## **NIST Technical Note 1973**

# A Methodology for Estimating Seismic Retrofit Costs

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

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.1973



## **NIST Technical Note 1973**

# A Methodology for Estimating Seismic Retrofit Costs

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

Siamak Sattar Steven L. McCabe Earthquake Engineering Group Engineering Laboratory

This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.1973

September 2017



U.S. Department of Commerce Wilbur L. Ross, Jr., Secretary

National Institute of Standards and Technology Kent Rochford, Acting NIST Director and Under Secretary of Commerce for Standards and Technology Certain commercial entities, equipment, or materials may be identified in this document in order 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.

> National Institute of Standards and Technology Technical Note 1973 Natl. Inst. Stand. Technol. Tech. Note 1973, 40 pages (September 2017) CODEN: NTNOEF

> > This publication is available free of charge from: https://doi.org/10.6028/NIST.TN.1973

#### Abstract

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. However, while guidance on evaluation and retrofit practices is readily available, a standard approach to estimating retrofit costs does not exist. Moreover, the absence of easily obtainable estimates can make retrofits prohibitive for decision-makers.

This paper develops a cost-estimating methodology for seismic retrofits that (1) captures the essential factors that drive seismic retrofit costs, such as building construction and square footage; and (2) is reproducible using data available to decision-makers.

The methodology builds on FEMA 156 and 157, *Typical Costs for Seismic Rehabilitation of Existing Buildings*, Volumes 1 and 2. A series of regression models is fit to the data used for the FEMA reports, with the models varying in the level of data required; e.g., a decision-maker may not have information on building construction for each asset in their inventory. Thus, the trade-off from estimating retrofit costs subject to data limitations can be quantified in terms of prediction error, providing decision-makers with a set of options for estimating costs together with a measure of predictive performance. We find that a simple model, in terms of data requirements, can deliver reliable predictions.

### Key words

Building economics; earthquake risk reduction; retrofit; resilience.

#### 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**

1	Intr	bauction	1
	1.1	Seismic Risk to Federally Owned and Leased Buildings	1
	1.2	Existing Literature and Alternative Methodologies	5
	1.3	Applying the Methodology to Federal Buildings	6
2	Cost	Estimation: Methodology	7
	2.1	The FEMA 156/157 Retrofit Cost Data	7
		2.1.1 Using the FEMA Data to Estimate Current Retrofit Costs	9
		2.1.2 The FEMA 156/157 "Super Database"	11
	2.2	The FEMA 156/157 Linear Regression Methodology	13
	2.3	Extending the FEMA 156/157 Model	14
	2.4	Summary of the General Methodological Approach	16
3	Cost	Estimation: Main Results	17
3	<b>Cos</b> t 3.1	Estimation: Main Results FEMA 156/157 Replication Results	<b>17</b> 17
3	<b>Cos</b> 3.1 3.2	Estimation: Main Results FEMA 156/157 Replication Results Prediction Error and Overfitting on the Synthetic Data	<b>17</b> 17 18
3	Cost 3.1 3.2 3.3	Estimation: Main Results FEMA 156/157 Replication Results Prediction Error and Overfitting on the Synthetic Data Predicting Cost Using the Main Model and Its Extensions	<b>17</b> 17 18 22
3	Cost 3.1 3.2 3.3 3.4	Estimation: Main Results FEMA 156/157 Replication Results Prediction Error and Overfitting on the Synthetic Data Predicting Cost Using the Main Model and Its Extensions How Data Limitations Affect Prediction Error	<b>17</b> 17 18 22 24
3	Cost 3.1 3.2 3.3 3.4 Con	Estimation: Main Results FEMA 156/157 Replication Results Prediction Error and Overfitting on the Synthetic Data Predicting Cost Using the Main Model and Its Extensions How Data Limitations Affect Prediction Error clusion	<b>17</b> 17 18 22 24 <b>28</b>
3	Cost 3.1 3.2 3.3 3.4 Con 4.1	Estimation: Main Results FEMA 156/157 Replication Results Prediction Error and Overfitting on the Synthetic Data Predicting Cost Using the Main Model and Its Extensions How Data Limitations Affect Prediction Error clusion Limitations	<b>17</b> 17 18 22 24 <b>28</b> 28
3	Cost 3.1 3.2 3.3 3.4 Con 4.1 4.2	E Estimation: Main Results FEMA 156/157 Replication Results Prediction Error and Overfitting on the Synthetic Data Predicting Cost Using the Main Model and Its Extensions How Data Limitations Affect Prediction Error clusion Limitations Guidance and Future Directions	<ul> <li>17</li> <li>17</li> <li>18</li> <li>22</li> <li>24</li> <li>28</li> <li>28</li> <li>29</li> </ul>
3 4 Re	Cost 3.1 3.2 3.3 3.4 Con 4.1 4.2	Estimation: Main Results FEMA 156/157 Replication Results Prediction Error and Overfitting on the Synthetic Data Predicting Cost Using the Main Model and Its Extensions How Data Limitations Affect Prediction Error clusion Limitations Guidance and Future Directions ces	<ul> <li>17</li> <li>17</li> <li>18</li> <li>22</li> <li>24</li> <li>28</li> <li>29</li> <li>30</li> </ul>

## List of Tables

Table 1	Hazard, buildings, and area by agency (Source: FRPP).	4
Table 2	Summary statistics for SRCE data: cost and select building characteristics	
	(N = 1083).	9
Table 3	Proportions of each seismicity-performance objective pair in the SRCE	
	data (N = $1083$ ).	10
Table 4	MSE of predictions, raw vs synthetic data.	18
Table 5	MSE of predictions for Canadian data, from training on raw vs synthetic	
	data.	22
Table 6	MSE relative to base model.	23
Table 7	MSE relative to base model on Canada data.	24
Table 8	MSE subject to data limitations, relative to main model.	25
Table 9	Building groups and types in SRCE data.	35
Table 10	Summary statistics for SRCE data: building characteristics used in ex-	
	tended models ( $N = 1083$ ).	35
Table 11	Replication results: building groups 1 - 4	38
Table 12	Replication results: building groups 5 - 8	39
Table 13	Coefficient estimates from training main model and variations	40

# List of Figures

Fig. 1	Peak ground acceleration (pga) with a 10 % probability of exceedance in	
	50 years for the continental United States (Source	2
Fig. 2	Federal building density map for the United States	3
Fig. 3	County-level pga, weighted by 2010 Census tract population	12
Fig. 4	Prediction error across true values (vertical line is the true mean), for raw	
	and synthetic data	19
Fig. 5	Prediction across true values (vertical line is the true mean), for raw and	
	synthetic data	20
Fig. 6	Prediction error distributions	21
Fig. 7	Prediction errorm with and without age, stories (vertical line is the true	
	mean)	26
Fig. 8	Prediction error with and without building group (vertical line is the true	
	mean)	27
Fig. 9	Seismicity-performance objective interaction term in FEMA 156/157 (Source	ce:
	FEMA 156).	32
Fig. 10	ATC-3 seismicity map (Source: FEMA 156).	33
Fig. 11	Building groups and associated building types (Source: FEMA 156).	34
Fig. 12	Super Database Algorithm (Source: FEMA 157).	36
Fig. 13	Original regression results (Source: FEMA 157).	37

#### Glossary

DC: Damage Control

FEMA: Federal Emergency Management Agency

FRPP: Federal Real Property Profile

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

IEBC: International Existing Building Code

*IO*: Immediate Occupancy

NEHRP: National Earthquake Hazards Reduction Program

LS: Life Safety

*pga*: Peak ground acceleration (with a 10 % probability of exceedance in 50 years), in % g

RP-8: Recommended Practice 8 [1]

SRCE: Seismic Rehabilitation Cost Estimator

USD: United States Dollars

#### 1. Introduction

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<sup>1</sup> However, while guidance on evaluation and retrofit practices is readily available (e.g., ASCE/SEI 41-13 [2] and FEMA 547 [3]), a standard approach to estimating retrofit costs does not exist. Moreover, the absence of easily obtainable estimates can make retrofits prohibitive for decision-makers.

In this report, we develop a cost-estimating methodology for seismic retrofits that (1) captures the essential factors that drive seismic retrofit costs, such as building construction and square footage; and (2) is reproducible using data available to decision-makers.

Our methodology builds on FEMA 156 [4] and 157 [5], *Typical Costs for Seismic Rehabilitation of Existing Buildings*, Volumes 1 and 2, respectively—hereafter FEMA 156/157. We fit a series of regression models to the data used for the FEMA reports, with the models varying in the level of data required; e.g., a decision-maker may not have information on building construction for each asset in their inventory, or the decision-maker may have an expectation for duration of construction. In this way, we can quantify the trade-off (in terms of prediction error) from estimating retrofit costs subject to data availability, providing decision-makers with a set of options for estimating costs together with a measure of predictive performance. We find that a simple model, in terms of data requirements, can deliver reliable cost predictions.

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.

#### 1.1 Seismic Risk to Federally Owned and Leased Buildings

Significant earthquake risk—the potential for damage from an earthquake—threatens federal buildings and, consequently, a functioning federal government. To get a picture of seismic risk to federal buildings, Fig. 1 maps peak ground acceleration (*pga*) with a 10 % probability of exceedance in 50 years for the continental United States.<sup>2</sup>

Figure 2 maps the concentration of federally owned and leased buildings in the continental United States. Note the overlap of high building density with high seismic hazard in many parts of the country, particularly in the west coast.

Table 1 provides a more detailed picture of seismic risk to federal buildings. The table presents summaries of seismic hazard (pga, as discussed above), as well as total number of buildings and total area by agency.<sup>3</sup>

<sup>&</sup>lt;sup>1</sup>Typically the International Existing Building Code (IEBC).

<sup>&</sup>lt;sup>2</sup>We discuss the use of this measure to capture seismic hazard (*seismicity*) in Section 2.1.

<sup>&</sup>lt;sup>3</sup>Federal building data in Fig. 2 and Table 1 is based on 130 819 buildings from the Federal Real Property (FRPP) for Fiscal Year 2015 (FY15). See https://www.realpropertyprofile.gov/FRPPMS for more information.









ω

Reporting agency	Mean pga: %g	Max pga: %g	Number of buildings	Area: 1000 sq ft (sq m)
Department of Homeland Security	9.11	80	9563	52 258 (4855)
Department of Energy	9.09	60	10 682	115 430 (10 723)
Tennessee Valley Authority	8.88	60	2442	28 317 (2631)
Department of Agriculture	8.31	80	24 193	56 657 (5263)
National Aeronautics And Space Administration	8.29	80	2700	47 202 (4385)
Federal Communications Commission	8.18	40	47	108 (10)
Department of the Interior	7.59	80	43 127	103 228 (9590)
Environmental Protection Agency	6.91	30	172	4215 (392)
Department of Commerce	6.67	60	614	8937 (830)
Department of Labor	6.67	60	2371	25 464 (2366)
Department of Veterans Affairs	6.17	60	7970	170 361 (15 827)
Department of the Treasury	5.93	60	106	6371 (592)
Department of Transportation	5.90	80	11 088	25 733 (2391)
General Services Administration	5.64	60	8628	427 270 (39 693)
Broadcasting Board of Governors	5.60	15	30	137 (13)
Department of Justice	5.30	40	3879	70 434 (6543)
National Archives and Records Administration	4.70	30	27	4910 (456)
National Science Foundation	3.35	30	208	1274 (118)
Department of Health and Human Services	2.99	25	2790	35 738 (3320)
Department of State	2.59	30	140	1437 (133)
Smithsonian Institution	2.58	5	25	1595 (148)
Office of Personnel Management	2.50	3	4	75 (7)
DC Court Services & Offender Supervision Agency	2.00	2	10	239 (22)
John F. Kennedy Center for the Performing Arts	2.00	2	1	1500 (139)
United States Holocaust Memorial Council (Holocaust Museum)	2.00	2	2	320 (30)

Table 1. Hazard, buildings, and area by agency (Source: FRPP).

#### **1.2** Existing Literature and Alternative Methodologies

The most obvious approach to estimating retrofit costs is to hire an consulting engineering or construction firm with the appropriate expertise. In principle, a decision-maker can obtain better estimates with access to better data. In particular, if a decision-maker has the time and money, an engineering consulting firm can provide precise estimates for each building in his inventory.

For a federal agency with hundreds, if not thousands, of buildings in its inventory, this is time consuming and prohibitively costly. Absent the availability of such *primary* data, a decision-maker who wants to estimate retrofit costs must rely on *secondary* data. Unfortunately, secondary data on seismic retrofit costs are not readily available, either for free or for purchase through a vendor.

The existing literature on estimating seismic retrofit costs reflects these inherent data challenges. The literature itself is fairly small, relying on a handful of datasets collected specifically for research. These few data sets offer the most realistic option for using secondary data to estimate seismic retrofit costs. Methodologically, the papers largely predict retrofit costs using some form of regression analysis.

FEMA 156/157 [4, 5] introduce three alternative methodologies for estimating retrofit cost, the most sophisticated being a linear regression model. The methodologies are applied to a unique database collected specifically for the reports. In Section 2.1, we discuss the data in more detail.

The FEMA 156/157 data appears to be the only large database of seismic retrofit costs for the United States that is readily available. Lew [6], for instance, applies the FEMA methodology to estimate retrofit costs for ten existing buildings owned by FEMA. Note that the application of this methodology *implicitly* depends on the FEMA 156/157 data, as the cost estimates are obtained using the regression coefficients from the original report.

In a series of papers, Jafarzadeh et al. [7–9] collect and analyze a database on retrofit costs for 158 public schools in Iran. Jafarzadeh et al. [9] provides the associated data (in the appendix), as well as a detailed discussion of the data collection effort and a description of the data.

Jafarzadeh et al. [7] analyzes the data using standard linear regression techniques, while Jafarzadeh et al. [8] applies artificial neural networks to predicting costs, a non-parametric regression approach. The main objective of the papers is to explore which predictors matter most for retrofit cost (the former) and the parameterizations that are likely to minimize prediction error (the latter).

More recently, Nasrazdani et al. [10] collected their own database of 167 school retrofits in Iran. The authors use Bayesian linear regression in order to predict retrofit costs. The main result is that the increase in lateral strength, together with pre-retrofit building *value*, are the most important predictors of retrofit cost.

Chen and Huang [11] analyze retrofit costs and *duration* for schools in Taiwan. The authors also compare linear regression to neural networks. Note that, unlike the papers by Jafarzadeh et al., these are retrofits that occur due to *reconstruction* following earthquake damage.

Hopkins and Stuart [12] adopt a different methodological approach, with a very different goal. They analyze data on buildings throughout New Zealand in order to conduct a benefit-cost analysis. To this extent, they collect data on buildings, including seismicity and floor area, and compute retrofit costs based on an undisclosed formula (likely based on typical engineering assessments). Thus, retrofit costs are essentially taken as *given* and the objective is to use them as inputs in computing a benefit-cost ratio.

The main takeaways from the existing literature are that the data used is either: (1) not available to the public; or (2) not applicable to construction practices in the United States. The central motivation for using the FEMA 156/157 data is that it is a very detailed data set of cost estimates, compiled from engineering consulting firms, for a large inventory of buildings, and is available to the public. Moreover, the data is most applicable to federal buildings as a majority of the buildings built in the United States.

Updating the FEMA 156/157 data is desirable, but would be a tremendous effort. Thus, we propose that a decision-maker can use this data, together with the approach developed in this report, in order to easily obtain retrofit cost estimates. Our approach, introduced in Section 2.3, relies on this data and builds upon the FEMA 156/157 methodology. Our methodology is designed to be:

- 1. **Plausible**, in the sense that the model captures the essential elements of the cost to seismically retrofit an existing building;
- 2. **Tractable**, in the sense that both data and computational requirements are reasonable, and thus, e.g., a decision-maker at a federal agency can use the approach;
- 3. **Generalizable**, in the sense that the approach applies to non-federal buildings as well.

Short of collecting primary data, we believe this offers a reasonable alternative for obtaining retrofit cost estimates.

#### **1.3** Applying the Methodology to Federal Buildings

In a forthcoming report, we apply the methodology developed in this report to estimating retrofit costs for federally owned and leased buildings. The objective is to obtain cost estimates for each reporting agency in the FRPP data in order to provide federal agencies with some guidance for retrofit decisions. Federal agencies can directly apply our methodology themselves and potentially obtain more accurate estimates.

#### 2. Cost Estimation: Methodology

Retrofit costs vary with a building's unique characteristics; e.g., building construction type, age, and location. Collecting such data is time consuming and expensive, and it is unlikely to be collected for large building inventories unless done so specifically for a large construction project such as a seismic retrofit.

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 [4], published in 1994, and FEMA 157 [5], published in 1995, provide both a methodology for estimating seismic retrofit costs and an extensive database of retrofit costs.

#### 2.1 The FEMA 156/157 Retrofit Cost Data

FEMA 156/157 [4, 5] update a 1988 FEMA report on the seismic rehabilitation of existing buildings. The first edition introduced a simple methodology for estimating retrofit costs, based on a sample of cost estimates for 614 buildings. However, as noted in FEMA 156 [4], "most of the original data points were derived from rather limited studies." For instance, the data predominantly consisted of unreinforced masonry (URM) buildings, which are known to be at much higher risk than other building types.

The second edition not only expands the methodology, but also improves the quality of the data on which the analysis is based. This is accomplished in two ways:

- The collection of a larger sample (2088 cost estimates);
- A quality control process for the data collection.

Moreover, significant advances in retrofit construction are cited in the intervening period, highlighting the importance of updating data on retrofit costs.

The cost estimates collected for FEMA 156/157 [4, 5] represent actual or estimated ("by an experienced design professional") retrofit project costs for buildings in the United States and Canada. The data collected includes a detailed survey of building characteristics (e.g., age, location, and building type). Importantly, the data is "validated" through follow-ups with survey respondents, and assigned a quality control rating (or "quality factor") to ensure that each response is "objective and reliable." Cost estimates with low quality control ratings are removed from the database, with the intent to improve the overall quality of the data.

In addition to building characteristics, retrofit costs are influenced by two other, important factors: *building seismicity*, the level of seismic hazard exposure; and *performance objective*, building anticipated performance in a seismic event.

Building seismicity is assigned using a map featured in the the 1991 edition of the *NEHRP Recommended Provisions for the Development of Regulations of New Buildings* [13]. Based on our examination of the map as reproduced in FEMA 156 [4], the measure

of seismicity used in the report appears to be "Effective Peak Acceleration" (EPA), developed in the report Applied Technology Council (1978) (ATC-3), a measure loosely related to peak ground acceleration.<sup>4</sup> The FEMA model, discussed in Section 2.2, depends on four seismicity categories based on EPA: *Low* (*L*), *Medium* (*M*), *High* (*H*), and *Very High* (*VH*).<sup>5</sup> In Section 2.1.1, we discuss how seismicity can be updated using current USGS seismic hazard maps.

The performance objective categories represented in the FEMA data are, as defined in FEMA 156 [4]:

- Life Safety (*LS*): Allows for unrepairable damage as long as life is not jeopardized and egress routes are not blocked.
- Damage Control (*DC*): Protects some feature or function of the building beyond life-safety, such as protecting building contents or preventing the release of toxic material.
- Immediate Occupancy (*IO*): Allows only minimal post-earthquake damage and disruption, with some nonstructural repairs and cleanup done while the building remains occupied and safe.

The loose definitions provided in FEMA 156 [4] suggest that *DC* is equivalent to *LS* plus *nonstructural* improvements, and would more likely coincide with current definitions of *LS* (e.g., in RP-8 [1]. In Section 3, we compare retrofit cost estimates for each performance objective.

A version of the data used in FEMA 156/157 is available online through FEMA's Seismic Rehabilitation Cost Estimator (SRCE),<sup>6</sup> a web-based app for estimating retrofit costs that is no longer maintained.

The most important discrepancy between the SRCE data and the data cited in FEMA 156/157 [4, 5] is that the data on *occupancy class* (whether a building is an office or a warehouse, for instance) is missing. This is an important component of the FEMA model, as discussed in Section 2.2.

Several other, less consequential, discrepancies are worth noting. First, the SRCE data is normalized to 1993 US dollars (USD) for typical construction costs in California. In contrast, FEMA 156/157 normalize the data using typical construction costs in Missouri, in 1993 USD. This is not critical to our results, but we do caution strict comparisons between the original FEMA 156/157 results and our replication attempt in Section 3.1.

<sup>6</sup>https://www.fema.gov/media-library/assets/documents/30220

<sup>&</sup>lt;sup>4</sup>We were unable to obtain the 1991 NEHRP report. The map published in FEMA 156 [4] defines seismicity based on a coefficient, *Aa*, and states that the map was "prepared by Applied Technology Council." A visit to https://earthquake.usgs.gov/hazards/learn/technical.php reveals that the *Aa* coefficient originates in ATC-3, and quotes ATC-3, stating that *Aa* coefficients "do not at present have precise definitions in physical terms."

<sup>&</sup>lt;sup>5</sup>In particular, FEMA 156/157 [4, 5] and the ATC-3 map of EPA suggest each building is assigned a *zone*, ranging from zone 1 (EPA < 0.05) through zone 7 (EPA >= 0.4), with zones and 1 and 2 belonging to *L*; zones 3 and 4 belonging to *M*; zones 5 and 6 belonging to *H*; and zone 7 being *VH*. See Fig. 10.

156 [4], this category is defined as "rehabilitating parts or portions of a structure without considering the entire structure for life-safety or greater performance." Thus, *RR* is *weaker* than *LS* and is not relevant to estimating current retrofit costs. Moreover, *RR* is not discussed in the reports and accounts for only 0.031 % of the data. Finally, the SRCE data consists of 1978 observations, compared to 2088 in the original reports. While it does not appear to be a substantial loss of data in *absolute* terms, it does represent a 5 % loss of data relative to the original report. Nevertheless, the SRCE data provides a tremendous amount of information and thus offers a reasonable starting point for estimating retrofit costs based on secondary data. In order to provide a picture of the SRCE data, summary statistics for some of the key variables are about in Tables 2 and 3. Table 2 summarizes the date on post (atructure).

In order to provide a picture of the SRCE data, summary statistics for some of the key variables are shown in Tables 2 and 3. Table 2 summarizes the data on cost (structural construction cost in USD per square foot), age (in years), area (in thousands of square feet), and stories (above ground plus below ground stories, if any). Note, in particular, the large variation in costs per square foot. The table also includes summary statistics for building height (in feet), duration of retrofit construction (in months), and peak ground acceleration (in %g where g is the acceleration of gravity),<sup>7</sup> which we discuss in Section 2.3.

Another discrepancy is that a fourth performance category appears in the SRCE data: Risk Reduction (RR). Based on the sample survey presented in the Appendix of FEMA

Stat	Cost: USD/sq ft (sq m)	Age: years	Area: 1000 sq ft (sq m)	Stories
Min	0.3 (3.1)	0	0.2 (0)	1
Mean	29.7 (320.1)	44	64.8 (6)	3
Median	16 (172.4)	40	25 (2.3)	2
Max	1011.4 (10887)	153	1430.3 (132.9)	38
Std dev	47.3 (509)	22	109.1 (10.1)	3

Table 2. Summary statistics for SRCE data: cost and select building characteristics (N = 1083).

Table 3 summarizes the other two variables that are critical to estimating retrofit costs: building seismicity, as defined in ATC-3 and discussed above, and the performance objectives of *LS*, *DC*, and *IO*. The table presents the frequencies of each seismicity-performance objective pair. Note that the data used in Tables 2 and 3 exclude those observations with performance objective *RR*.

Descriptive statistics for other characteristics, including building type, as well as other information from FEMA 156/157, can be found in Appendix 4.2.

#### 2.1.1 Using the FEMA Data to Estimate Current Retrofit Costs

Our methodology, discussed in Section 2.3, is based on fitting a series of regression models on a *subset* of the SRCE data (the "training" step) and estimating the associated prediction

<sup>&</sup>lt;sup>7</sup>The acceleration of gravity is 9.8(m/s<sup>2</sup>). See https://earthquake.usgs.gov/learn/glossary/ for a complete definition of g.

	LS	DC	ΙΟ
L	0.009	0.038	0.021
Μ	0.026	0.100	0.048
Н	0.160	0.026	0.020
VH	0.371	0.112	0.069

Table 3. Proportions of each seismicity-performance objective pair in the SRCE data (N = 1083).

error on the rest of the data (the "testing" step). Minimizing prediction error in the testing step depends on having a good training data set.

Although the SRCE data is the most realistic option for training a model to predict retrofit costs for the United States, the data is rather dated.<sup>8</sup> Updating the database would be arduous. Nevertheless, with proper adjustments the data can provide reasonable cost estimates.

In order to estimate *current* retrofit costs, two considerations are the adjustment of costs to current dollars and the adjustment of seimisicy to reflect current seismic hazards.

**Cost adjustment.** The SRCE data includes raw costs, as well as time and location adjustment factors. The time factor normalizes costs to 1993 USD. The location factor normalizes costs to California construction costs, to account for regional differences in construction and materials costs.

Time adjustment factors in FEMA 156/157 [4, 5] are based on the Engineering News Record (ENR) 20-city average Building Construction Index (BCI).<sup>9</sup> Location adjustment factors are constructed from the Means Construction Guide, another index of construction costs.<sup>10</sup>

Since the SRCE cost data is already normalized to a single state, we create an annual California BCI by averaging the BCI for Los Angeles and San Francisco. Given costs (per square foot) in California for 1993,  $C_{CA,1993}$ , the average construction cost in the United States for 2016 is:

$$C_{US,2016} = \frac{BCI_{US,2016}}{BCI_{CA,1993}} C_{CA,1993}$$
(1)

where  $BCI_{CA,1993}$  is the California BCI for 1993 and  $BCI_{US,2016}$  is the 20-city average BCI for 2016. Thus, we normalize costs to be average *national* costs in 2016 USD.

**Seismicity adjustment.** The ATC-3 map based on EPA is outdated. Moreover, if a decision-maker wants to make predictions for a specific building inventory, it is difficult to assign seismicity based on this map.

<sup>9</sup>ENR's BCI is an index that tracks labor costs and the costs of three construction components (steel, cement, and limber) in each of twenty cities in the United States. See http://www.enr.com/economics/faq for more.

<sup>&</sup>lt;sup>8</sup>Only 9.45 % of the buildings are Canadian.

<sup>&</sup>lt;sup>10</sup>The Means Construction Guide, now called the RSMeans Construction Cost Index, tracks construction costs in 150 cities. Location adjustment factors are computed by averaging city indexes at the state level.

Given that EPA is correlated with peak ground acceleration, using current seismic hazard maps based on the latter is a reasonable way to assign seismicity to a building. We focus on peak ground acceleration with a 10 % probability of exceedance in 50 years (pga), because it appears most closely related to EPA, but a user can choose whatever measure they like. The objective is to have some criteria for distinguishing high seismic-hazard buildings from low seismic-hazard buildings.

Given approximate location information (e.g., address or latitude and longitude), assigning *pga* using current USGS seismic hazard maps is fairly straightforward.

Location information, however, may be imprecise or missing altogether. In order to assign seismicity, it is worth exploring a suitable *aggregate* measure. For instance, if we know the building's county (and state), we can assign the building some measure of county-level seismicity.

We propose using a population weighted-average. The motivation for using a population weighted-average is that population density is a proxy for *building* density, which increases seismic risk by increasing exposure to the hazard. Of course, a decision-maker may reject this assumption, and a simple average may suffice. As mentioned earlier, the goal is to distinguish buildings that face high seismic hazard from those that face low seismic hazard.

Suppose we have population information at the Census tract level. If we overlay our pga map on a Census tract map, the Census tracts will be split according to the pga contours (that is, the contours in Fig. 1). Then given population information for each tract, we compute a weighted average of pga at the desired level (e.g., county), using the population in each sub-tract as weights for the pga.

Figure 3 maps county-level *pga*, weighted by 2010 Census tract population, for the continental United States. The seismic hazard map is the same as the one in Fig. 1, based on the 2014 USGS long-term model.

#### 2.1.2 The FEMA 156/157 "Super Database"

In order to leverage the quality of information in the data, the analyses in FEMA 156/157 [4, 5] are based on a *synthetic* data set derived from the original database. The result is a "super database" consisting of 83 000 observations. The quality control ratings in the original data are used as weights in simulating new cost samples.

The main motivation for creating a synthetic data set appears to be to increase the sample size, with the intent of increasing the quality of the observed cost estimates. On the surface, the procedure bears some similarity to standard re-sampling techniques such as the bootstrap, except that it is applied *before* the analysis.<sup>11</sup>

<sup>&</sup>lt;sup>11</sup>The procedure is referred to as a generic "Monte Carlo Simulation" in FEMA 156/157 [4, 5].



Fig. 3. County-level *pga*, weighted by 2010 Census tract population.

Note, however, that by design the synthetic data may be highly correlated and lead to overfitting in the training step based on the over-representation of "high quality" cost estimates.<sup>12</sup> In Section 3.2, we provide empirical results that suggest overfitting is a real concern. Thus, our main results are based on the original *raw* data.

#### 2.2 The FEMA 156/157 Linear Regression Methodology

FEMA 156/157 [4, 5] present three methods for estimating costs, differing in the amount of information needed to compute estimates (and thus in their general applicability).

- Option 1: mean cost by building type;
- Option 2: mean cost by building type, conditional on performance objective and seismicity;
- Option 3: linear regression (predicting the conditional mean).

FEMA 156 also provide a clear definition of cost as being "mean structural cost of the seismic rehabilitation of a building...and does not include the cost of replacing architectural finishes" ([4], page 1-3).

The most general method (their "Option 3") is based on the following linear regression model. The structural cost per square foot, C, of retrofitting a particular building, conditional on observable building characteristics and target seismic performance category (the *performance objective*), is log-normally distributed. That is,  $\ln(C)$  is a linear function of observable (and unobservable) characteristics:

$$\ln(C)_{sp} = \alpha + \zeta_{sp} + \beta_1 \ln(\text{Area}) + \beta_2 \ln(\text{Age}) + \beta_3 \ln(\text{Stories}) + \beta_4(\text{Occupancy class}) + \beta_5(\text{Occupancy condition}) + \varepsilon$$
(2)

where *s* denotes the seismicity of the building, *p* denotes the performance objective of the retrofit, and  $\varepsilon$  is the unobserved error term with normal distribution,  $\varepsilon N(0, \sigma^2)$ .

The term  $\zeta_{sp}$  is the combined fixed effect (interaction) of seismicity and performance. The FEMA base model considers the interaction between seismicity and performance objective, but not their individual effects. This captures the possibility that, for instance, the expected cost to retrofit for life safety is higher the larger the seismic risk. The parameters of interest are thus { $\alpha$ ,  $\zeta_{sp}$ ,  $\beta_1$ ,..., $\beta_5$ }.

The model given in equation (2) is fitted *separately by building group*, *b*. The 15 building construction types in the data are assigned to one of eight building groups, as shown in the Appendix, Fig. 11—presumably based on similarities.<sup>13</sup>

The main motivation for this approach is to allow all coefficients to vary by building group, b. Another advantage of this approach is that it does not make assumptions about

<sup>&</sup>lt;sup>12</sup>Overfitting refers to the situation in which a trained model does a very good job of explaining the observed (training) data (e.g., in terms of prediction error), but does very poorly in explaining *new* data.

<sup>&</sup>lt;sup>13</sup>Strictly speaking, all three methods in FEMA 156/157 [4, 5] use building group rather than type.

the relationship *between* building groups; that is, no assumption is made on the relationship between  $\varepsilon_b$  and  $\varepsilon_{b'}$  for  $b \neq b'$ .<sup>14</sup>

Note the additional predictors in equation (2). *Occupancy class* describes the primary building use (e.g., office, warehouse, or hospital). *Occupancy condition* is defined as follows:

- In-place (*IP*): work is scheduled around normal hours of occupancy;
- Temporarily removed (*TR*): occupants are moved to another room in building during construction; and
- Vacant (V): building completely vacated during construction.

Completely vacating the building is the most expensive option, while leaving occupants in-place is the lowest cost option.

A natural question is whether it is necessary for any, let alone all, coefficients to vary by building group. While this modeling approach does not make as strong assumptions about the error term as does pooling the data (perhaps with a fixed effect  $\gamma_b$  for building group), it does fail to capture potential "clustering" or other structure between groups. Moreover, it is likely that at least *some* effects are common across building groups (e.g., the effect of age on project cost). In the next section, we explore variations on the FEMA base model.

#### 2.3 Extending the FEMA 156/157 Model

If we accept the premise that the FEMA model is a reasonable benchmark, given the paucity of retrofit cost estimates and data, it is natural to ask if we can do better. Thus, we evaluate whether a better performing predictive model can be obtained. The metric we will use to compare our models to the FEMA model is prediction error (that is, the Mean Squared Error of predictions on our test set). See 3.1 for more details.

Note that occupancy class is missing in the SRCE data, as discussed in Section 2.1. Thus, our version of the benchmark FEMA model, which we will call the **base model**, is:

$$\ln(C)_{sp} = \alpha + \zeta_{sp} + \beta_1 \ln(\text{Area}) + \beta_2 \ln(\text{Age}) + \beta_3 \ln(\text{Stories}) + \beta_4 (\text{Occupancy condition}) + \varepsilon$$
(3)

where equation (3) is estimated separately for each building group b.

Given the discussion in the previous section, we consider whether it is necessary to vary all slopes and coefficients by building group. We estimate a single model with building group fixed effects, rather than fitting separate regressions for each building group. Moreover, we include individual seismicity and performance objective fixed effects,  $\eta_s$  and  $\delta_p$ , in addition to their interaction,  $\zeta_{sp} \equiv \eta_s \times \delta_p$ . Finally, it is worth noting that the SRCE data includes information on a building's historical status. The Historic Status Indicator is

<sup>&</sup>lt;sup>14</sup>Fitting separate regressions is equivalent to fitting a single regression in which each variable is interacted with the building group, b.

used to designate whether a retrofit requires "special consideration...for preserving historic character of the building" (FEMA 157 [5], pp. 38).

Thus, the main model we propose is:

$$\ln(C)_{s,p,b} = \alpha + \eta_s + \delta_p + \zeta_{sp} + \gamma_b + \beta_1 \ln(\text{Area}) + \beta_2 \ln(\text{Age}) + \beta_3 \ln(\text{Stories}) + \beta_4(\text{Occupancy condition}) + \beta_5(\text{Historic}) + \varepsilon$$
(4)

The main motivation for including the Historic indicator is that retrofits of historic buildings are likely to be very different, and more expensive, than retrofits of non-historic indicators. Our main model thus acknowledges these potential differences.

In addition to the Historic indicator, the SRCE data includes a wealth of additional information. It is worth asking whether any other factors are important for predicting costs.

One potential extension to the model given in equation (4) is to include information on building height, measured in feet (see Table 2):

$$\ln(C)_{s,p,b} = \alpha + \eta_s + \delta_p + \zeta_{sp} + \gamma_b + \beta_1 \ln(\text{Area}) + \beta_2 \ln(\text{Age}) + \beta_3 \ln(\text{Stories}) + \beta_4 (\text{Occupancy condition}) + \beta_5 (\text{Historic}) + \beta_6 \ln(\text{Height}) + \varepsilon$$
(5)

Another possibility is to use information on the duration of retrofit construction. Thus, another alternative to (4) is:

$$\ln(C)_{s,p,b} = \alpha + \eta_s + \delta_p + \zeta_{sp} + \gamma_b + \beta_1 \ln(\text{Area}) + \beta_2 \ln(\text{Age}) + \beta_3 \ln(\text{Stories}) + \beta_4 (\text{Occupancy condition}) + \beta_5 (\text{Historic}) + \beta_6 \ln(\text{Duration}) + \varepsilon$$
(6)

Table 2 summarizes duration, given in months.

Alternatively, suppose a decision-maker has information on *past* retrofit costs. Such lagged costs could prove an important predictor for current costs. Given the SRCE data's limited sample size, we consider a very simple lag, defined as the average of retrofit costs for the previous year:

$$\log(C) \equiv \frac{\sum_{j \in N_{-1}} \ln(C_j)}{|N_{-1}|}$$

where  $N_{-1}$  is the set of all observations in the year before *C*. Then another extension is the following:

$$\ln(C)_{s,p,b} = \alpha + \eta_s + \delta_p + \zeta_{sp} + \gamma_b + \beta_1 \ln(\text{Area}) + \beta_2 \ln(\text{Age}) + \beta_3 \ln(\text{Stories}) + \beta_4(\text{Occupancy condition}) + \beta_5(\text{Historic}) + \beta_6(\log(C)) + \varepsilon$$
(7)

Finally, we consider training our model using our updated definition of seismicity. In particular, following the discussion in Section 2.1.1, we use peak ground acceleration, *pga*, with a 10 % probability of exceedance in 50 years as our measure of seicmicity.

Since the SRCE data only provides location to the county level, we use county-level *pga*, weighted by population. Thus, a third extension using *pga* is given below:

$$\ln(C)_{\underline{\lambda},p,b} = \alpha + \gamma_{\underline{\lambda}} + \delta_p + \zeta_{\underline{\lambda},p} + \gamma_b + \beta_1 \ln(\operatorname{Area}) + \beta_2 \ln(\operatorname{Age}) + \beta_3 \ln(\operatorname{Stories}) + \beta_4(\operatorname{Occupancy condition}) + \beta_5(\operatorname{Historic}) + \beta_6(\operatorname{pga}) + \varepsilon$$
(8)

A key difference between using the measure of seismicity given in the FEMA data and our county-level *pga* is that the latter is a *continuous* predictor rather than categorical. Thus, the question of interest regarding the model in equation (8) is whether a finer measure of seismicity improves prediction.

#### 2.4 Summary of the General Methodological Approach

This section provides both a summary of the methodology, as discussed above, as well as a preview of the application of this methodology in Section 3.

- We train a series of linear regression models, given by equations (3)–(8), on a subset of the SRCE data (the "training" step).
- We estimate prediction error (MSE of predictions) on the test set (the SRCE data that is not used for training) (the "testing" step).
- We compare prediction error for our proposed models to the benchmark FEMA model.

In addition, we explore how data limitations (e.g., no information on building type) affect prediction error in Section 3.4.

Note that the training and test sets do not include those observations with performance objective *RR*. In addition, since long-term USGS maps for *pga* are only available for the continental United States, we exclude the US Virgin Islands and territories (no buildings in the SRCE data are in Alaska or Hawaii).<sup>15</sup> Finally, cost predictions based on the training data will be for 1993 USD in Missouri. To update costs to 2016 USD, simply multiply the cost per square foot by  $BCI_{US,2016}$ .

<sup>&</sup>lt;sup>15</sup>In Section 3, we use the hold-out Canadian data as another test set.

#### **3.** Cost Estimation: Main Results

Before presenting the main results, we attempt to replicate the FEMA 156/157 [4, 5] results in order to have a baseline for comparison. The results are presented in Section 3.1. Note that the replication is performed *on the full synthetic data set* so that results are as close to comparable as possible.

Section 3.2 then compares prediction error from training a model on the synthetic data to prediction error obtained from training on the *raw* data. The results suggest that training on the synthetic data leads to overfitting: that is, the model only does well in terms of prediction error on the synthetic data it is trained on (i.e., good in-sample performance but poor out-of-sample performance). This is likely because the re-sampling approach used to generate the synthetic data leads to highly correlated observations.

In Section 3.3, we present the results of training the models proposed in Section 2.3. The results demonstrate that fitting separate regressions based on building group is worse, in terms of prediction error, than fitting a single regression with suitable fixed effects. We also illustrate how data limitations (e.g., no information on building type) affect prediction error in Section 3.4.

#### 3.1 FEMA 156/157 Replication Results

FEMA 157 [5] presents a detailed description of the "super database" construction. A diagram from FEMA 157 that summarizes the steps is shown in Fig. 12 in the Appendix. The main idea is to use the quality factors for each observation to construct both a sampling distribution and to determine the number of samples drawn from the distribution.<sup>16</sup> Samples are drawn so that each building group, seismicity, and performance objective are roughly equally represented in the synthetic data.

Note that only costs are sampled; building characteristics remain fixed. Thus, for example, if building i has a quality factor of 10, then 100 cost samples are drawn from the sampling distribution based on building i, with each of the 100 samples having the same building characteristics as building i.

We attempt to replicate the original FEMA 156/157 results, using the base model given in equation (3), on our synthetic data set. The original regression results from FEMA 156/157 are presented in the Appendix, Fig. 13. Replication results are presented in the Appendix, Tables 11 and 12, with standard errors in parentheses. Note that the coefficient estimates are not identical, but for the most part coincide with the original results in terms of magnitude and sign.<sup>17</sup> Moreover, one might be tempted to simply use the benchmark FEMA model after examining the high statistical significance of almost all coefficient estimates.

<sup>&</sup>lt;sup>16</sup>In particular, a log-normal sampling distribution is used, with mean equal to the observed cost and coefficient of variation decreasing with the quality factor. Number of samples increase linearly with the quality factor. See Chapter 4, FEMA 157 [5].

<sup>&</sup>lt;sup>17</sup>Recall that occupancy class is missing, in addition to the differences with respect to the location adjustment and the sample size of the raw data.

#### **3.2** Prediction Error and Overfitting on the Synthetic Data

Statistical significance, however, does not tell the whole story. If the objective is to obtain reliable predictions, a natural metric is *prediction error*. In this section, we explore the possibility that the base model is overfitting the synthetic data by comparing prediction errors on both the raw and synthetic data. In addition, we compare prediction error for the base and main models on each data set.

We randomly split our raw and synthetic data sets into two subsets:<sup>18</sup>

- a **training set** of size *n* for fitting the model (i.e., estimating the coefficients); and
- a **test set** of size *m* for estimating prediction error.

We define prediction error as the Mean Squared Error (MSE) of predictions,

$$MSE \equiv \frac{1}{n} \sum_{i=1}^{m} (\hat{y}_i - y_i)^2,$$
(9)

where  $y_i \equiv \ln(C_i), i = 1, ..., m$ , is the true (log) cost in our training set of size *m*, and  $\hat{y}_i, i = 1, ..., m$ , are the associated predictions based on a model fit to the training set.

The key to estimating prediction error is that the test set simulates "new data," as the data has not been seen by the model used for estimating the coefficients. If we expect the synthetic data to be a reliable proxy for data generated from the *true* data generating process, then prediction error should not differ substantially between the raw and synthetic data. Thus, the main question with respect to overfitting is whether prediction error behaves differently on the raw data than it does on the synthetic data.

To get a sense of the base model's *relative* performance, we compare prediction error for the base model to that for our main model, equation (4). Thus, we train the base and main models on training sets derived from either the raw or the synthetic data, and estimate prediction errors based on the respective test sets.

Table 4 shows the results of this experiment, together with the mean predicted and true (log) costs. Figures 4–6 provide more detail.

Data	Model	MSE	Predicted (log) cost	Actual (log) cost
raw	base	1.06	2.77	2.72
raw	main	1.01	2.76	2.72
synthetic	base	0.69	2.95	2.96
synthetic	main	1.08	2.96	2.96

Table 4. MSE of predictions, raw vs synthetic data.

<sup>&</sup>lt;sup>18</sup>The synthetic data, sample size  $N_s = 75\ 000$ , is split as 60 % training and 40 % testing. Given the significantly smaller sample size of the raw data,  $N_r = 1716$ , the raw split is 75 % training and 25 % testing.



**Fig. 4.** Prediction error across true values (vertical line is the true mean), for raw and synthetic data.



Fig. 5. Prediction across true values (vertical line is the true mean), for raw and synthetic data.



Fig. 6. Prediction error distributions. Midline of box is median and labeled point is the MSE.

The results reveal that the base model only outperforms the main model on the synthetic data. If we had accepted the coefficient estimates from the replication exercise, and simply compared prediction errors for the base and main models on the synthetic data, we would conclude that the FEMA model is better. For instance, in Fig. 4, we see that both models, on each data set, achieve a fairly low prediction error around the true mean. On the other hand, the base model dominates the main model in terms of prediction error on the synthetic data.

However, the sizable difference in prediction error between the two models disappears when we train our models on the raw data. As Figures 4 and 5 show, predictions further from the mean are sometimes much worse when we use the base model.

Figure 6 illustrate the prediction error distributions. The boxes represent the lower and upper quartiles, with the midline representing the median. The mean (i.e., the MSE) is also shown. The main observation is that the distributions of prediction error for both models are more disperse on the synthetic data.

Comparing prediction error on another hold-out set reveals an even starker contrast. Recall that we are training our models on data for buildings within the contiguous United States. In particular, we omit the FEMA data on Canadian buildings. Thus, evaluating predictions on this set of Canadian buildings provides a stronger test of predictive power as the data is, in a very real sense, truly different from the training data. Table 5 shows how much worse the base model performs relative to the main model *on both* the raw and synthetic datasets.<sup>19</sup>

Data	Model	MSE	Predicted (log) cost	Actual (log) cost
raw	base	2.507	2.98	2.81
raw	main	0.709	2.65	2.81
synthetic	base	2.432	3.19	2.81
synthetic	main	0.993	2.85	2.81

Table 5. MSE of predictions for Canadian data, from training on raw vs synthetic data.

Together with Table 4, these results suggest that the base model is overfitting on the synthetic data. Thus, we caution against using the synthetic data method in FEMA 156/157 [4, 5]. Moreover, the relative performance of the main model on the raw data suggests that training models separately by building group is unnecessary. Next, we explore whether our proposed extensions can reduce prediction error even further.

#### 3.3 Predicting Cost Using the Main Model and Its Extensions

We consider each of the extensions discussed in Section 2.3, equations (5)–(8), in addition to the main model given in equation (4). Each model is trained on the *raw* training set and

<sup>&</sup>lt;sup>19</sup>The hold-out Canada data constitutes 187 observations.

prediction errors are estimated on both the raw test set and the hold-out Canadian test set.<sup>20</sup> We also consider a model without the interaction between the seismicity and performance objective fixed effects,  $\zeta_{s,p} \equiv \eta_s \times \delta_p$ .

Table 6 presents prediction error for each model, relative to the base model. The test error is estimated on the raw test set. Note that each model outperforms the base model, in terms of reducing prediction error.

Model	MSE	Relative Error
duration model; equation (6)	0.925	-13.05 %
pga model; equation (8)	0.997	-6.22 %
main model; equation (4)	1.013	-4.77 %
lag model; equation (7)	1.017	-4.35 %
no interactions model	1.029	-3.2 %
height model; equation (5)	1.044	-1.81 %

Table 6. MSE relative to base model.

The main model provides a reduction in prediction error of over 4 % relative to the base model. Note that omitting the interaction between seismicity and performance,  $\zeta_{sp}$ , still reduces prediction error by 3 % relative to the base model. However, the lower MSE of the main model suggests that the interaction is important for prediction.

The height model, the worst performer, still reduces prediction error by a little less than 2 %. Nevertheless, the results suggest that adding height to the main model as in equation (5) does not improve-and may actually hurt-prediction.

While the lag reduces prediction error relative to the base model, it is outperformed by the main model-though not by much. In choosing between the main model in equation (4) and the lag model in equation (7), an important consideration is the availability of data on past retrofit costs. Given the relative data demands of the models, the simplicity of the main model makes it the more attractive alternative.

The model in equation (8) that uses pga (see Section 2.1.1), a continuous measure of seismicity, reduces prediction error by over 6 % relative to the base model, outperforming the main model. This is encouraging for a decision-maker looking to apply this methodology to his own data, since it is reasonable that the decision-maker would use current USGS seismic hazard data for assigning seismicity.

The best performing model includes information on duration of the retrofit, as given in equation (6). The improvement relative to the base model is remarkable at 13 %, far

<sup>&</sup>lt;sup>20</sup>The raw data is randomly split as: 75 % for training and 25 % for testing. Given the restrictions discussed in Section 2.4 (the exclusion of US territories and performance objective *RR*) the raw data sample size is N = 1716, and the split results in n = 1287 training samples and m = 429 test samples. Note that certain observations are missing age, occupancy condition, or historic indicator, so the actual degrees of freedom are much lower. Given the missing data (and excluding buildings with age equal to zero), the true training size is n = 812.

outperforming the main model. That the length of construction is an important predictor of cost is perhaps not too surprising in hindsight. Moreover, in applying the methodology a decision-maker does not have to collect data on duration, but may instead produce a range of estimates parameterized by duration. Thus, in terms of data demands, the duration model retains the simplicity of the main model while potentially improving prediction.

We also estimate prediction error on the hold-out Canada data, as shown in Table 7. Note that the ordering is very different. The duration model is now the worst performer, while the model without interaction term  $\zeta_{sp}$  is the best. Nevertheless, all of the models outperform the base model on the Canadian data.<sup>21</sup> The main model, with MSE almost identical to the model with no interaction term, still stands out for its predictive performance, relative simplicity, and data demands.

Model	MSE	Relative Error
no interactions model	0.701	-72.04 %
main model	0.709	-71.74 %
lag model	0.729	-70.93 %
height model	0.831	-66.84 %
duration model	1.572	-37.3 %

 Table 7. MSE relative to base model on Canada data.

#### 3.4 How Data Limitations Affect Prediction Error

It may not always be possible to collect all of the relevant building characteristics to train the models from Section 2.3. A decision-maker may not have the time or resources to collect the relevant data, while a researcher may not have access to all of the data. Information on building type, for instance, may be particularly difficult to obtain.

In this section, we explore how such data limitations affect prediction error. For the sake of illustration, we consider the following scenarios:

- We do not have information on building type (or group);
- We do not have information on building age;
- We do not have information on number of stories.

In addition, we consider combinations of these scenarios.<sup>22</sup>

<sup>&</sup>lt;sup>21</sup>Note that the *pga* model is not tested on the Canada data, since we use *pga* for the United States.

<sup>&</sup>lt;sup>22</sup>Rather than exploring all possible data limitations, we assume that building area and historic status are basic building characteristics that should be easy to obtain. A similar exercise may be performed on these characteristics.

We train the main model, equation (4), subject to these data limitations. For example, with no building age, the model is:

$$\ln(C)_{s,p,g} = \alpha + \eta_s + \delta_p + \zeta_{sp} + \gamma_g + \beta_1 \ln(\text{Area}) + \beta_2 \ln(\text{Age}) + \beta_3 \ln(\text{Stories}) + \beta_4(\text{Occupancy condition}) + \beta_5(\text{Historic}) + \varepsilon$$
(10)

Table 8 presents the prediction error of these models, relative to the main model with no data limitations.

Model	MSE	Relative Error
No age	1.00	-1.2 %
No age, stories	1.02	0.36 %
No stories	1.02	0.42~%
No building group	1.07	5.46 %
No age, stories, building group	1.10	8.8 %

Table 8. MSE subject to data limitations, relative to main model.

For the most part, training the main model, equation (4), subject to data limitations leads to an increase in prediction error. This is perhaps not too surprising, and suggests that the main model is the best alternative to the FEMA model.

The model with no stories yields a higher prediction error than the main model without data limitations, though not by much. Interestingly, the main model with no age information performs *better* than the main model, though only by about 1 %. Thus, if we have no information on age or stories, the increase in prediction error is almost negligible.

As expected, our predictive performance is penalized without information on building type. The model without building group increases prediction error by about 5.5 %, a significant loss in predictive ability. Without information on age, stories, or building group, we see an almost 9 % increase in prediction error.

Figures 7 and 8 illustrate the relative trade-off in prediction error from excluding age and stories, which are only marginally important, and excluding building group, which appears to be more important. In particular, note the discrepancy from the main model in Fig. 8 away from the mean. On the other hand, prediction errors are fairly close across the range of true values when we only exclude age and stories, as seen in Fig. 7.

The preceding results, together with the results in Section 3.3, suggest that the main model given by equation (4) provides the most reasonable option for estimating retrofit costs.





Fig. 7. Prediction errorm with and without age, stories (vertical line is the true mean).





Fig. 8. Prediction error with and without building group (vertical line is the true mean).

#### 4. Conclusion

We present a methodology for estimating retrofit costs that builds on the methodology developed in FEMA 156 [4] and 157 [5]. Our approach uses the data collected for the FEMA reports to train a series of regression models, varying in their data requirements. In a forthcoming report, we apply this methodology to obtain retrofit cost estimates for federal buildings.

Estimates of prediction error allow comparison across models. The results suggest that a simple modification to the FEMA base model—equation (3)—can improve prediction in many applications. In particular, our main model—equation (4)—is a *single* regression with building group fixed effects, in contrast to fitting the base model separately for each building group, that also includes an indicator for whether a building is deemed historical. Our results suggest that this model yields a lower prediction error than the FEMA base model.

Prediction error estimates also allow us to quantify the penalty for data limitations. For instance, information on building type seems to be important for prediction, while information on building age does not. Thus, information on building type appears to be crucial for accurate cost predictions.

One result of particular interest for researchers or decision-makers interested in using the FEMA data is that models should be trained on the *raw* data itself and not on a synthetic data set generated from the raw data using the approach in FEMA 156/157 (see Section 2.1.2 for details). Our results suggest that using the synthetic data may lead to overfitting: models trained on the synthetic data may produce good in-sample predictions but poor out-of-sample predictions, and so cannot be be expected to perform well on new data. The FEMA base model is a striking example of this pitfall, as shown in Section 3.2.

#### 4.1 Limitations

Our results should be taken with caution. The major caveat is that the FEMA data used for training our models is outdated. Adjusting cost estimates using the BCI, as suggested in Section 2.1.1, is a suitable, though imperfect fix. Access to more recent data on seismic retrofit costs should improve the models' predictions. However, given the relative scarcity of data—especially for the United States—training models on the FEMA data provides a sensible starting point.

This limitation highlights the importance of sharing cost data. Our results suggest that retrofit cost information that includes a building's area, historical status, seismicity, and type, in addition to the retrofit performance objective and occupancy condition, would be sufficient for updating the retrofit cost database. Data shared in this way should not compromise sensitive or private information.

#### 4.2 Guidance and Future Directions

We focus on structural construction costs. The FEMA data also includes non-construction costs (e.g., permits, fees, and relocation costs). These could easily be added to structural retrofit estimates *ex post*. However, this approach does not help with predictions of *total cost* for new data.

An alternative approach is to train a (potentially new) set of models to predict total cost costs. The models may differ from ours, as the predictors of total cost may differ from those for structural costs. In particular, building characteristics may be poor predictors of non-construction costs. We explore models for total costs in future work.

More importantly, our focus on construction costs ignores *indirect costs*, such as loss of productivity during a retrofit. Construction projects may also impose externalities on neighbors, in the form of noise or disruption to traffic. Nevertheless, obtaining reasonable construction cost estimates is an important first step. In future work, we attempt to quantify direct and indirect costs of seismic retrofits.

Finally, cost estimates provide important information to decision-makers, but may not help in *making decisions*. Given a range of estimates, how does a decision-maker prioritize buildings in his inventory for retrofit? Should the decision-maker only retrofit the most vulnerable buildings? How does a decision-maker choose between *LS* and *IO*? We examine decision-making protocols, using our estimates as input data, in future work.

#### References

- 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).
- [2] 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.
- [3] FEMA (2006) Techniques for the seismic rehabilitation of existing buildings. Federal Emergency Management Agency, Technical Report FEMA 547.
- [4] FEMA (1994) Typical costs for seismic rehabilitation of existing buildings, volume 1: Summary. Federal Emergency Management Agency, Technical Report FEMA 156.
- [5] FEMA (1995) Typical costs for seismic rehabilitation of existing buildings, volume2: Supporting documentation. Federal Emergency Management Agency, Technical Report FEMA 157.
- [6] Lew HS (1999) Supporting document for rehabilitation cost estimates of fema existing buildings. *NIST Interagency/Internal Report (NISTIR)-6315*.
- [7] Jafarzadeh R, Wilkinson S, Gonzalez V, Ingham J, Amiri GG (2013) Predicting seismic retrofit construction cost for buildings with framed structures using multilinear regression analysis. *Journal of Construction Engineering and Management* 140(3):04013062.
- [8] Jafarzadeh R, Ingham J, Wilkinson S, González V, Aghakouchak A (2013) Application of artificial neural network methodology for predicting seismic retrofit construction costs. *Journal of Construction Engineering and Management* 140(2):04013044.
- [9] Jafarzadeh R, Ingham JM, Wilkinson S (2014) A seismic retrofit cost database for buildings with a framed structure. *Earthquake Spectra* 30(2):625–637.
- [10] Nasrazadani H, Mahsuli M, Talebiyan H, Kashani H (2017) Probabilistic modeling framework for prediction of seismic retrofit cost of buildings. *Journal of Construction Engineering and Management* 143(8):04017055.
- [11] Chen WT, Huang YH (2006) Approximately predicting the cost and duration of school reconstruction projects in taiwan. *Construction management and Economics* 24(12):1231–1239.
- [12] Hopkins DC, Stuart G (2003) Strengthening existing new zealand buildings for earthquake: An analysis of cost benefit using annual probabilities. 2003 Pacific Conference on Earthquake Engineering, Christchurch, New Zealand, pp 13–15.
- [13] FEMA (1991) Nehrp recommended provisions for the development of regulations of new buildings. Federal Emergency Management Agency, Technical Report FEMA 222-223.

#### **Appendix A: Supplemental Materials**

This Appendix includes the following supplemental materials:

- Supplemental information from FEMA 156/157 and additional descriptive statistics for SRCE data.
- Replication results and coefficient estimates for trained models from Section 2.3.

#### Supplemental Information: FEMA 156/157 and SRCE data

Figure 9 presents the original definition for the seismicity-performance objective interaction term (S-P) used in FEMA 156/157. We only use these terms in the replication. See Section 3.1.

The ATC-3 seismicity map, based on Effective Peak Acceleration (EPA), used in FEMA 156/157, is shown in Fig. 10. Note that the EPA coefficient, *Aa*, is binned into seven categories, corresponding to the seven seismicity "zones" in the SRCE data. See the second footnote in Section 2.3 for a description of the seismicity categories and zones.

Figure 11, from FEMA 156, provides a description of the building types represented in the FEMA 156/157 data. The SRCE data reflects these building types. Note that FEMA 178, the source for the building type definitions, has since been updated and now includes 16 rather than 15 types; see FEMA 547 [3].

Table 9 provides a summary of the building groups and models represented in the SRCE data. Note the prevalence of Unreinforced Masonry (URM) buildings in the data.

Table 10 provides additional summary statistics for the SRCE data. Specifically, the table summarizes the predictors used in the extensions of the main model discussed in Section 2.3: construction duration (in months), building height (in feet and meters), and pga as defined in Section 2.1.1.

Finally, Fig. 12, from FEMA 157, illustrates the step-by-step process for creating the synthetic data set (the "Super Database") from the raw data. Section 2.1.2 provides more details.

#### **Coefficient Estimates: Replication and Training**

Figure 13 presents the original coefficient estimates from FEMA 156/157. The variables  $C_c, X_1, \ldots, X_6$  are the constant, area, age, stories, seismicity-performance (S-P) interaction, occupancy class, and occupancy condition, respectively. Replication results are shown in Tables 11 and 12. See Section 3.1 for discussion.

For sake of completeness, we also present coefficient estimates for the models presented in Section 2.3 and trained in Section 3.3. The models are trained on the raw test set. Standard errors are given in parentheses.

SEISMICITY	PERFORMANCE OBJECTIVES				
	LIFE SAFETY	DAMAGE CONTROL	IMMEDIATE OCCUPANCY		
Low	1	5	9		
Moderate	2	6	10		
High	3	7	11		
Very High	4	8	12		

#### TABLE 7.3.2 CATEGORY FOR CONSTANT X4

Fig. 9. Seismicity-performance objective interaction term in FEMA 156/157 (Source: FEMA 156).



Fig. 10. ATC-3 seismicity map (Source: FEMA 156).

BUILDING GROUP	MODEL	FEMA 178 BUILDING TYPES		
1	URM	Unreinforced Masonry		
2	W1 W2	Wood Light Frame Wood (Commercial or Industrial)		
3	PC1 Precast Concrete Tilt Up Walls RM1 Reinforced Masonry with Metal or W 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	Concrete Shear Wall Precast Concrete Frame with Concrete Walls		
и 1 – 12	RM2 S4	Reinforced Masonry with Precast Concrete Diaphragm Steel Frame with Concrete Walls		

Fig. 11. Building groups and associated building types (Source: FEMA 156).

Building Group	Model	Count	Percentage
1	URM	572	33.33 %
_	W1	52	3.03 %
2	W2	52	3.03 %
_	PC1	55	3.21 %
3	RM1	57	3.32 %
	C1	108	6.29 %
4	C3	264	15.38 %
5	<b>S</b> 1	78	4.55 %
	S2	32	1.86 %
6	<b>S</b> 3	16	0.93 %
7	S5	110	6.41 %
	C2	258	15.03 %
	PC2	14	0.82 %
8	RM2	11	0.64 %
	S4	37	2.16 %

Table 9. Building groups and types in SRCE data.

**Table 10.** Summary statistics for SRCE data: building characteristics used in extended models (N = 1083).

Stat	Duration: months	Height: ft (m)	pga: %g
Min	1	8 (2.4)	1.0
Mean	9	45.2 (13.8)	27.9
Median	6	30 (9.1)	36.6
Max	60	444 (135.3)	58.5
Std dev	8	41.5 (12.6)	15.1



### Figure 4.4.1 Extended Data Set Creation Process

Fig. 12. Super Database Algorithm (Source: FEMA 157).

COEFF.	CATE- GORY	BUILDING GROUP							
		1	2	3	4	5	6	7	8
Cc	-	151.9	1.2	13.5	36.9	182.5	137.6	59.2	86.5
X1	-	-0.23	-0.02	-0.26	-0.15	-0.30	-0.11	-0.26	-0.28
X2	-	0.02	0.52	0.60	0.18	0.19	-0.50	0.40	0.14
ХЗ	-	0.28	-0.28	1.06	0.43	0.21	-0.71	0.40	0.53
X4 (See Table 4.5.2 below)	1	0.28	0.48	0.51	0.48	0.53	0.58.	0.47	0.61
	2	2.65	0.61	0.41	2.55	0.46	0.73	1.20	0.64
	3	1.16	0.72	1.25	0.72	1.07	1.27	0.97	0.43
	4	0.57	1.31	0.70	1.03	1.22	0.90	1.74	1.02
	5	0.69	0.40	0.35	0.52	0.76	0.83	0.67	0.44
9.5	6	0.57	0.67	1.03	0.52	0.14	0.30	0.32	2.27
	7	0.76	1.17	0.96	1.01	1.23	0.42	0.81	1.42
	8	2.30	2.53	1.01	1.02	1.30	0.43	1.40	1.61
ĺ	9	1.48	1.12	1.20	1.17	1.25	1.35	1.10	1.86
190 - C	10	1.28	1.31	1.16	0.62	2.71	3.21	1.25	1.38
	11	1.60	1.24	3.23	1.28	1.89	2.12	1.57	0.46
	12	2.09	1.10	2.15	2.10	1.44	2.36	1.54	1.89
X5	Р•	4.27	1.09	1.09	0.26	1.19	1.48	1.15	0.45
[	м	0.76	0.43	0.59	4.50	0.45	0.56	0.85	0.36
	R	0.48	0.90	2.19	0.75	2.72	1.11	0.32	1.09
[	F	0.98	0.91	0.99	1.03	0.39	0.54	0.96	2.21
	1	0.97	1.35	1.00	0.82	1.29	0.47	1.17	0.96
	С	0.82	0.94	1.47	1.01	0.81	0.73	2.48	1.25
	А	0.83	2.22	.53	1.33	0.91	4.77	1.33	2.16
X6	IP**	0.69	1.78	1.00	0.77	1.11	0.63	0.93	0.69
Γ	TR	1,12	1.13	0.96	1.44	1.28	1.94	1.08	1.21
Γ	v	1.30	0.50	1.04	0.90	0.70	0.81	0.99	1.20

### TABLE 7.3.1 VALUES OF REGRESSION VARIABLES

Fig. 13. Original regression results (Source: FEMA 157).

	Dependent variable:						
		Cost	per sf				
	BG 1	BG 2	BG 3	BG 4			
Area	-0.137*** (0.010)	0.354*** (0.014)	-0.116*** (0.009)	-0.126*** (0.011)			
Age	-20.500*** (0.856)	-38.200*** (0.871)	$-64.000^{***}$ (2.540)	-16.000*** (1.190)			
Stories	-0.026(0.022)	-0.057* (0.031)	0.904*** (0.027)	0.105*** (0.018)			
Occup: Temp removed	-0.862*** (0.024)	-0.015 (0.027)	$-0.114^{***}$ (0.027)	-1.000*** (0.026)			
Occup: In-place	-1.350*** (0.029)	0.597*** (0.040)	$-0.208^{***}$ (0.032)	-1.420*** (0.031)			
S-P 2	-0.074* (0.042)						
S-P 3	0.451*** (0.043)		0.778*** (0.038)	-0.179*** (0.043)			
S-P 4	-0.743*** (0.043)	$-0.822^{***}$ (0.041)	0.922*** (0.035)	0.206*** (0.044)			
S-P 5	-0.201*** (0.048)	$-1.540^{***}$ (0.048)	1.280*** (0.050)	-0.044(0.046)			
S-P 6	-0.200*** (0.044)	-0.109*** (0.041)		-0.284*** (0.048)			
S-P 7	$-0.400^{***}$ (0.042)	1.900*** (0.047)	0.893*** (0.040)	0.895*** (0.048)			
S-P 8	0.813*** (0.041)	0.311*** (0.041)	1.050*** (0.036)	0.238*** (0.047)			
S-P 9	0.692*** (0.046)			1.170*** (0.050)			
S-P 10	0.254*** (0.042)		1.720*** (0.047)	-0.046(0.049)			
S-P 11	0.986*** (0.046)	$-0.946^{***}$ (0.042)		0.649*** (0.047)			
S-P 12	1.780*** (0.042)	1.610*** (0.045)	1.710*** (0.039)	0.956*** (0.046)			
Constant	160.000*** (6.480)	288.000*** (6.570)	488.000*** (19.300)	126.000*** (8.990)			
Observations	12 000	8 000	8 000	11 000			
$\mathbb{R}^2$	0.497	0.561	0.563	0.340			
Adjusted R <sup>2</sup>	0.496	0.560	0.562	0.339			
Notes:	***Significant at the	e 1 percent level.					
	**Significant at the 5 percent level.						

 Table 11. Replication results: building groups 1 - 4

\*Significant at the 10 percent level.

	Dependent variable:						
		Cost	per sf				
	BG 5	BG 6	BG 7	BG 8			
Area	-0.651*** (0.009)	$-0.872^{***}$ (0.015)	-0.057*** (0.008)	0.063*** (0.015)			
Age	-0.091 (1.300)	41.500*** (1.840)	$-27.900^{***}$ (0.849)	$-48.400^{***}$ (1.580)			
Stories	0.342*** (0.013)	-0.571*** (0.026)	0.214*** (0.015)	0.035 (0.023)			
Occup: Temp removed	0.304*** (0.032)	0.223*** (0.076)	0.172*** (0.017)	$-0.352^{***}$ (0.033)			
Occup: In-place	0.079** (0.038)	0.027 (0.065)	$-0.088^{**}$ (0.036)	$-1.240^{***}$ (0.038)			
S-P 2			-0.123*** (0.029)				
S-P 3	-0.269*** (0.042)		$-0.682^{***}$ (0.029)	$-0.768^{***}$ (0.046)			
S-P 4	-0.003(0.038)	-0.715*** (0.033)	$-0.360^{***}$ (0.029)	$-0.330^{***}$ (0.047)			
S-P 5			1.630*** (0.048)	$-0.591^{***}$ (0.047)			
S-P 6	2.510*** (0.040)	$-0.694^{***}$ (0.045)	-1.240*** (0.036)	0.601*** (0.052)			
S-P 7	-0.467*** (0.042)		-0.355*** (0.030)	0.560*** (0.053)			
S-P 8	-0.230*** (0.036)	0.240*** (0.053)	1.270*** (0.032)	-0.050(0.048)			
S-P 9			0.323*** (0.027)	0.450*** (0.052)			
S-P 10	2.000*** (0.040)	2.920*** (0.100)	$-0.285^{***}$ (0.028)				
S-P 11	0.727*** (0.047)			-3.350*** (0.050)			
S-P 12	-0.022(0.040)	$-1.460^{***}$ (0.044)	0.336*** (0.029)	0.148*** (0.047)			
Constant	10.200 (9.840)	-304.000*** (14.000)	215.000*** (6.420)	369.000*** (11.900)			
Observations	9 000	6 000	11 000	10 000			
$\mathbb{R}^2$	0.625	0.596	0.689	0.515			
Adjusted R <sup>2</sup>	0.624	0.595	0.688	0.514			
Notes:	***Significant at th	e 1 percent level.					

 Table 12. Replication results: building groups 5 - 8

\*\*Significant at the 5 percent level. \*Significant at the 10 percent level.

			Depender	ıt variable:			
	Cost per sf						
	no interactions	main	height	pga	duration	lag	
Area	-0.170*** (0.038)	-0.164*** (0.038)	-0.264*** (0.050)	-0.164*** (0.038)	-0.441*** (0.076)	-0.164*** (0.038)	
Age	-3.500 (4.590)	-6.080(4.680)	-4.990 (5.790)	-2.690 (4.510)	-10.400(8.820)	-6.120 (4.680)	
Stories	0.303*** (0.068)	0.306*** (0.068)	0.041 (0.124)	0.307*** (0.067)	0.122 (0.122)	0.303*** (0.068)	
Occupancy: TR	-0.326*** (0.095)	-0.308*** (0.096)	-0.267** (0.110)	-0.318*** (0.094)	-0.268* (0.151)	-0.304*** (0.096)	
Occupancy: IP	-0.986*** (0.117)	-0.947*** (0.119)	-0.995*** (0.136)	-0.989*** (0.116)	-0.594*** (0.173)	-0.948*** (0.119)	
Seismicity: M	-0.278 (0.177)	-0.064 (0.404)	-0.043 (0.425)		-0.488 (0.681)	-0.054 (0.405)	
Seismicity: H	0.002 (0.193)	-0.038 (0.337)	0.141 (0.355)		-0.064 (0.493)	-0.049(0.338)	
Seismicity: VH	0.055 (0.165)	-0.132(0.322)	-0.260(0.327)		-0.390 (0.399)	-0.132(0.322)	
Performance: DC	0.267*** (0.103)	0.040 (0.371)	-0.182 (0.423)	0.272*** (0.097)	-0.019 (0.517)	0.070 (0.373)	
Performance: IO	0.685*** (0.112)	0.635 (0.479)	0.683 (0.525)	0.653*** (0.107)	0.462 (0.624)	0.657 (0.480)	
BG: 2	-0.110 (0.174)	-0.129 (0.174)	-0.310 (0.228)	-0.155 (0.174)	0.447 (0.345)	-0.123 (0.174)	
BG: 3	-0.123 (0.190)	-0.139 (0.191)	-0.074(0.221)	-0.150(0.189)	1.200*** (0.346)	-0.134 (0.191)	
BG: 4	0.306** (0.137)	0.326** (0.139)	0.487*** (0.178)	0.278** (0.136)	1.220*** (0.267)	0.324** (0.140)	
BG: 5	0.389* (0.205)	0.330 (0.208)	0.336 (0.237)	0.322 (0.205)	1.410*** (0.450)	0.335 (0.208)	
BG: 6	-0.661*** (0.254)	-0.668*** (0.254)	-0.684** (0.288)	-0.730*** (0.254)	0.826 (0.624)	-0.664*** (0.254)	
BG: 7	0.534*** (0.175)	0.505*** (0.177)	0.547** (0.233)	0.525*** (0.174)	0.762* (0.456)	0.498*** (0.177)	
BG: 8	0.170 (0.140)	0.145 (0.141)	0.313* (0.168)	0.139 (0.140)	1.150*** (0.256)	0.137 (0.142)	
Historic	1.040*** (0.150)	0.975*** (0.153)	0.922*** (0.180)	1.040*** (0.149)	0.934*** (0.243)	0.981*** (0.153)	
Height	× ,		0.475*** (0.156)				
Duration					0.931*** (0.108)		
Lag						0.074 (0.101)	
M x DC		-0.095 (0.464)	0.407 (0.580)		0.872 (0.885)	-0.121 (0.466)	
H x DC		-0.127 (0.484)	0.246 (0.643)		-0.741 (1.130)	-0.135 (0.485)	
VH x DC		0.325 (0.395)	0.630 (0.446)		0.331 (0.543)	0.291 (0.398)	
M x IO		-0.512 (0.560)	-0.604 (0.636)		-0.365 (0.926)	-0.534 (0.561)	
V x IO		-0.336 (0.581)	0.386 (0.766)		-0.515 (0.875)	-0.317 (0.582)	
VH x IO		0.271 (0.500)	0.253 (0.550)		0.099 (0.687)	0.243 (0.501)	
pga				0.009*** (0.003)		· · · ·	
Constant	30.800 (34.700)	50.400 (35.400)	41.700 (43.800)	24.300 (34.100)	83.800 (66.600)	50.400 (35.400)	
Observations	812	812	579	812	257	812	
$\mathbb{R}^2$	0.313	0.321	0.384	0.316	0.572	0.321	
Adjusted R <sup>2</sup>	0.298	0.300	0.356	0.302	0.525	0.300	
Notes:	***Significant at th	e 1 percent level.					

Table 13. Coefficient est	timates from	training r	nain mode	el and	variations
	innaces moni	ti anning i	mann moa	or and	, and another

\*\*\*Significant at the 1 percent level. \*\*Significant at the 5 percent level. \*Significant at the 10 percent level.