Identifying Residential Fires Involving Upholstered Furniture within the National Fire Incident Reporting System

Douglas S. Thomas
David T. Butry

NIST

National Institute of Standards and Technology

U.S. Department of Commerce

Identifying Residential Fires Involving Upholstered Furniture within the National Fire Incident Reporting System

Douglas S. Thomas
David T. Butry
Applied Economics Office
Engineering Laboratory

Abstract

This Technical Note develops a method for identifying residential fire incidents, reported to the National Fire Incident Reporting System (NFIRS), with missing information regarding the "item first ignited" (IFI) or "item most contributing to flame spread" (IMCFS) that were likely to have involved upholstered furniture. Using NFIRS data, upholstered furniture involvement can only be identified using the IFI or IMCFS fields; however, it is not uncommon for these fields to contain missing values or to be coded as unknown (undetermined). Current statistical methods address this issue by allocating the missing and unknown incidents proportionally over the distribution of those known. The appropriateness of this depends on the representativeness of the known incidents. This Technical Note explores a data imputation method to determine the likelihood of upholstered furniture involvement based on statistical correlations with other factors describing the fire incident. A number of competing statistical models are evaluated using the data imputation approach, with the best performing models selected based on the minimum out-of-sample root mean squared prediction error. The optimum models based on this criterion are then used to estimate the number of U.S. residential upholstered furniture fires, related occupant fatalities, injuries, and property and content losses, and the results are compared to the conventional scaling-based techniques currently used by the National Fire Protection Association.

Keywords: fire; fire statistics; NFIRS; upholstered furniture; residential upholstered furniture; economics; imputation

Preface

This study was conducted by the Applied Economics Office in the Engineering Laboratory at the National Institute of Standards and Technology. The study develops an imputation approach to determine the expected values of missing or unknown data elements, found within the National Fire Incident Reporting System, based on observed patterns found within the non-missing data fields.

Acknowledgements

The authors wish to thank all those who contributed so many excellent ideas and suggestions for this report. Special appreciation is extended to Dr. William Pitts, Dr. Richard Gann, and Dr. Anthony Hamins of the Engineering Laboratory's Fire Research Division; and William Guthrie of the Information Technology Laboratory's Statistical Engineering Division. Special appreciation is also extended to Dr. Stanley Gilbert and Dr. Robert Chapman of the Engineering Laboratory's Applied Economics Office for their thorough reviews and many insights and to Shannon Takach for her assistance in preparing the manuscript for review and publication. The authors also wish to thank Dr. Nicos Martys, of the Materials and Structural Systems Division, for his review. Finally, we to acknowledge the thoughtful comments and suggestions provided by: Christina Holcroft, Kathleen Almand, and members of the National Fire Protection Association; Brad Peabody of the U.S. Fire Administration; and Rohit Khanna, Stephen Hanway, David Miller, Andrew Stadnik, and Kathleen Stralka from the U.S. Consumer Product Safety Commission.

Table of Contents

List of Tables

List of Figures

Executive Summary

In 2011, there were an estimated 386 000 residential structure fires, representing nearly 28 % of all fires in the United States. These fires resulted in approximately 2550 deaths, 14 360 injuries, and 7.1 billion dollars in losses (Karter 2013). Many of these fires involved upholstered furniture, which is ignitable and can contribute significantly to the growth/spread of a fire. As a result, the National Fire Protection Association (NFPA) has been summarizing the consequences of fires in which upholstered furniture was the first item ignited. This report develops performance metrics, enabling comparisons between the use of new fire mitigation technologies and their impact on the U.S. fire losses, with a particular focus on residential fires involving upholstered furniture.

The U.S. Fire Administration's National Fire Incident Reporting System (NFIRS), the most comprehensive accounting of individual fire incidents in the world, provides detailed information on more than one million fire incidents, requiring a fire department response, in the U.S. each year. Yet, it represents only a partial census, as many fire incidents, are not reported. In addition, a large number of the reported fire incidents contain incomplete descriptions. The current ('scaling') method used to produce national fire statistics assumes the fire incidents with complete data are representative of those with incomplete data—i.e., the occurrence of missing or unknown data is a completely random process.

This Technical Note offers an alternative ("data imputation") method for handling the unknown data. It exploits correlations between the fire incident's characteristic of interest (e.g., item first ignited) and other reported fire incident factors (e.g., area of origin) found in the data, to impute values for those incidents with missing or unknown data. Specifically, this Technical Note develops a method for identifying residential fire incidents, reported to the NFIRS, with missing or reported as unknown information regarding the "Item First Ignited" (IFI) or "Item Contributing Most to Flame Spread" (ICMFS) that were likely to have involved upholstered furniture. A number of competing statistical models are evaluated using the data imputation approach, with the best performing models selected based on the minimum out-of-sample root mean squared

prediction error. The optimum models based on this criterion are then used to develop alternate estimates of the number of incidents, fatalities, injuries, and direct property losses from residential fires involving upholstered furniture from 2006 to 2010.

Methodologically, the scaling and data imputation approaches will yield similar results should the fire incidents with known data be representative of the data with unknown and missing information. However, if there are systematic differences between the known and unknown data (e.g., if upholstered furniture involvement occurs more frequently in the known data), and if these differences are correlated with other, more frequently reported data, the two approaches will yield different results. While the data imputation approach offers the potential for greater accuracy, it is computationally expensive in comparison to the scaling approach.

The imputation method estimates 10 100 home fires involving upholstered furniture resulted in 440 occupant fatalities, 1100 occupant injuries, and $510 million, on average each year, from 2006 to 2010. (Annual estimates, with uncertainty ranges, are shown in Section 4.) By comparison, the scaling method estimates 8900 home fires involving upholstered furniture resulted in 610 occupant fatalities, 1120 occupant injuries, and $565 million, on average each year, from 2006 to 2010.

In general, the scaling and imputation methods produce similar statistical results for the number of residential home fires involving upholstered furniture, as well as for the resulting number of occupant injuries and direct property losses. However, the disagreement between estimated numbers of occupant fatalities using these two methods is statistically significant. The imputation method estimates about a third fewer occupant fatalities, the difference being in the number of fire fatalities resulting from upholstered furniture involvement due to a direct ignition (i.e., not from flame spread).

1 Introduction

In 2011, there were an estimated 386 000 residential structure fires, representing nearly 28 % of all fires in the United States. These fires resulted in approximately 2550 deaths, 14 360 injuries, and 7.1 billion dollars in losses (Karter 2013). Many of these fires involved upholstered furniture, which is ignitable and can contribute significantly to the growth/spread of a fire. In the past, upholstered furniture has contributed significantly to residential fires. In recent years, however, there has been efforts to reduce the fire risk of these products, and the National Fire Protection Association (NFPA) has been summarizing the consequences of fires in which upholstered furniture was the first item ignited (Ahrens, 2011). Unfortunately, it is difficult to understand the fire risk of upholstered furniture or the effect that reduced flammability might have on these fires, as a complete database of residential fire incidents does not include information on details of the type, size, or material composition of fires involving furniture.

The U.S. Fire Administration's National Fire Incident Reporting System (NFIRS) provides detailed information on more than one million fire incidents each year (on average) from 2002 to 2010. While NFIRS is the most comprehensive accounting of individual fire incidents in the world, it represents only a partial census. Many incidents are not reported. Out of an estimated 30 125 fire departments, approximately 23 000 departments report in NFIRS each year; however, some have occasional participation and many do not report at all, as it is a voluntary system. In addition, many of the reported fire incidents contain incomplete descriptions.

The conventional national fire statistics are based on NFIRS incident data—with complete/known data—that are scaled using two ratios, based on the work of Hall and Harwood (1989), although the approach has been modified over time to account for versioning changes in NFIRS (e.g., see Ahrens [2011]). (The conceptual approach has remained consistent.) The first ratio scales the NFIRS-based fire statistics to account for the sizeable number of reported fire incidents with unknown or missing data ("allocating the unknowns"). A proportional scaling is used to ensure that the total number of fire

1

incidents (or fatalities, injuries, economic loss) with known values equals the total number of fire incidents (or fatalities, injuries, economic loss) reported to NFIRS. The second ratio ("national scaling ratio") scales the NFIRS-based fire statistics into national estimates using fire department survey data from the National Fire Protection Association (NFPA).

Survey data is often weighted using an adjustment technique such as iterative proportional fitting (i.e., raking), to address statistical issues resulting from missing (incomplete) data (Izrael et al. 2004). Raking has been applied to NFIRS data (Greene 2001), and offers a number of advantages over the scaling method. Raking allocates the missing and unknown observations over multi-dimensional cross tabulation data (e.g., item first ignited *by* cause of ignition) using an iterative procedure. This procedure eliminates results that are sensitive to the adjustment (allocation) order, a known issue with the non-iterative, scaling method (Greene 2001). In addition, raking preserves the conditional probability distribution of the cross tabular data, meaning correlations between the data dimensions are the same for the input and imputed datasets. However, this technique poses some statistical challenges, as discussed in Greene (2001), when there is high dimensionality and sparse counts. Further, the iterative process requires the data to be arranged in a cross tabular structure, making it inappropriate to allocate unknowns over a single data dimension (e.g., item first ignited). Finally, while raking preserves the conditional probability distribution of the data dimensions included, it is susceptible to sample selection bias should correlated data dimensions be omitted from the procedure.

In this report, we explore the use of a new data imputation technique as an alternative method to "allocate the unknowns." To accomplish this, we evaluate the likelihood of involvement of residential upholstered furniture (RUF) in residential home fires. Using NFIRS data, upholstered furniture involvement can only be identified using the "Item First Ignited" (IFI) or "Item Contributing Most to Flame Spread" (ICMFS) fields; however, it is not uncommon for these fields to contain missing values or to be coded as

2

unknown. Table 1.1 shows the number of non-confined[1], residential fire incidents with IFI and ICMFS data listed as upholstered furniture, missing or unknown, or known, but not upholstered furniture ("other"). Both IFI and ICMFS data fields contain a significant number of incidents with missing or unknown (undetermined) values. Over the five-year period, 3363 fires or 673 on average per year had both IFI and ICMFS indicating upholstered furniture was involved in the fire.

Table 1.1 Number of non-confined, residential fire incidents reported to NFIRS identified by Item First Ignited and Item Contributing Most to Flame Spread (U.S. Fire Administration, National Fire Data Center, 2010).

Year	Item First Ignited			Item Contributing Most to Flame Spread*		
	Upholstered Furniture	*Missing / Unknown*	*Other*	*Upholstered Furniture*	*Missing / Unknown*	*Other*
2006	3105	51441	83523	659	74 096	16 888
2007	3053	58114	89540	641	80 903	17 441
2008	2937	61908	88941	641	85 160	16 683
2009	2566	62371	81261	579	79 995	16 508
2010	2714	71131	91011	608	92 699	18 085

* Excluded are incidents (A) with the Item Contributing Most to Flame Spread the same as the Item First Ignited, and (B) without fire spread.

Given the relative number of incidents with missing or unknown values and their potential influence on national RUF estimates, we investigate the usefulness of allocating the unknowns based on statistical correlations with other, non-missing, information describing the fire incident. Unlike raking, the allocation approach can be performed on a single data array, and limits selection bias through the inclusion of all other, relevant, non-missing information. A number of competing statistical models are evaluated, with the best performing models selected based on the minimum out-of-sample mean squared prediction error (Makridakis, 1998). The optimum models based on this criterion are then used to estimate the number of residential upholstered furniture fires, related occupant fatalities and injuries, and property and content losses, by year, from 2006 to 2010. In

---

[1] Confined fires include fires confined to a cooking container, chimney or flue, incinerator, burner, compactor, or trash area.

addition, nonparametric bootstrapping techniques (e.g., see Kennedy 2003) can be used to produce uncertainty ranges associated with the estimates.

This analysis is a part of a larger research effort to develop statistics and uncertainty measures for characterizing, tracking, and better understanding the root causes of the total burden of fire in the United States. The purpose of these statistics will be to provide details on the circumstances, causes, and development of fires and the related deaths, injuries, and property damage by major fire incident category (residential structure, non-residential structure, natural vegetation, vehicle), as well as to provide details on the costs related to fire protection and loss mitigation (e.g., fire protection of constructed facilities; standards and codes development, testing, and implementation; wildland fuel treatments, etc.). This information will be used to develop performance metrics, enabling comparisons between the use of new fire mitigation technologies and their impact on the U.S. fire losses, with a particular focus on residential fires involving upholstered furniture.

2 Data and Methods

2.1 Data

This analysis utilizes data from the National Fire Incident Reporting System (NFIRS), which is a product of the National Fire Data Center, an entity of the U.S. Fire Administration (USFA). The USFA is authorized to gather, analyze, and standardize fire information and data through the Federal Fire Prevention and Control Act of 1974 (FEMA 2010). The reporting format in NFIRS is, generally, consistent with the National Fire Protection Association Standard 901, "Uniform Coding for Fire Protection" (National Fire Protection Association 2011). Approximately 1 million fires are reported to NFIRS each year by over 22 000 fire departments. NFIRS is the world's largest, national database of fire incident information; however, fire department participation in NFIRS is voluntary. Consequently, not all fire departments participate in NFIRS, making the incident data a partial, and potentially non-random, census.

As described in the NFIRS reference guide, the NFIRS data system is broken into 11 modules[2] with the *Basic Module* as the primary one (FEMA 2010). In this module, general information is captured for every reported incident (emergency call). There are additional modules for different types of incidents. Information is generally obtained at the scene by emergency response personnel. Within each module there are required fields and optional fields. Frequently, fire departments will provide aid to one another. In these instances, an indication is made in the NFIRS system. For this analysis, only the non-aid fires recorded in NFIRS are counted as a fire incident.

Not every module is relevant to all incidents (e.g., the *Structure Fire* module might not be relevant to an outside fire incident); therefore, only some modules are mandatory (e.g., the *Basic Module*), while others are not (e.g., the *Arson Module*). Even within the mandatory modules, only some of the field elements are mandatory (e.g., Incident Type

---

[2] The modules include the following: basic module, fire module, structure fire module, civilian fire casualty module, fire service casualty module, EMS module, hazardous materials module, wildland fire module, apparatus or resources module, personnel module, and the arson module.

within the *Basic Module*). The *Basic Module* is completed for every reported emergency call to which the participating department responds. It requests information about the location of the incident, incident type, alarm time and response time, actions taken (e.g., extinguishment), property losses, casualties, and property use.

For non-confined fires (see footnote 1), the fire department is required to fill out the *Fire Module* or the *Wildland Fire Module*. These modules request information about the ignition (e.g., Area of Origin, Heat Source, Item First Ignited), cause, and factors contributing to ignition. The *Structure Fire Module* is required for all non-confined structure fires. Structures can include "buildings, open platforms, bridges, roof assemblies over open storage or process areas, tents, air-supported structures, and grandstands" (p. 5-3, U.S. Fire Administration 2010). This module requests information about the structure type, building status, building height and floor size, fire origin, fire spread, and the presence and operation of smoke detectors and automatic fire sprinklers. There are additional fields in each of the modules above; however, most of them are optional fields.

For this analysis, a number of variables (i.e., incident factors) are utilized, which are reported within the module categories described above. They are listed below with the definition provided by the NFIRS reference guide (U.S. Fire Administration 2010):

Area of Fire Origin: "The primary use of the area where the fire started within the property. The area of origin may be a room, a portion of a room, a vehicle, a portion of a vehicle, or an open area devoted to a specific use. Every fire has an area of fire origin."

Cause of Ignition: "The general causal factor that resulted in a heat source igniting a combustible material. The cause could be the result of a deliberate act, mechanical failure, or act of nature."

Factors Contributing to Ignition: "The contributing factors that allowed the heat source and combustible material to combine to ignite the fire."

6

Fire Spread: "The extent of fire spread in terms of how far the flame damage extended. The extent of flame damage is the area actually burned or charred and does not include the area receiving only heat, smoke, or water damage."

Heat Source: "The heat source that ignited the Item First Ignited to cause the fire."

Human Factors Contributing to Ignition: "The human condition or situation that allowed the heat source and combustible material to combine to ignite the fire."

Incident Type: "This is the actual situation that emergency personnel found on the scene when they arrived. These codes include the entire spectrum of fire department activities from fires to EMS to public service."

Item First Ignited: "The use or configuration of the item or material first ignited by the heat source. This [variable] identifies the first item that had sufficient volume or heat intensity to extend to uncontrolled or self-perpetuating fire."

Item Contributing Most to Flame Spread: "The item contributing most to flame spread, if different from the Item First Ignited." This includes objects such as a sofa, bedding, or a structural component.

Type of Material First Ignited: "The composition of the material in the item first ignited by the heat source. The type of material ignited refers to the raw, common, or natural state of the material. The type of material ignited may be a gas, flammable liquid, chemical, plastic, wood, paper, fabric, or any number of other materials."

Type of Material Contributing Most to Flame Spread: "The type of material contributing most to flame spread, if different from the Type of Material First Ignited." This includes materials such as fabric, wood, or gasoline.

Property Use: "Each individual property has a specific use, whether a structure or open land. This entry refers to the actual use of the property where the incident occurred, not the overall use of mixed use properties of which the property is part. The intent of this entry is to specify the property use, not the configuration of the building or other details of the property."

Residential fires involving upholstered furniture can be identified two ways – based on the Incident Type and Property Use, to determine if the incident is a residential structure fire, and either by (1) Item First Ignited or (2) Item Contributing Most to Flame Spread. (Item Contributing Most to Flame Spread is an optional field, which tends to get populated less than mandatory fields.) For this analysis, residential refers to 1- and 2-family detached dwellings (NFIRS code: 419) and multifamily dwellings (NFIRS code: 429). Structure fires are categorized into non-confined (NFIRS codes: 111-112 and 120-123) and confined (NFIRS codes: 113-118). For both Item First Ignited and Item Most Contributing to Flame Spread, upholstered furniture is identified by code 21 ("upholstered sofa, chair, vehicle seats"). This category excludes items such as cabinetry, desks, and non-upholstered chairs. It also does not include mattresses, pillows, or bedding. For this analysis, data from 2006 to 2010 are used. A summary of the (raw) NFIRS data for non-confined residential fires involving upholstered furniture is shown in Table 2.1 while that for the confined fires is shown in Table 2.2.[3]

---

[3] A confined fire in NFIRS is a fire confined to 1) a container, 2) chimney, 3) incinerator, 4) burner/boiler, 5) compactor/rubbish or 6) trash. If a fire is confined to these items it is not clear how upholstered furniture is involved, but there are fires with these designations; therefore, they are calculated separately.

Table 2.1: Non-Confined Residential Fires Recorded in NFIRS that Involve Upholstered Furniture

		2006	2007	2008	2009	2010
IFI	Incidents	3105	3053	2937	2566	2714
	Injuries	346	291	354	322	275
	Fatalities	120	126	106	110	95
	Losses ($ million)	198.3	101.8	106.9	93.5	81.0
ICMFS	Incidents	659	641	641	579	608
	Injuries	94	81	81	110	78
	Fatalities	29	35	38	16	36
	Losses ($ million)	29.2	39.1	33.9	29.3	26.6

Table 2.2: Confined Residential Fires Recorded in NFIRS that Involve Upholstered Furniture

		2006	2007	2008	2009	2010
IFI	Incidents	25	29	18	30	24
	Injuries	0	1	1	0	0
	Fatalities	0	0	0	0	0
	Losses (dollars)	3282.0	7207.0	8650.0	12305.0	10913.0
ICMFS	Incidents	1	0	1	0	2
	Injuries	0	0	1	0	0
	Fatalities	0	0	0	0	0
	Losses (dollars)	501.0	0.0	800.0	0.0	22000.0

In addition to NFIRS data, both the conventional scaling method and the data imputation method developed in this report utilize data from the National Fire Protection Association (NFPA) to scale the aggregated NFIRS data into national statistics. (The two methods only differ in how they allocate the unknowns within the NFIRS data.) The NFPA collects annual survey data from municipal fire departments. It is a survey sent to more than 20 000 of the approximate 30 000 total fire departments. The data from the approximate 3000 fire departments that respond is used to produce nationwide estimates of residential and non-residential structure fires, vehicle fires, and outside and other fires (Karter 2010). NFPA's estimates for residential fires, which are appropriate for scaling NFIRS data involving upholstered furniture, are shown in Table 2.3. The NFIRS data for residential fires between 2006 and 2010 (i.e., fires with incident type 111 through 123 and have property use 419 or 429) represents approximately 67 % of the fires estimated by the NFPA.

Table 2.3: NFPA Residential Fire Estimates

	2006	2007	2008	2009	2010
Fires	412500	414000	403000	377000	384000
Injuries	12925	14000	13560	13050	13800
Deaths	2620	2895	2780	2590	2665
Losses ($billion)	7.0	7.5	8.6	7.8	7.1

2.2 Imputation Method

The data imputation technique ultimately allocates the unknowns using (two) statistical models relating non-missing, reported fire incident factors to the likelihood that the (1) Item First Ignited (IFI) or the (2) Item Contributing Most to Flame Spread (ICMFS) was upholstered furniture. The parameters in each statistical model are estimated using a sample of incidents with known IFI or ICMFS data.

The conventional scaling and data imputation approaches will yield similar results should the fire incidents with known data be representative of the data with unknown and missing information. However, if there are systematic differences, and if these differences are correlated with other, more frequently reported data, the two approaches will yield different results. While the data imputation approach offers the potential for greater accuracy, it is computationally expensive in comparison.

2.2.1 Statistical Model

This analysis uses a type of statistical model known as a logistic regression model (StataCorp 2007), often referred to as a "logit" regression. The logistic regression models used here relate the probability that residential fire incident *j* has upholstered furniture as IFI (or ICMFS)—as a function of other reported incident factors (e.g., Area of Origin). The logit model is fit to the data via Maximum Likelihood (Greene 2008) and has the following general form:

$$Pr(y_{j,t} = 1|x_j) = \frac{\exp(x_{j,t}\beta_t)}{1 + \exp(x_{j,t}\beta_t)} \qquad (1)$$

where y is a (0/1) binary variable indicating the involvement of upholstered furniture by fire incident factor t (IFI, ICMFS) for residential fire incident j, x is a vector of variables correlated with the involvement of upholstered furniture, and β is a vector of parameters.

Each model is fit to a sample of the NFIRS data that has non-missing data entered for Incident Factor, Property Use, Item First Ignited, and Item Contributing Most to Flame Spread. Each of these models is estimated for non-confined structure fires.[4] To avoid double-counting, incidents with reported upholstered furniture as both the item first ignited and item contributing most to flame spread were included in the IFI model.

The IFI variable sets include the following categories ($\mathbf{x}_{IFI}$):
- Factors Contributing to Ignition
- Heat Source
- Area of Fire Origin
- Type of Material First Ignited
- Item Contributing Most to Spread
- Type of Material Contributing Most to Spread
- Human Factors Contributing to Ignition
- Cause of Ignition
- Incident Type
- Fire Spread

The ICMFS variable sets include the following categories ($\mathbf{x}_{ICMFS}$):
- Item First Ignited
- Factors Contributing to Ignition

---

[4] Due to the limited number of confined fires, the imputation method was only applied to non-confined fires.

- Heat Source

- Area of Fire Origin

- Type of Material First Ignited

- Type of Material Contributing Most to Spread

- Human Factors Contributing to Ignition

- Cause of Ignition

- Incident Type

- Fire Spread

The variable sets are derived from the NFIRS categories. Each NFIRS category has a list to select from. For example, there are six items to select from for the cause of ignition: 1) Intentional; 2) Unintentional; 3) Failure of equipment or heat source; 4) Act of nature; 5) Cause under investigation; 6) Cause undetermined after investigation. For statistical analysis, the categorical and ordinal items were converted into indicator (binary) variables. The data conversion meant that each variable set included multiple binary variables (e.g., cause of ignition set included seven binary variables—the six items above plus one for incidents with missing values—with one held out of the model to ensure the parameters have unique values that can be identified[5]). If data were missing for a specified variable set, then all the indicator variables in the model were zero and a separate variable indicated a missing value. This resulted in hundreds of potential variables (covariates) to include in the statistical modeling. A constraint of the logit model is it cannot include variables that perfectly predict an outcome—here, the presence or absence of upholstered furniture involvement. Thus, binary variables that describe fire conditions that only occur either with or without upholstered furniture involvement were excluded from analysis.[6] However, this information could be exploited to narrow the

---

[5] If the models are allowed to be over-parameterized, the values of the β's in Eq. (1) cannot be determined uniquely. Instead there will be an infinite number of different solutions. This situation is sometimes referred to as the dummy variable trap (Greene 2008) when it arises in regression analysis with categorical predictor variables.

[6] It is important to note that a "perfect predictor" is a misleading term, as the variable could not be used as a proxy for IFI or ICMFS. For more information on this, please see the Stata manual (StataCorp 2007).

universe of possible fire incidents involving upholstered furniture. This is left for future analysis.

2.2.2 Estimated Probabilities

After estimating the values of the parameters in the model given by Eq. 1 using data randomly sampled from the NFIRS database, the fitted model can be used to produce estimated probabilities of upholstered furniture involvement for any particular fire using Eq.2:

$$\hat{p}_{j,t} = \frac{\exp(x_{j,t}\widehat{\beta_t})}{1 + \exp(x_{j,t}\widehat{\beta_t})} \tag{2}$$

where $\hat{p}$ denotes the predicted probability of incident factor t being upholstered furniture, and $\widehat{\beta}$ is the vector of estimated parameters. These estimated probabilities can be interpreted as the likelihood of upholstered furniture involvement, given other fire-related information.

2.2.3 Estimated NFIRS Residential Fires Involving Upholstered Furniture

Based on the estimated probabilities, the total number of non-confined fire incidents, F, of factor t (F_t), involving upholstered furniture can be computed as:

$$F_t = \sum_{i}^{n_{n,t}} y_{i,t} + \sum_{j}^{n_{u,t}} \hat{p}_{j,t} \tag{3}$$

where $n_{n,t}$ is the total number of fire incidents with known incident data on factor t, $n_{u,t}$ is the total number of fire incidents with missing or unknown incident data on factor t, and $n_n + n_u = N$, the total number of NFIRS-reported non-confined fire incidents.

The total number of fatalities, injuries, and economic loss (property and content loss) from fires involving residential upholstered furniture can be computed as:

$$O_{t,k} = \sum_{i}^{n_{n,t}}(y_{i,t} * o_{i,t,k}) + \sum_{j}^{n_{u,t}}(\hat{p}_{j,t} * o_{j,t,k}) \tag{4}$$

where $O_{t,k}$ is the total number of outcomes of k (fatalities, injuries, losses) resulting from fire incidents of factor t, involving residential upholstered furniture, and $o_{i,t,k}$ and $o_{j,t,k}$ are the incident-level outcomes. As shown in Eq. 4, the missing and unknown outcomes are allocated based on the likelihood that the incident involved upholstered furniture, and resemble expected values. The likelihood is determined by other variables in the NFIRS data.

The values resulting from Eqs. 3 and 4 can be scaled using the "national scaling ratio," as detailed above to produce national fire statistics, at least for the non-confined fires. (See Appendix A for more detail.) Again, because the number of non-missing confined fires with known IFI or ICMFS information is so few, it does not lend itself to the imputation technique. Therefore, we allocate the unknown and missing confined fires involving upholstered furniture similar to the method used by NFPA.

2.3 Model Specification & Selection

2.3.1 Model Specification

For IFI, fourteen different model specifications were evaluated—i.e., different combinations of predictor variables were included in Eq. 1 to determine the combination that produced the most accurate prediction of upholstered furniture involvement. Combinations of variables were created from the 10 variable sets (listed in Section 2.2.1) plus a variable indicating the state where the fire occurred along with month and year dummy variables. Some variables were removed due to collinearity. The model comparison exercise required in excess of seven days of computation time using Stata SE 12 (64-bit) on a desktop computer with a 64-bit Windows 7 Enterprise operating system with two processors running at 2.4GHz.

14

The comparison of multiple model specifications (i.e., the fourteen discussed above) was only done for the IFI model. Multiple ICMFS model specifications were not compared due to the observed lack of variation in IFI accuracy (results not shown) and to time constraints.[7] Thus, the model specification selected for IFI is used in modeling ICMFS.

2.3.2 Model Selection

Each model specification was evaluated using random samples of 20 000 residential fire incidents with known IFI from the five years examined. These training samples were used to fit the statistical models (Eq. 1). Validation samples were created by randomly sampling an additional 20 000 observations with known IFI data that were not part of the training data.[8] The validation samples were used to create estimated probabilities of upholstered furniture involvement based on the statistical modeling results from the training data (Eq. 2). The sum of the estimated probabilities of the validation sample was then compared to the total number of fire incidents involving upholstered furniture in the validation sample. This process was repeated 1000 times.

To identify the most accurate specifications, for the purposes of this analysis, the root mean squared error (RMSE) (Makridakis, 1998) was computed.[9] Again, due to computing demands, this was limited to the IFI model. The best performing model (i.e., lowest RMSE related to the number of fire incidents) was selected. This model had an average relative prediction error rate of 3.6 % for the number of fire incidents.[10] Applying a similar specification to the ICMFS resulted in an average relative prediction error rate of 5.4 %.[11] For the estimation of deaths, injuries, and losses the average relative

---

[7] In addition, an earlier version of the ICMFS model, for which multiple model specifications were compared, produced results qualitatively similar to the IFI results.

[8] Occasionally, training and validation samples were less than 20 000 observations each due to missing covariate data.

[9] There are two primary categories of modeling that utilize regression: explanatory models and predictive models. This paper presents the latter. Extreme caution should be exercised when drawing conclusions regarding causation from a predictive model, as it is not designed for this purpose. The predictor variables in the regression part of this model do not in any way explain why or how fires with upholstered furniture occur. Any correlating variable can be used as a predictor.

[10] Relative prediction error rate is computed as (100 %)*|(predicted value − actual value)/(actual value)|.

[11] The final specifications differ between the IFI and ICMFS models in the following ways: (1) the IFI model includes variables describing the item contributing most to flame spread; the ICMFS model includes

prediction error rates for the IFI model were 23.9 %, 12.1 %, and 14.5 %, respectively. For the estimation of deaths, injuries, and losses the average relative prediction error rates for the ICMFS model were 42.3 %, 27.9 %, and 10.5 %. The average relative prediction error rates for deaths and injuries are higher than for incidents because there are fewer occurrences of deaths and injuries, making them more difficult to estimate. The average numbers of deaths were 25 for IFI and 19 for ICMFS out of 20 000 observation samples. The average number of injuries was 75 for the IFI and 59 for the ICMFS. Losses also have a higher average relative prediction error rate, which is possibly related to sporadic reporting of losses and variance in loss assessment. The indicator variables that were statistically significant at the 5 % level in the final statistical models are listed in Table 2.4.

Appendix B presents the z-scores[12] from the logit models averaged over the 1000 runs for each model type. The z-scores provide a measure of association between each of the variables and the probability of upholstered furniture involvement, with their signs indicating direction, and their absolute magnitude indicating strength. For IFI, the three variables with the largest positive (also in absolute terms) z-scores are (1) type of material fire ignited – *fabric, fiber, cotton, blends, rayon, wool*; (2) area of fire origin – *common room, den, family room, living room, lounge*; and (3) item contributing most to flame spread – *upholstered sofa, chair, vehicle seat*. For ICMFS, the three variables with the largest positive (also in absolute terms) z-scores are (1) area of fire origin – *common room, den, family room, living room, lounge*; (2) area of fire origin – *function area, other*; and (3) item first ignited – *undetermined*.

---

variables describing the item first ignited; and (2) due to collinearity, some variables in the IFI model were dropped in the ICMFS model, while some variables in the ICMFS model were dropped in the IFI model.
[12] For each variable its z-score ('standard score') equals its estimated parameter divided by the estimated parameter's standard error. Scores equal to or exceeding the absolute value of 1.96 indicate statistical significance at the 5% level.

Table 2.4: Indicator Variables Statistically Significant at the 5 % Level

IFI	NFIRS Category
Common room, den, family room, living room, lounge	Area of Fire Origin
Function area, other	Area of Fire Origin
Cooking area, kitchen	Area of Fire Origin
Vehicle storage area; garage, carport	Area of Fire Origin
Exterior balcony, unenclosed porch	Area of Fire Origin
Courtyard, patio, porch, terrace	Area of Fire Origin
Sawn wood, including all finished lumber	Type of Material First Ignited
Fabric, textile, fur, other	Type of Material First Ignited
Fabric, fiber, cotton, blends, rayon, wool	Type of Material First Ignited
Undetermined	Type of Material First Ignited
Upholstered sofa, chair, vehicle seats	Item Contributing Most to Flame Spread
Mattress, pillow	Item Contributing Most to Flame Spread
Bedding; blanket, sheet, comforter	Item Contributing Most to Flame Spread
Heat source too close to combustibles.	Factors Contributing to Ignition
Mechanical failure, malfunction, other	Factors Contributing to Ignition
Electrical failure, malfunction, other	Factors Contributing to Ignition
Equipment unattended	Factors Contributing to Ignition
ICMFS	**NFIRS Category**
Common room, den, family room, living room, lounge	Area of Fire Origin
Function area, other	Area of Fire Origin
Cooking area, kitchen	Area of Fire Origin
Laundry area, wash house (laundry)	Area of Fire Origin
Item First Ignited, Other	Item First Ignited
Furniture, utensils, other	Item First Ignited
Non-upholstered chair, bench	Item First Ignited
Electrical wire, cable insulation	Item First Ignited
Magazine, newspaper, writing paper	Item First Ignited
Rubbish, trash, or waste	Item First Ignited
Undetermined	Item First Ignited
Sawn wood, including all finished lumber	Type of Material First Ignited
Fabric, textile, fur, other	Type of Material First Ignited
Fabric, fiber, cotton, blends, rayon, wool	Type of Material First Ignited

2.3.4 Imputation of Unknown & Missing Data

Based on the IFI model specification, a procedure analogous to the validation procedure was replicated for residential fire data with unknown and missing IFI fields; then the fitted models, using the specification with the lowest RMSE from the validation procedure, were used with this data to impute the missing values. Some observations were dropped due to issues of collinearity. This process produced, for each fire incident with missing information, the probability that the IFI was upholstered furniture, given the

other fire data reported for that incident. (The same procedure was performed using the ICMFS, using a similar model specification.) The probabilities were then used to estimate the number of fatalities, injuries, and economic losses. For each model, this process was repeated 1000 times. Iterating this process provides a distribution of values of the total number of fire incidents, fatalities, injuries, and economic losses, rather than just point estimates, as produced from the conventional scaling method. This information is useful in assessing significance of time trends, for example.

18

3 Upholstered Furniture Fire Estimates

Applying the model specification to fire incidents with known Item First Ignited (IFI) and Item Contributing Most to Flame Spread (ICMFS), two statistical models were used to estimate the probability that IFI or ICMFS were upholstered furniture fires for those fire incidents with missing IFI or ICMFS data. The predicted probabilities were used to estimate the total number of incidents, fatalities, injuries, and losses from non-confined fires that involved upholstered furniture, but had missing or unknown data. Combining the incident, fatality, injury, and loss data from those incidents with known IFI and ICMFS information with the allocated confined fires with missing IFI and ICMFS data, these numbers were 'scaled-up" to represent statistics for the U.S., using the procedure consistent with Ahrens (2011).

The confined fires with missing or unknown IFI and ICMFS data [see Table 2.2] were allocated over the distribution of the known, similar to Ahrens (2011).
The small number of confined fires made it difficult to estimate the number of fires using a logit regression; thus, the number of confined fires involving upholstered furniture was estimated using the method used by NFPA. The annual totals of confined fires that involved upholstered furniture as the Item First Ignited (IFI) ranged between 180 and 390 (see Table 3.1), making them a relatively small proportion of the total number of upholstered furniture fires. Those that involved upholstered furniture as the Item Contributing Most to Flame Spread (ICMFS) ranged between 0 and 1300. The small number of confined fires with a known ICMFS combined with the very few fires with ICMFS values that were designated as upholstered furniture resulted in a wide range for the confined fires where upholstered furniture was the item contributing most to flame spread. This is true for all four categories (i.e., incidents, injuries, fatalities, and losses).

The number of non-confined fires was scaled into national estimates using a method consistent with Ahrens (2011), but the allocation of incidents with missing or unknown IFI or ICMFS data was performed using a logit regression, as described in Section 2. The estimated number of non-confined residential fires involving upholstered furniture as IFI,

Table 3.1: Estimated Number of Confined Residential Fires Involving Upholstered Furniture

		2006	2007	2008	2009	2010
IFI	Incidents	390	320	180	290	190
	Injuries	0	6	5	0	0
	Fatalities	0	0	0	0	0
	Losses ($thousand)	21	33	45	51	57
ICMFS	Incidents	1300	0	450	0	400
	Injuries	0	0	110	0	0
	Fatalities	0	0	0	0	0
	Losses ($thousand)	96	0	83	0	1000

for the U.S., is shown in Table 3.2. Since the allocation of the missing or unknowns were made using a sampling technique, the minimum, maximum, average, and median value were selected from the 1000 iterations described in Section 2.3 to provide some context. For example, in 2010, the average number of non-confined residential fire incidents with upholstered furniture as the Item First Ignited is estimated at 5000. Given the uncertainty in allocating the missing or unknown, based on the modeling approach, the actual number of fires is likely between 4500 and 5700.

Table 3.2: Estimated Number of Non-Confined Residential Fires Involving Upholstered Furniture as Item First Ignited (Imputation Approach)

		2006	2007	2008	2009	2010
Incidents	Minimum	6200	5700	5300	4600	4500
	Maximum	7400	7300	6700	5700	5700
	Average	6800	6400	5900	5100	5000
	Median	6800	6400	5900	5100	5000
Injuries	Minimum	780	680	760	650	570
	Maximum	870	810	890	740	670
	Average	820	740	810	690	610
	Median	820	740	810	690	610
Fatalities	Minimum	290	300	270	240	210
	Maximum	330	370	340	280	260
	Average	310	330	290	260	230
	Median	310	330	290	260	230
losses ($million)	Minimum	500	230	280	240	200
	Maximum	570	340	390	320	270
	Average	530	280	330	280	230
	Median	530	280	330	280	230

Table 3.3 is similar to Table 3.2, but for Item Contributing Most to Flame Spread. The values in Table 3.3 exclude any fires where the Item First Ignited was upholstered furniture. The values from Table 3.2 and Table 3.3 are, therefore, additive. Together, these represent the total number of fires where upholstered furniture played a significant role in either the ignition or spread of the fire. The number of fires incidents that involved upholstered furniture as the Item Contributing Most to Flame Spread had a larger range than that for the Item First Ignited. Table 3.4 provides the total estimated number, for the U.S., of confined and non-confined residential fires involving upholstered furniture either as the Item First Ignited or as the Item Contributing Most to Flame Spread. Figure 3.1 graphs the minimum, maximum, and average estimated number of residential fires involving upholstered furniture from 2006 to 2010. Based on the figure, it appears the numbers of fires are slightly decreasing. Despite this decrease, residential upholstered furniture fires cause more deaths per fire than the average residential fire. According to NFPA data in Table 2.3, there were 6.9 deaths per thousand residential fires. According to the estimates in Table 3.4, there are 39.2 deaths per thousand residential upholstered furniture fires.

Table 3.3: Estimated Number of Non-Confined Residential Fires Involving Upholstered Furniture as Item Contributing Most to Flame Spread (Imputation Approach)

		2006	2007	2008	2009	2010
Incidents	Minimum	2700	2500	2800	2500	2400
	Maximum	7000	7200	7500	7200	7800
	Average	4400	4300	4600	4100	4100
	Median	4300	4300	4500	4000	4100
Injuries	Minimum	290	280	290	310	260
	Maximum	390	380	430	410	380
	Average	330	330	350	360	320
	Median	330	330	350	360	320
Fatalities	Minimum	110	120	140	77	120
	Maximum	150	180	220	130	180
	Average	130	150	170	100	140
	Median	130	150	170	100	140
losses ($million)	Minimum	140	130	160	140	130
	Maximum	220	210	290	250	230
	Average	170	170	210	190	170
	Median	170	170	210	180	170

Table 3.4: Estimated Number of Confined and Non-Confined Residential Fires Involving Upholstered Furniture either as Item First Ignited or Item Contributing Most to Flame Spread

		2006	2007	2008	2009	2010
Incidents	Minimum	11000	8600	8700	7400	7500
	Maximum	16000	15000	15000	13000	14000
	Average	13000	11000	11000	9400	9700
	Median	13000	11000	11000	9400	9700
Injuries	Minimum	1100	960	1200	960	830
	Maximum	1300	1200	1400	1100	1100
	Average	1200	1100	1300	1100	930
	Median	1200	1100	1300	1100	930
Fatalities	Minimum	400	420	410	310	330
	Maximum	490	550	550	410	440
	Average	440	480	460	360	380
	Median	440	480	460	360	380
losses ($million)	Minimum	640	360	450	380	320
	Maximum	790	560	680	570	500
	Average	700	450	540	460	390
	Median	700	450	540	460	390

Figure 3.1: Minimum, Maximum, and Average of the Estimated Number of Confined and Non-Confined Residential Fires Involving Upholstered Furniture either as Item First Ignited or Item Contributing Most to Flame Spread

4 A Comparison of Estimates Produced by the Scaling and Imputation Approaches

As discussed previously, the conventional national fire statistics are based on NFIRS incident data—with complete/known data—that are scaled using two ratios (e.g., see Ahrens [2011]). In this report, we explore the use of a data imputation technique as an alternative method to "allocate the unknowns." There are some differences between the estimates using the conventional method compared to the method in this report.

Item First Ignited

Fire incidents with missing or unknown information regarding the Item First Ignited (IFI) represented 40 % of all non-confined home structure fires reported to NFIRS from 2006 to 2010. These fire incidents accounted for 57 % of all civilian (non-fire service) fatalities.

The NFPA estimates (National Fire Protection Association 2013), using the scaling method, indicate that between 2006 and 2010 there was an annual average of 6700 home structure fires with upholstered furniture being the Item First Ignited. These fires caused, on average, 480 civilian deaths, 840 civilian injuries, and $427 million in direct property damage (National Fire Protection Association 2013). This report, using the imputation method, estimates these numbers to be 5800 fires, 300 civilian deaths, 740 civilian injuries, and $328 million in direct property damage.

Item Contributing Most to Flame Spread

Fire incidents with missing or unknown information regarding the Item Contributing Most to Flame Spread (ICMFS) represented 75 % of all non-confined home structure fires reported to NFIRS from 2006 to 2010 that were reported to have spread beyond the object of origin. (Fires that spread beyond object of origin were reported in 90 % of all non-confined home fire incidents.) These fire incidents accounted for 69 % of all civilian fatalities of non-confined fires that were reported.

23

Hall estimates, using the scaling method,[13] that there was an annual average of 2200 home fires with upholstered furniture as Item Contributing Most to Flame Spread and Item First Ignited identified as something other than upholstered furniture (Hall 2014). These fires produced 130 civilian deaths, 280 civilian injuries, and $138 million in losses. These figures are directly comparable to the statistics produced in this report, and can be added to the IFI figures above. This report, using the imputation method, estimates that there was an annual average of 4300 home fires with upholstered furniture as Item Contributing Most to Flame Spread and Item First Ignited identified as something other than upholstered furniture. These fires produced 140 civilian deaths, 340 civilian injuries, and $182 million in losses.

Upholstered Furniture Involvement in Residential Home Fires

The imputation method estimates 10 100 residential home fires involving upholstered furniture resulted in 440 occupant fatalities, 1100 occupant injuries, and $510 million, on average each year, from 2006 to 2010, based on data reported in Table 3.4. Using nonparametric bootstrapping (Kennedy 2003), 95 % confidence intervals were produced for each year's estimate. Figures 4.1 to 4.4 present the estimated annual number of fires involving upholstered furniture (Figure 4.1) and the number of resulting occupant fatalities (Figure 4.2), occupant injuries (Figure 4.3), and total direct losses (Figure 4.4). Plotted along with these values are the combined values published by NFPA, including Hall's estimates (2014), produced using the method of scaling NFIRS data, discussed in the previous section. (Due to data limitations, the yearly estimates pertaining to Item First Ignited were combined with averaged yearly estimates pertaining to Item Contributing Most to Flame Spread.)

Figure 4.1 demonstrates the imputation method estimates more fires involving upholstered furniture than using the scaling approach, on average, although the combined scaling-based numbers often overlap the lower confidence interval. Confidence intervals

---

[13] It should be noted that the scaling method developed in Hall (2014) is different than that used in NFPA's report on upholstered furniture flammability for item first ignited. Hall's method employs allocating the missing and unknown within fire size groups (e.g., beyond object but confined to room of origin, beyond room but confined to floor of origin).

for the NFPA estimates are not produced, but assuming even a small interval band, suggests that the scaling and imputation methods produced results that are not statistically distinguishable. Figure 4.2 shows that the combined scaling-based numbers produce estimates of occupant fatalities that are statistically larger than those produced using the imputation method. Figure 4.3 graphs the number of annual occupant injuries and shows a high degree of agreement between the two methods. The patterns are largely similar and the annual values tend to be statistically indistinguishable. Figure 4.4 presents losses by year; only in years 2006 and 2010 are the values statistically different between the methods, although the difference in 2010 is far less than in 2006.

In general, the scaling and imputation methods produce similar statistical results for the number of residential home fires involving upholstered furniture, as well as for numbers of occupant injuries and direct property losses. However, the disagreement between the estimated numbers of occupant fatalities is statistically significant. The imputation method estimates about a third fewer occupant fatalities, the difference being in the number fires involving upholstered furniture from direct ignition rather than from flame spread. Overall, the imputation method suggests that fires involving upholstered furniture resulted in fewer deaths on a per reported fire basis than currently reported, although residential upholstered furniture fires cause more deaths per fire than the average residential fire. For fires involving upholstered furniture from direct ignition, there were 7 deaths per 100 home fires involving upholstered furniture (via scaling) versus 5 deaths per 100 home fires (via imputation). For fires involving upholstered furniture from flame spread, the fires were less deadly, on a per reported fire basis, but produced the same number of fatalities, as twice as many fires were estimated – 6 deaths per 100 home fires with significant flame spread (via scaling) versus 3 deaths per 100 home fires (via imputation).

A difference in the estimates produced from the two methods occurs when the proportional allocation assumption used by the scaling approach is not supported by the data. The magnitude of the difference depends on the robustness of the assumption along with the number of unknowns requiring allocation. Of non-confined home fires, 30 % of

25

the injuries, 40 % of the fire incidents, 49 % of the economic losses, and 57 % of the fatalities are associated with fire incidents with an unknown or missing IFI. Therefore, if the proportionality assumption fails to hold (for each category), we would expect the greatest agreement in estimates to occur between injuries, which has the lowest percentage of unknowns, and the least between fatalities, which has the highest percentage of unknowns, with fires and economic losses falling somewhere in-between.

Further, the imputation method evaluates the likelihood that an incident involved upholstered furniture and keeps the outcome linked (fatality, injury, economic loss) to produce estimates, unlike the scaling approach, where each estimate is produced independent of the others. Of the known home fires with IFI identified as upholstered furniture, the rate of fatalities, injuries, and economic losses *occurred* at rates of 3.9, 11.0 and $4.0 million per 100 home fires, respectively. Of the unknown fires attributed to upholstered furniture by the scaling method, the *implied* rates of fatalities, injuries, and economic losses were 7.6, 7.1, and $5.7 million per 100 home fires, respectively. (The rates are implied in the sense they are an artifact of the scaling.) The scaling approach implies that the unknown fires were deadlier and costlier, but less injurious than the home fires with known IFI identified as upholstered furniture. However, it has no mechanism to verify these implications. By comparison, of the unknown fires likely to have involved upholstered furniture by the imputation method, the *observed* rates of fatalities, injuries, and economic losses were actually 3.6, 7.0, and $4.9 million per 100 home fires, respectively. There is pretty good agreement between the two methods for injuries and economic losses, suggesting the proportionality assumption held for these, but not for fatalities, which is reflected in Figures 4.2 to 4.4.

26

Figure 4.1: Estimated Number of Fire Incidents, Along with their 95 % Confidence Interval, of Confined and Non-Confined Residential Fires Involving Upholstered Furniture either as Item First Ignited or Item Contributing Most to Flame Spread, Compared with NFPA Published Estimates.

Figure 4.2: Estimated Number of Civilian Fire Fatalities, Along with their 95 % Confidence Interval, of Confined and Non-Confined Residential Fires Involving Upholstered Furniture either as Item First Ignited or Item Contributing Most to Flame Spread, Compared with NFPA Published Estimates.

Figure 4.3: Estimated Number of Civilian Fire Injuries, Along with their 95 % Confidence Interval, of Confined and Non-Confined Residential Fires Involving Upholstered Furniture either as Item First Ignited or Item Contributing Most to Flame Spread, Compared with NFPA Published Estimates.

Figure 4.4: Estimated Direct Fire Losses, Along with their 95 % Confidence Interval, of Confined and Non-Confined Residential Fires Involving Upholstered Furniture either as Item First Ignited or Item Contributing Most to Flame Spread, Compared with NFPA Published Estimates.
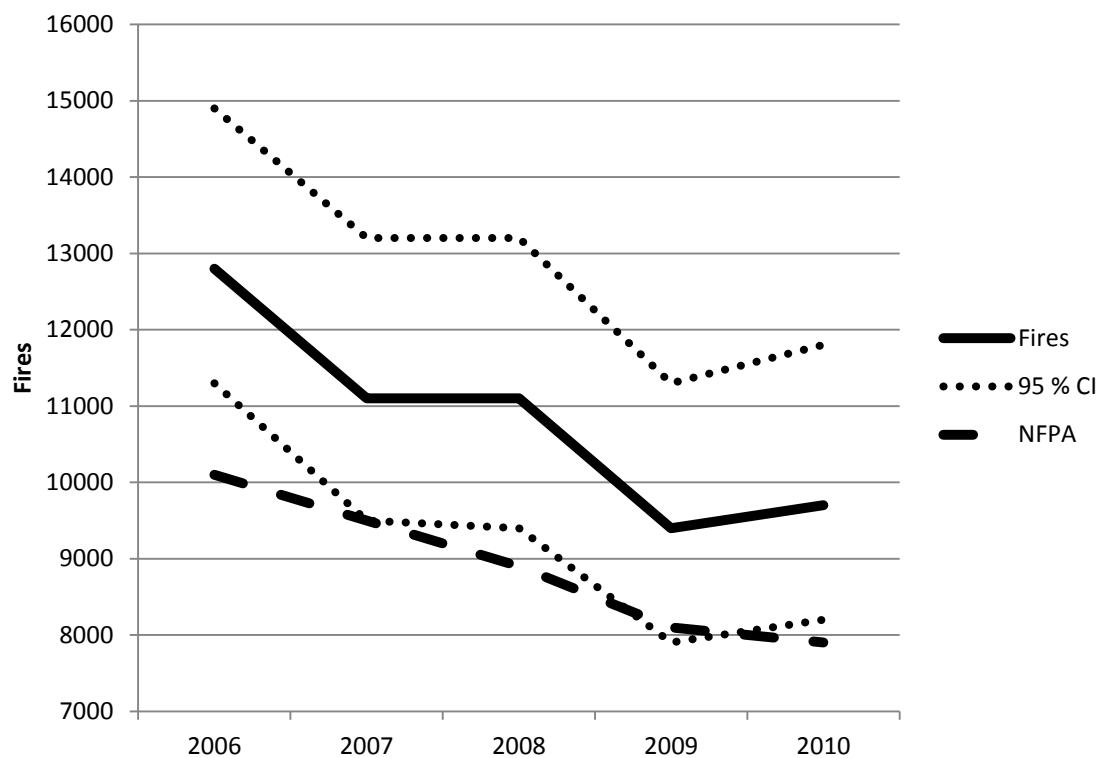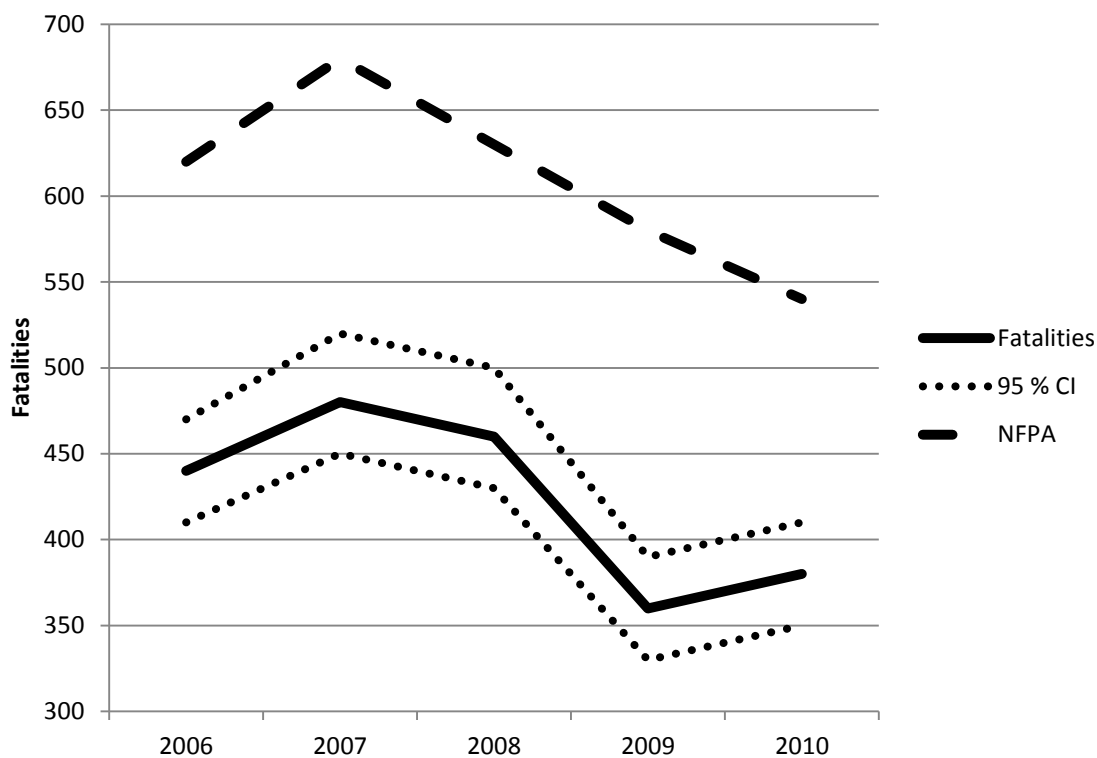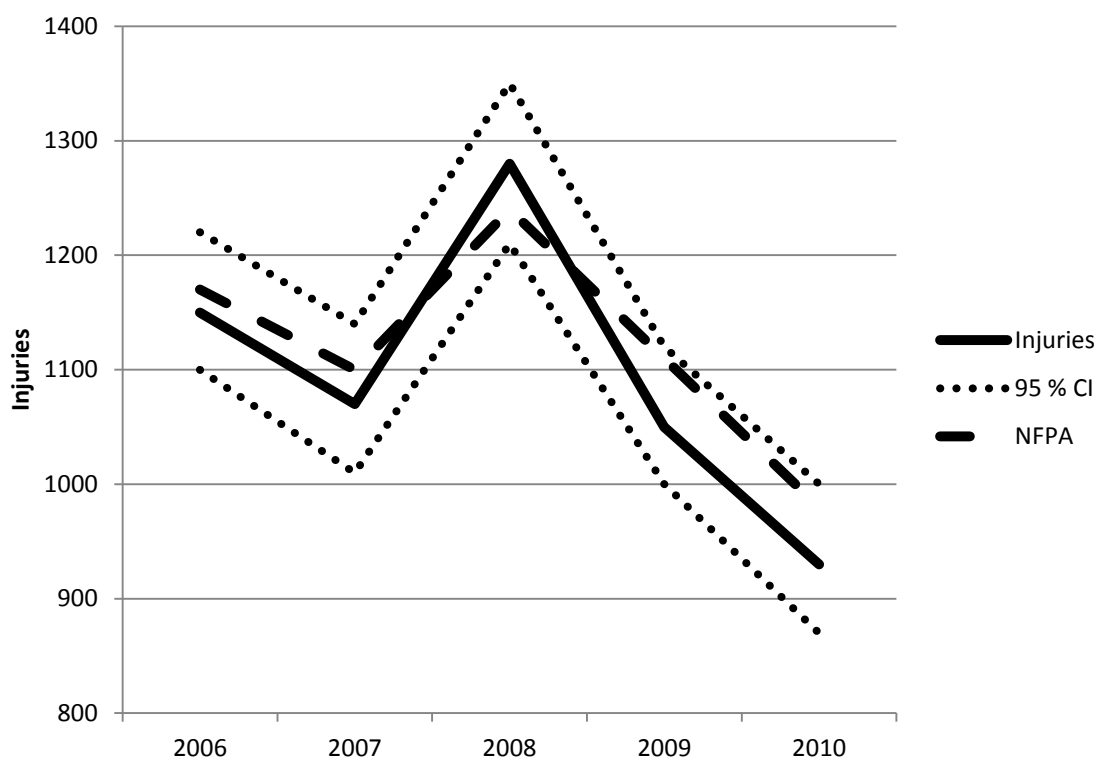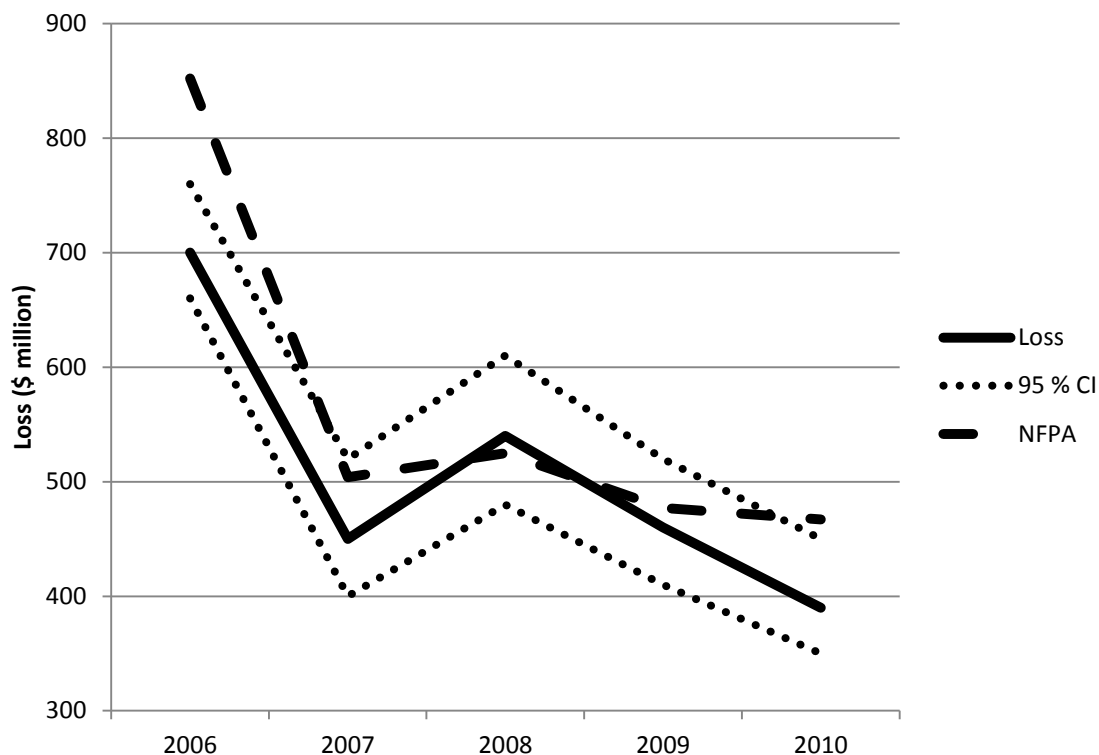
5 Summary

This Technical Note was developed in an attempt to explore alternative ways to identify the missing information regarding the "item first ignited" (IFI) or "item most contributing to flame spread" (IMCFS) that were likely to have involved upholstered furniture for residential fire incidents reported to the National Fire Incident Reporting System (NFIRS). This approach differs from the conventional scaling approach that allocates the missing data using the distribution of incidents with known values. The approach described employed a data imputation technique to determine the likelihood of upholstered furniture involvement based on statistical correlations with other, non-missing, information describing the fire incident. This model utilized data from 2006 through 2010, but did not examine the effect of including more or less data from adjacent years. More data often results in more accuracy; however, the correlations could potentially change over time, making it unclear whether additional data from adjacent years would improve accuracy. Extensive validation data was used for selection of the form of the IFI model and demonstrated that the imputation method is relatively accurate, with an average relative prediction error rate of $\pm 3.6 \%$ for IFI and $\pm 5.4 \%$ for ICMFS at predicting the total number of fires by year. The average relative prediction error rates for fatalities, injuries, and losses are higher than for incidents. These were $\pm 23.9 \%$ for fatalities, $\pm 12.1 \%$ for injuries, and $\pm 14.5 \%$ for economic losses for the IFI model, and $\pm 42.3 \%$ for fatalities, $\pm 27.9 \%$ for injuries, and $\pm 10.5 \%$ for economic losses for the IFI model. Improvements for these estimates might be made with different modeling techniques and specifications. The average relative prediction error rates are not available for statistics of the current scaling method.

It is estimated that the number of residential fires involving upholstered furniture averaged between 9400 to 13 000 a year during the five-year period from 2006 to 2010. [14] These fires included those for which upholstered furniture was either the item first ignited or the item contributing most to flame spread (and not the item first ignited). These fires resulted in civilian (non-fire service) fatalities that averaged between 360 to 480 per year,

---

[14] The averages in this section are the minimum and maximum average values from Table 3.4.

civilian injuries that averaged between 930 to 1300 per year, and property and content losses that averaged between $450 million to $700 million per year.

Comparisons between national statistics produced by the imputation method developed within this report with the conventionally-used scaling method produced both similar and dissimilar results. Annual trends in the number of fires involving upholstered furniture, along with the results for occupant injuries and total direct property losses, were largely similar (generally not statistically different) between the two methods. The only significant difference appeared between occupant fatalities, with the imputation method implying fires involving upholstered furniture were less deadly on a per reported fire basis. Overall, the imputation method resulted in estimating the same number of occupant fatalities resulting from fires where upholstered furniture was the item contributing most to flame spread, but suggested there are twice as many fires as previously thought. The number of estimated occupant fatalities caused by fires involving upholstered furniture started by a direct ignition were fewer. The number of fires estimated using the conventional method was lower than that of the imputation method; however, it overlaps the lower confidence interval of the imputation method.

Future research is needed to understand why the imputation and scaling methods diverge in estimating the total number of occupant fatalities. Unlike the number of fires, occupant injuries, and direct property losses, fatalities occur in smaller numbers, which makes statistical modeling challenging. Fewer occurrences can leave statistical methods vulnerable to influential unique or rare fire incidents, which do not represent typical fire conditions or occupant behaviors. Further research can evaluate the robustness of competing methods to outliers or abnormal reporting results.

Reporting issues related to Item First Ignited and Item Contributing Most to Flame Spread were the only NFIRS fields considered. Further research could be performed to evaluate the relative performance of the scaling and imputation method to other critical fields contained within the National Fire Incident Reporting System. A systematic comparison would allow users of NFIRS-based fire statistics to better understand the

robustness and limitations of findings based on a partial census of all fire incidents reported in the United States. Finally, both methods scaled statistics using a national scaling ratio, which assumes reported fire incidents contained within NFIRS are representative of the fires within municipalities that do not report. This may partially explain the similarity of the results. How much this assumption influences national fire statistics is not well-known and is left for future exploration.

References

Ahrens, Marty, 2011, Home Fires that Began with Upholstered Furniture, National Fire Protection Association, Quincy, MA.

Federal Emergency Management Agency. 2010. *National Fire Incident Reporting System: Complete Reference Guide*.

Greene, Michael A., Linda E. Smith, Mark S. Levenson, Signe Hiser, Jean C. Mah. 2001 "Raking Fire Data." Presented at the Federal Committee on Statistical Methodology Research Conference 2001.
<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.159.7090>

Greene, William H. Econometric Analysis. 6th edition. (Pearson Education, Inc Upper Saddle River, NJ 2008) 482-572, 108.

Hall, John, and Beatrice Harwood. "The National Estimates Approach to U.S. Fire Statistics." *Fire Technology,* May 1989. 13.

Hall, John R. 2014. Estimating Fires When a Product is the Primary Fuel but not the First Fuel, with an Application to Upholstered Furniture. *Fire Technology*, February 2014.
<http://link.springer.com/article/10.1007%2Fs10694-014-0391-8>.

Izrael, David, David C. Hoaglin, and Michael P. Battaglia. 2004 "To Rake or Not to Rake is Not the Question Anymore with the Enhanced Raking Macro." SUGI 29 Proceedings. Montreal, Canada. <http://www2.sas.com/proceedings/sugi29/toc.html>

Karter, Michael J. 2010. *Fire Loss in the United States During 2009*. National Fire Protection Association. <http://www.nfpa.org/assets/files/pdf/os.fireloss.pdf>

Karter, Michael J, 2013, *Fire Loss in the United States 2012*, and previous reports in the series.

Kennedy, Peter. A Guide to Econometrics. 5th edition. (MIT Press: Cambridge, MA, 2003)

Makridakis, Spyros, Steven C. Wheelwright, and Rob J. Hyndman. Forecasting: Methods and Applications. 3rd edition. (John Wiley & Sons, Inc: New York, NY 1998) 45-46.

National Fire Protection Association. 2011 "NFPA 901: Standard Classifications for Incident Reporting and Fire Protection Data." <http://www.nfpa.org/codes-and-standards/document-information-pages?mode=code&code=901>

National Fire Protection Associations. 2013 "Upholstered Furniture Flammability." <http://www.nfpa.org/~/media/Files/Research/NFPA%20reports/Major%20Causes/UpholsteredFurnitureWhitePaper.pdf>

StataCorp. Stata Base Reference Manual. Volume 2: I-P. Release 10. (Stata Press: College Station, TX 2007) 178-189.

U.S. Fire Administration, National Fire Data Center, 2010. *National Fire Incident Reporting System 5.0*. Federal Emergency Management Agency, Washington, DC.

Appendix A: A Comparison of the Scaling & Imputation Approaches

The purpose of this Appendix is to compare differences between the method developed within this Technical Note and a scaling method employed by NFPA and U.S. Fire Administration. The scaling method is largely based on the work of Hall and Harwood (1989), although the approach has been modified over time to accommodate changes in NFIRS (e.g., see Ahrens [2011]). An example of the scaling method is shown below. It is not representative of all the ways the scaling method is employed, but the general approach is conceptually the same.

To simplify the comparison, the focus is on estimating the total number of home fire fatalities in the U.S. for fire incidents involving upholstered furniture (RUF) as the item first ignited (IFI). Differences exist only for non-confined fires. Due to the limited number of confined home fires that involve upholstered furniture, the data imputation technique cannot be used.

Scaling Method

The scaling method uses a simplifying assumption that the proportion of all fire deaths that is associated with RUF as the IFI is the same for the incidents with unknown IFI— i.e., *assumes the occurrence of unknown IFI is a completely random process*. A national estimate of non-confined RUF fire fatalities (NCRUFFF) is computed with the scaling method as follows,

$$NCRUFFF_{NFPA} = NCRUFFF_{NFIRS} * (Fatalities_{NFPA}/Fatalities_{NFIRS}) *$$
$$(NCFF_{NFIRS}/KNCFF_{NFIRS})$$

where,

NCRUFFF$_{NFIRS}$ is the total number of non-confined RUF fire fatalities reported to NFIRS

Fatalities$_{NFPA}$ is the total number of residential fire fatalities estimated by NFPA

Fatalities$_{NFIRS}$ is the total number of residential fire fatalities reported to NFIRS

NCFF$_{NFIRS}$ is the total number of non-confined home fire fatalities reported to NFIRS

KNCFF$_{NFIRS}$ is the total number of non-confined home fire fatalities, with known IFI, reported to NFIRS

This can be rearranged into

$$NCRUFFF_{NFPA} = (NCRUFFF_{NFIRS} + \{[UNCFF_{NFIRS}] * [NCRUFFF_{NFIRS}/KNCFF_{NFIRS}]\}) * SRF$$

where,

UNCFF$_{NFIRS}$ is the total number of non-confined home fire fatalities, with unknown or missing IFI, reported to NFIRS (i.e., NCFF$_{NFIRS}$ − KNCFF$_{NFIRS}$)

SRF is the national scaling ratio for residential fire fatalities (i.e., Fatalities$_{NFPA}$/Fatalities$_{NFIRS}$)

It can be seen that the number of fire fatalities with unknown or missing IFI information is allocated to RUF based on the proportion of non-confined RUF fire fatalities to the total number of non-confined home fire fatalities with known IFI information.

38

Imputation Method

The method develop in this Technical Note relaxes the assumption that the proportion of all fire deaths that is associated with RUF as the IFI is the same for the incidents with unknown IFI (i.e., the occurrence of unknown IFI is a completely random process). A data imputation technique is used to directly estimate the number of deaths with unknown IFI that were likely to have involved upholstered furniture as item first ignited. Instead of scaling, it allocates the fire fatalities reported with unknown or missing IFI information based on the distribution of other, known factors found correlated with the reporting of upholstered furniture as the IFI. (Scaling is still used to inflate the NFIRS based statistics into national estimates.) The imputation method assumes the fire behavior exhibited across those fire incidents with known IFI and other particular fixed factor values is similar to that of the incidents with unknown IFI with the same fixed factor values—i.e., *it assumes the occurrence of unknown IFI is a random process conditioned on the correlated factors (covariates).*

NIST's estimate of NCRUFFF is computed as follows,

$$\text{NCRUFFF}_{\text{NIST}} = (\text{NCRUFFF}_{\text{NFIRS}} + \{\textstyle\sum_i \Pr(\text{RUF}_i = 1 | X_i) * \text{UNCFF}_{\text{NFIRS},i}\}) * \text{SRF}$$

where,

> $\Pr(\text{RUF}_i = 1 | X_i)$ is the estimated probability that fire incident *i*, with unknown or missing IFI information, involved upholstered furniture given other, known factors, *X*.

The scaling and imputation method produce equal estimates when the distribution of other, known factors is the same between the sample of fire incidents with known IFI and those with unknown or missing IFI—i.e., when the occurrence of unknown IFI is a random process.

39

If, for instance, the other, known factors positively correlated with RUF fires occur less (more) frequently in the sample of incidents with unknown or missing IFI information, then the estimated number of RUF fire fatalities will be less (greater) than the scaling method.

The two methods offer different advantages. The scaling method is straightforward and quicker to compute. The imputation method utilizes additional information on the reported fire incidents, making it more accurate if correlations exist between upholstered furniture involvement and other factors. In addition, the number of fires with known IFI indicating RUF and unknown IFI estimated as RUF could trend differently.

Appendix B: COVARIATES & Z-SCORES

Table B.1. List of included covariates and their average z-score resulting from the 1000 fits of the logit model predicting (1) the probability of upholstered furniture as Item First Ignited or (2) the probability of upholstered furniture as Item Contributing Most to Flame Spread.

COVARIATES	ITEM FIRST IGNITED		ITEM CONTRIBUTING MOST TO FLAME SPREAD	
	INCLUDED	Z-SCORE	INCLUDED	Z-SCORE
AREA OF FIRE ORIGIN				
Assembly, sales areas (groups of people)				
Common room, den, family room, living room, lounge	X	16.820	X	13.649
Function area				
Function area, other	X	13.233	X	9.662
Bedroom - < 5 persons; included are jail or prison	X	0.385	X	1.492
Bedroom - 5+ persons; including barrack/dormitory	X	0.832		
Cooking area, kitchen	X	-2.982	X	-3.076
Laundry area, wash house (laundry)			X	-2.096
Storage areas				
Closet			X	-0.912
Storage: supplies or tools; dead storage	X	0.692		
Vehicle storage area; garage, carport	X	3.105	X	-0.444
Service, Equipment Areas				
Heating room or area, water heater area			X	-0.574
Structural Areas				
Exterior balcony, unenclosed porch	X	4.336	X	-1.008
Ceiling & floor assembly, crawl space b/t stories	X	0.497	X	0.354
Wall assembly	X	0.153	X	0.473
Wall surface: exterior			X	-0.977
Other area of origin				
Courtyard, patio, porch, terrace	X	2.891	X	-0.654
Item First Ignited				
Item First Ignited, Other				
Item First Ignited, Other			X	2.363
Structural Component, Finish				
Structural component or finish, other			X	-0.431
Floor covering or rug/carpet/mat			X	0.613
Interior wall covering excluding drapes, etc.			X	0.871
Structural member or framing			X	-0.974
Furniture, utensils, including built-in furniture				
Furniture, utensils, other			X	5.492
Non-upholstered chair, bench			X	2.237
Cabinetry (including built-in)			X	1.236
Appliance housing or casing			X	1.338
Soft goods, wearing apparel				
Soft goods, wearing apparel, other			X	0.717
Mattress, pillow			X	-0.045
Bedding; blanket, sheet, comforter			X	0.541
Wearing apparel not on a person			X	-0.297
Curtains, blinds, drapery, tapestry			X	0.266
Storage supplies				
Box, carton, bag, basket, barrel			X	0.947

Liquids, piping, filters				
Flammable liquid/gas - uncontained			X	0.136
Organic materials				
Cooking materials, including edible materials			X	-0.841
General materials				
Electrical wire, cable insulation			X	2.561
General materials continued				
Magazine, newspaper, writing paper			X	2.736
Rubbish, trash, or waste			X	1.727
Multiple items first ignited			X	1.049
Undetermined			X	5.901
Type of Material First Ignited				
Plastics				
Plastic	X	0.300	X	0.319
Natural Product				
Natural product, other	X	1.171	X	-0.351
Wood or paper - processed				
Wood or paper, processed, other			X	-0.937
Sawn wood, including all finished lumber	X	-3.005	X	-1.991
Fiberboard, particleboard, and hardboard			X	-0.447
Paper, including cellulose, waxed paper			X	0.268
Fabric, textiles, fur				
Fabric, textile, fur, other	X	15.518	X	1.790
Fabric, fiber, cotton, blends, rayon, wool	X	19.521	X	4.141
Undetermined	X	7.309	X	-0.447
Item Contributing Most to Flame Spread				
Structural Component, finish				
Interior ceiling cover or finish	X	0.391		
Furniture, utensils, including built-in furniture				
Upholstered sofa, chair, vehicle seats	X	15.987		
Soft goods, wearing apparel				
Mattress, pillow	X	-2.288		
Bedding; blanket, sheet, comforter	X	-2.548		
Cause of Ignition				
Cause, other			X	-0.528
Intentional			X	-1.290
Unintentional			X	-0.448
Failure of equipment or heat source			X	-0.956
Act of nature			X	-0.754
Cause under investigation			X	0.076
Incident Type				
Fires in structures other than in a building	X	-0.030	X	-0.076
Fire in mobile prop. used as a fixed struc., other	X	0.323	X	0.184
Fire in mobile home used as fixed residence	X	0.782	X	1.042
Fire in motor home, camper, recreational vehicle	X	0.753	X	0.948
Fire in portable building, fixed location			X	1.180
Fire Spread				
Confined to object of origin	X	-1.133	X	-0.700
Confined to room of origin	X	-0.670	X	-0.292
Confined to floor of origin	X	-0.080	X	0.166
Confined to building of origin	X	0.027	X	-0.173
Beyond building of origin	X	-0.040		

Factors Contributing to Ignition				
Misuse of material or product				
Heat source too close to combustibles.	X	-2.512	X	-0.120
Improper container or storage	X	-1.459		
Playing with heat source	X	-0.720	X	-0.292
Mechanical failure, malfunction				
Mechanical failure, malfunction, other	X	-2.559	X	-0.661
Electrical failure, malfunction				
Electrical failure, malfunction, other	X	-2.613	X	0.571
Unspecified short-circuit arc	X	-1.679	X	1.633
Arc, spark from operating equipment	X	-0.644	X	0.092
Operational deficiency				
Equipment unattended	X	-1.804	X	-0.468
Equipment overloaded	X	-1.173	X	0.582
Fire spread or control				
Fire spread or control, other	X	0.281	X	-0.381