	Unreinforced Masonry
2	W1 W2	Wood Light Frame Wood (Commercial or Industrial)
3	PC1 RM1	Precast Concrete Tilt Up Walls Reinforced Masonry with Metal or Wood Diaphram
4	C1 C3	Concrete Moment Frame Concrete Frame with Infill Walls
5	S1	Steel Moment Frame
6	S2 S3	Steel Braced Frame Steel Light Frame
7	S5	Steel Frame with Infill Walls
8	C2 PC2 RM2 S4	Concrete Shear Wall Precast Concrete Frame with Concrete Walls Reinforced Masonry with Precast Concrete Diaphragm Steel Frame with Concrete Walls

Fig. 4. Building groups and associated building types (Source: FEMA 156).

Seismicity is defined in Sec. 1.1. Additional details. A more detailed discussion of the data, including summary statistics, is given in NIST TN 1973 [1].

23

A.2 Modifying the NIST TN 1973 model

In this report, we modify Eq. (2) as follows:

- Age: we use three age categories (pre-1950, 1960-1970, and post-1970);

- Height: we use three height categories (Low-Rise, Mid-Rise, and High-Rise);[27]

- *C*: we predict *structural* rather than total cost per square foot.

As discussed in Sec. 2.2, the FRPP does not collect building age or height. Our typical buildings use proxies for age and height, which are not available on a continous scale. Thus, we train our predictive model on categorical measures of age and height.

In NIST TN 1973 [1], we focus on predicting *total* construction costs, which include costs of structural mitigation as well as additional costs triggered by the retrofit. These additional costs include:

- costs associated with compliace with the Americans with Disabilites Act of 1990;

- costs associated with removal of asbestos and other hazardous material;

- costs associated with repairing damage or deterioration;

- non-structural mitigation costs.

See [2] for more details. These additional costs vary greatly. Moreover, it is difficult to say how they correspond to costs today (e.g., compliance with ADA).

In this report, we use a predictive model for *structural* retrofit costs, which are construction costs for structural seismic mitigation that do not include the costs listed above. Moreover, construction costs do not include permits, architect and engineering fees, and financing costs. In the training data, such costs might double the total cost, as shown in Table 13. We therefore expect that structural retrofit costs underestimate actual retrofit costs.

Table 13 presents select descriptive statistics for structural and total construction costs, to illustrate how much total costs vary relative to structural costs. Costs are given in 2016 $/square foot (square meter).[28]

The ratio of total to structural construction costs is about 2.01, on average. In fact, the largest ratio of total cost to structural cost in the FEMA data is 320. Table 14 summarizes the distribution of the ratio in the FEMA data.

---

[27]Low-rise is defined as 1-3 stories; Mid-Rise as 4-7 stories; and High-Rise as 8 stories or more.

[28]We omit the minimum structural and total costs because they are equal.

Table 13. Summary statistics: structural vs total construction costs in FEMA training data in $/sqft$ ($/$sq m).

Cost	Mean	Median	Max	s.d.
total	52.13 (561.14)	28.84 (310.41)	1688.55 (18175.95)	81.95 (882.12)
structural	36.03 (387.8)	23.33 (251.17)	675.42 (7270.38)	44.74 (481.62)

Table 14. Summary statistics: ratio of total to structural construction costs in FEMA training data.

1st quartile	Mean	Median	3rd quartile	Max	s.d.
1	2.01	1.01	1.24	320	8.88

Finally, the most important modification to Eq. (2) is the model itself. In contrast to modeling retrofit costs as log-normally distributed and training an ordinary least squares (OLS) model as in Eq. (2), we train a Generalized Linear Model (GLM). In Fung et al. [14], we discuss the advantages of a GLM over the standard OLS model that we develop in NIST TN 1973 [1].

Let $Y = \ln(E[C])$, where $E[C]$ is the expected (structural) construction cost per square foot. The GLM for Y used in this report includes *categorical* measures of building age and height, as well as a regional indicator for location:

$$Y = \beta_0 + \eta_s + \delta_p + \zeta_{sp} + \gamma_b + \beta_1 \ln(\text{Area}) + \beta_2 \text{Age} \\ + \beta_3 \text{Height} + \beta_4 \text{Occup} + \beta_5 \text{Historic} + \beta_6 \text{Region} \tag{3}$$

where Region is the Census Region in which the building is located.[29] Equation (3) specifies the first part of the GLM, the model of expected cost, $E[C]$. The model links the mean, $E[C]$, to the predictors through the log function; that is, the mean is modeled as $E[C] = \exp(X\beta)$ where X is the matrix of predictors and β is the vector of parameters. The second element of the GLM is the distribution of C. Since C should be nonnegative, and the distribution of C is unlikely to be symmetric, we assume that C follows a Gamma distribution.[30]

A.3 How we apply the methodology

We use Root Mean Square Error (RMSE) as our measure of out-of-sample prediction error. Let $\hat{\beta}$ denote the coefficient vector obtained in the training step. The RMSE is estimated

---

[29]The US Census groups the 50 states and the District of Columbia into four Census Regions (Northeast, Midwest, South, and West); see https://www.census.gov/geo/reference/gtc/gtc_census_divreg.html.

[30]One advantage of this specification is that it does not assume constant variance. The variance of C_i for observation i is modeled as proportional to $V(E[C_i])$.

25

from the test set as:

$$RMSE \equiv \left(\frac{1}{m} \sum_{i=1}^{m} (\exp(\hat{\beta}^T x_i) - C_i)^2 \right)^{\frac{1}{2}} \tag{4}$$

Note that we estimate RMSE on the original scale of the response variable, C. Thus, RMSE as defined in Eq. (4) is directly interpretable as prediction error in dollars per square foot.

In NIST TN 1973 [1], we use a random training-test set split for the data: 70 % of the data is used to train the model (i.e., estimate the parameters), and the other 30 % is used to test the model (i.e., estimate prediction error). The disadvantage of this approach is that the predictive model is only trained on a subset of the data (the training set).

In this report, we use K-fold cross validation. That is, we split the data into $K = 10$ "folds," and iteratively train the models on $K - 1$ folds while using the remaining fold to estimate prediction error. The final estimate of prediction error is obtained by averaging prediction error estimates across all K folds. The advantage of this approach is that the predictive model is trained on all of the data. Given that the FEMA data is not large to begin with,[31] using all of the available data to train the model is important for improving accuracy.

Finally, Table 15 presents the results from fitting Eq. (3) to the FEMA training data.

---

[31] N=1526 for the data used in this report.

Table 15. Coefficient estimates for region, discrete model.

	Cost per sf
Area	−0.108*** (0.028)
Occupancy: TR	0.128* (0.073)
Occupancy: IP	−0.285*** (0.092)
Seismicity: M	−0.089 (0.182)
Seismicity: H	−0.180 (0.180)
Seismicity: VH	0.364** (0.176)
Performance: DC	−0.092 (0.189)
Performance: IO	0.022 (0.197)
BG: 2	−0.484*** (0.135)
BG: 3	−0.463*** (0.140)
BG: 4	0.158 (0.100)
BG: 5	0.074 (0.155)
BG: 6	−0.903*** (0.192)
BG: 7	0.223 (0.136)
BG: 8	−0.088 (0.103)
Historic	0.884*** (0.104)
Height: Mid Rise	0.178** (0.090)
Height: High Rise	0.259** (0.125)
Age: 1950-1970	0.069 (0.076)
Age: post-1970	−0.097 (0.119)
Census Region: Midwest	−0.183 (0.149)
Census Region: Northeast	−0.202 (0.164)
Census Region: South	0.465*** (0.134)
M x DC	0.072 (0.244)
H x DC	0.583** (0.250)
VH x DC	0.290 (0.228)
M x IO	0.058 (0.307)
H x IO	0.495* (0.292)
VH x IO	0.513** (0.233)
Constant	3.710*** (0.321)
Observations	1,083
Log Likelihood	−4,178.000
Akaike Inf. Crit.	8,417.000

Notes:	***Significant at the 1 percent level.
	**Significant at the 5 percent level.
	*Significant at the 10 percent level.

27

B. Additional details for generating typical buildings

For a given building in the FRPP with observed characteristics x, we sample the unobserved characteristics z, generating a "typical" building with the features x, z. In our case, the missing features of interest are $z = \{$Age, Height, Type$\}$. Unobserved characteristics are sampled conditional on observed characteristics as follows:

$$\text{Age, Height, Type}|x \sim p(\text{Type}|\text{Age, Height}, x)p(\text{Height}|x)p(\text{Age}|x) \tag{5}$$

One approach to constructing the conditional distributions on the right-hand side of Eq. (5) is to use other data sources on building age, height, and type. We combine three disparate sources to construct the distributions in equation (5), as shown in Table 3. In our case, x includes building location (e.g., Census region or Hazus region), which is why we include the Region indicator in Eq. (3). The distributions derived for Eq. (5) are therefore conditional on a single observed characteristic, *location*.

For each building i in the FRPP, with observed characteristics x_i, we generate a typical building i' with observed characteristics x_i and unobserved characteristics $z_{i'}$ sampled according to Eq. (5). The result is an inventory of typical buildings, what we call a "building profile." Using this method, we generate $M = 1000$ typical building inventories, where each inventory contains $N = 117\,813$ buildings.

The following psuedo-code summarizes the procedure we use to generate typical buildings, as discussed in Sec. 2.3.

Algorithm 1: Pseudo-code to generate typical buildings using Eq. (5)

Data: A table InputData of N buildings
Result: A database OutDB of M tables of N buildings each
initialization;
for $i = 1$ **to** N **do**
    for $j = 1$ **to** M **do**
        x ← Observed features for building i;
        height ← Sample height given x;
        age ← Sample age given x;
        type ← Sample type given height, age, x;
        OutDB $[i, j]$ ← typical building i, inventory j, has observed features x
          and unobserved features { height, age, type };
    end
end

Recall that, in our case, the input table is the FRPP data, and that the observed features that matter for sampling are Census region and Hazus region.

Tables 16 and 17 show the building use categories, as well as number of buildings by agency, for buildings flagged as "essential" and "maybe essential" in Sec. 2.4.

28

Table 16. Buildings flagged as essential in FRPP.

Use	Acronym	Number of buildings
Aviation Security Related	GSA	462
	DHS	100
	VA	3
Border/ Inspection Station	DHS	635
	GSA	170
	USDA	3
Communications Systems	USDA	737
	DHS	472
	DOT	1433
	DOI	270
	NASA	93
	DOE	106
	VA	12
	FCC	42
	BBG	23
	DOC	4
	HHS	2
	GSA	5
	EPA	1
	DOS	5
	DOJ	3
	SI	1
Data Centers	VA	13
	GSA	11
	DOE	50
	DOJ	2
	DOC	1
	DHS	1
	TRE	1
	DOI	1

Table 16. Buildings flagged as essential in FRPP. *(continued)*

Use	Acronym	Number of buildings
Land Port of Entry	DHS	73
	GSA	325
	USDA	2

Table 17. Buildings flagged as maybe essential in FRPP.

Use	Acronym	Number of buildings
	USDA	1025
	HHS	54
	DHS	658
	DOJ	209
	VA	600
	DOI	2063
	GSA	17
	NASA	71
	NSF	11
	BBG	1
	DOE	1098
	TRE	8
	DOC	11
Industrial	DOT	5
	USDA	603
	HHS	217
	VA	166
	EPA	39
	NASA	691
	TVA	2
	DOI	179
	NSF	17
	DOC	110
	DOE	1243
	DHS	13
	GSA	73
	DOT	51
	SI	1
	DOJ	12
Laboratories	DOL	6

Table 17. Buildings flagged as maybe essential in FRPP. *(continued)*

Use	Acronym	Number of buildings
Prisons and Detention Centers	DOJ	1487
	DHS	31
	DOI	32

C.   Supplementary results

Finally, recall that in Sec. 3.6 we scale up our structural retrofit cost estimates to approximate total retrofit costs, as shown in Eq. (1). In discussing our results, we set $\alpha = 2$, consistent with the average ratio of total to structural costs in the training data.

However, it is possible that this still underestimates total construction costs, or that an approximation to total construction costs does not account for other costs associated with a retrofit and, therefore, that the retrofit-replace ratio is biased toward 0.

Table 18. Summary statistics: retrofit-replace ratio for range of α.

alpha	Mean	Median	Percent cost effective
2	0.299	0.175	97.51 %
4	0.598	0.350	88.20%
6	0.897	0.525	77.22 %
8	1.196	0.700	66.19 %
10	1.495	0.876	56.45 %

Table 18 presents average retrofit-replace ratio, as well as the percentage of buildings for which a retrofit is cost effective (i.e., the retrofit-replace ratio is less than 1, for a range of α values.[32] Note that even for very large values of α, a retrofit is cost effective for a majority of buildings. For instance, for $\alpha = 8$, nearly two thirds of buildings have a ratio less than 1. Of course, the values of α are chosen simply for illustration. While total most may be 10 or more times greater than structural cost occasionally, on average this is unlikely.

Figure 5 plots the empirical cumulative distribution function (CDF) of the retrofit-replace ratio as a function of α. The figure illustrates how the proportion of buildings for which the ratio is less than 1 varies by α, with a large proportion of buildings approaching a ratio of close to 1. Note that retrofit-replace ratio values greater than 50 are omitted for ease of presentation.

---

[32]As in Sec. 3.6, these estimates exclude buildings with very high and very low reported replacement values.
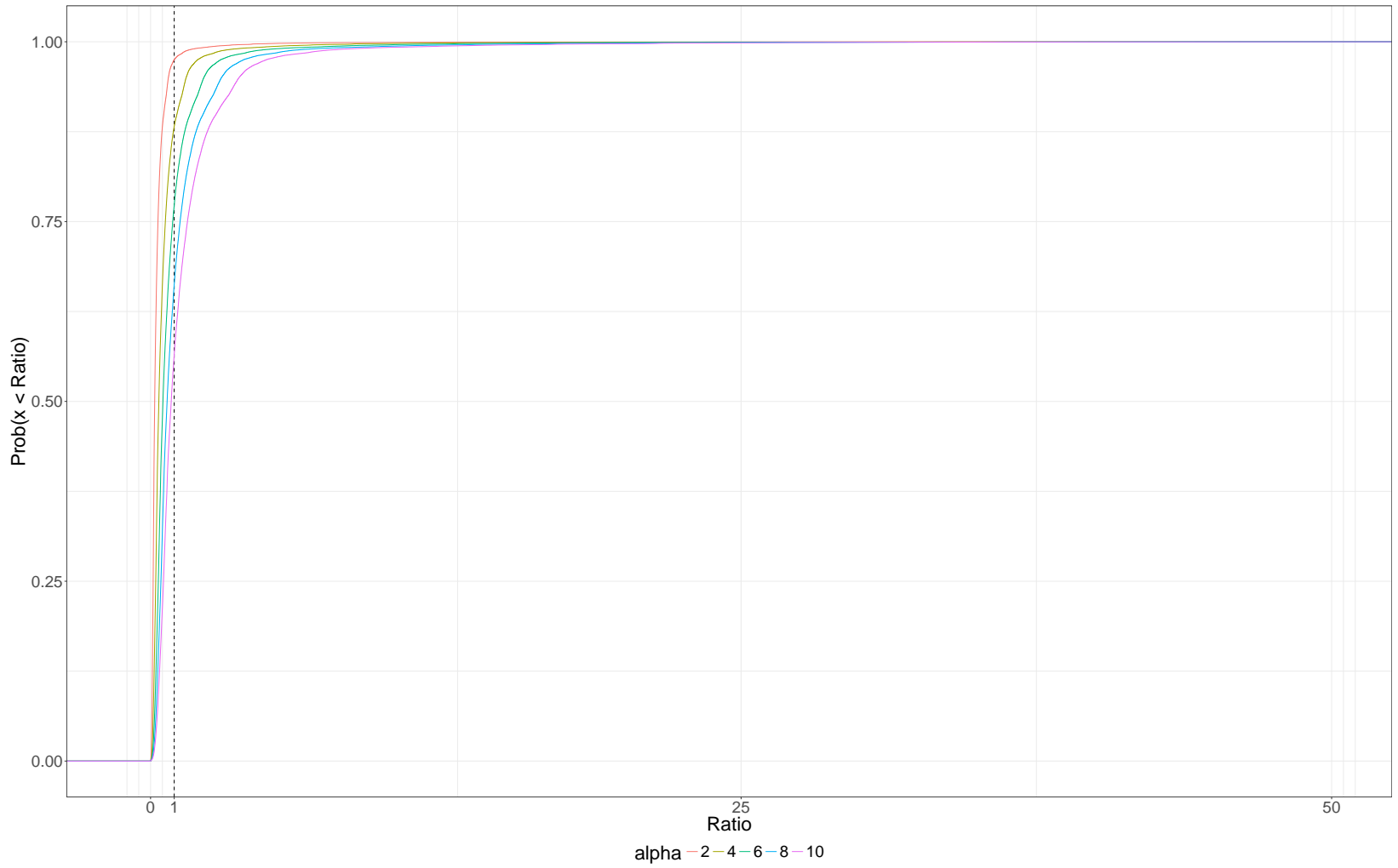
Fig. 5. Empirical CDF for retrofit-replace ratio as a function of alpha.