An Affordability Framework for the National Flood Insurance Program

April 17, 2018

Homeland Security

Message from the Administrator

I am pleased to submit this Affordability Framework for the National Flood Insurance Program (NFIP).

The Federal Emergency Management Agency (FEMA) developed options for an affordability framework for the NFIP pursuant to section 9 of the Homeowner Flood Insurance Affordability Act (HFIAA) of 2014, Pub. L. No. 113-89, 128 Stat. 1024. The Administration has submitted an affordability proposal that considers the findings and analysis in this Affordability Framework.

FEMA is sending this framework to the following Members of Congress:

The Honorable Mike Crapo
Chairman, Senate Committee on Banking, Housing, and Urban Affairs

The Honorable Sherrod Brown
Ranking Member, Senate Committee on Banking, Housing, and Urban Affairs

The Honorable Thad Cochran
Chairman, Senate Committee on Appropriations

The Honorable Patrick Leahy
Ranking Member, Senate Committee on Appropriations

The Honorable Jeb Hensarling
Chairman, House Committee on Financial Services

The Honorable Maxine Waters
Ranking Member, House Committee on Financial Services

The Honorable Rodney P. Frelinghuysen
Chairman, House Committee on Appropriations

The Honorable Nita M. Lowey
Ranking Member, House Committee on Appropriations

Please direct inquiries related to this framework to FEMA Congressional Affairs Division at (202) 646-4500.

Sincerely,
Brock Long

Administrator

Preface

Under the Biggert-Waters Flood Insurance Reform Act of 2012 (BW-12) Congress sought to build a more sound financial framework for the National Flood Insurance Program (NFIP) by directing the Federal Emergency Management Agency (FEMA), through the NFIP Administrator, to remove the discounts for some policyholders with homes insured by the NFIP, so that policyholders would realize flood insurance rates that more accurately reflected their expected flood losses. At that time, Congress recognized that removing discounts might cause flood insurance to become unaffordable for some households and mandated in BW-12 that FEMA study flood insurance affordability. Because of concerns about rising premiums from constituents in multiple communities, Congress later passed the Homeowner Flood Insurance Affordability Act of 2014 (HFIAA), which rolled back some of the changes implemented under BW-12 and recognized additional affordability challenges associated with increased premiums required by the BW-12 implementation. HFIAA mandated that FEMA develop an affordability framework aimed at providing targeted assistance for policyholders in addition to dealing with BW-12 affordability requirements rather than the current approach that primarily provides discounted rates to properties based on their date of construction.

To respond to the congressional mandate, FEMA engaged the broader policy community, including academia, and other government agencies to develop an affordability framework. The framework presented in this document is the result of FEMA efforts in this area.

An Affordability Framework for the National Flood Insurance Program

Table of Contents

List of Tables and Figures

I. Introduction

Background

The Federal Emergency Management Agency (FEMA) administers the National Flood Insurance Program (NFIP), a federally operated insurance program created by the National Flood Insurance Act of 1968. The NFIP is a voluntary program that enables property owners in participating communities to purchase insurance protection against losses from flooding. The NFIP collects premiums and fees from its policyholders and pays claims to those policyholders for costs associated with covered flood damages.[1] The NFIP provides discounts for some insured homes and the discounts are aimed at making flood insurance more affordable, but those discounts are not delivered based on need or ability to pay. These discounts, combined with several large loss years, contributed to revenue shortfalls and resulted in NFIP borrowing to pay claims in several instances. These factors caused the NFIP to be $20.525 billion in debt to the U.S. Treasury as of April, 2018.

In response to the debt accumulated largely from Hurricanes Rita, Wilma and Katrina in 2005, Congress passed the Biggert-Waters Flood Insurance Reform Act of 2012, (BW-12).[2] BW-12 focused on strengthening the NFIP's fiscal soundness and required FEMA to eliminate subsidies for some types of policyholders and to move further toward risk-based pricing of policies.[3] Through risk-based pricing, FEMA can communicate the risk of flooding by charging higher premiums in areas where the risk of flooding is greater. BW-12 also required FEMA to charge additional fees to policyholders to cover other program costs. As a result of this transition to higher rates and increased fees, premiums rose, and resulted in public concern that the prices stemming from BW-12 were unaffordable.[4] Congress reevaluated the rate increases as a result of the public concern and subsequently passed the Homeowner Flood Insurance Affordability Act of 2014, (HFIAA).[5] HFIAA rolled back some of the changes resulting from BW-12 and focused greater attention on the issue of flood insurance affordability. Congress was concerned that as NFIP rate discounts phased out, flood insurance premiums would become increasingly unaffordable and higher premiums would create financial hardship for some households, discouraging participation in the program. FEMA notes that despite higher fees, flood insurance claim payments may not cover the full replacement cost of housing damaged by flood disasters and those policyholders may have to rely on individual assistance and loans after a disaster.

BW-12 directed FEMA to examine options to aid individuals so they could afford risk-based premiums under the NFIP utilizing targeted assistance for policyholders rather than generally

[1]As of May 31, 2017, there were approximately five million policies insuring approximately $1.2 trillion in assets.

[2]Public Law 112-141, Div. F, Title II, Subtitle A.

[3]Prior to BW-12, approximately 80 percent of policies were risk-based.

[4]The HFIAA surcharge is $25 for primary residences and $250 for second homes. Under HFIAA, annual premium increases are capped at 18 percent for primary residences and 25 percent for secondary homes. On average, the increase was around 9 percent (Aon National Flood Service, 2016).

[5]Public Law 114-89.

subsidized rates, including means-tested vouchers.[6] In addition, HFIAA required FEMA to develop an affordability framework to help policymakers consider the impact of implementing risk-based premiums and determine how to provide targeted policyholder assistance rather than discounted rates across the entire NFIP portfolio.

Section 9 of HFIAA required FEMA to examine options and consider the following criteria:

1. Accurate communication to consumers of the flood risk associated with their properties;
2. Targeted assistance to flood insurance policyholders based on their financial ability to continue their participation in the NFIP;
3. Individual or community actions that mitigate or lower the cost of flood insurance;
4. The impact of increases in risk premium rates upon participation in the NFIP;
5. The impact flood insurance rate map updates will have on the affordability of flood insurance.[7]

Objective

This study's objective is to respond to HFIAA's direction for FEMA to develop an affordability framework proposing programmatic and regulatory changes that address affordability of flood insurance. As such, FEMA primarily focused on parts (1), (2), and (3) of the HFIAA statutory considerations cited above. FEMA retains a significant body of work focusing on considerations (4) and (5) cited above, and we incorporated knowledge gained from this work throughout the framework.[8] In addition, for the purposes of this framework, FEMA did not consider the effect of future rate changes on affordability, as FEMA is generally reconsidering the rating structure of the NFIP under its Risk Rating Redesign effort. After implementing Risk Rating Redesign, the

[6]Section 100236 of Biggert-Waters Flood Insurance Reform Act of 2012, Pub. L. 112-141 (July 6, 2012) directed the Administrator to conduct a study of "(1) methods to encourage and maintain participation in the National Flood Insurance Program; (2) methods to educate consumers about the National Flood Insurance Program and the flood risk associated with their property; (3) methods for establishing an affordability framework for the [NFIP], including methods for individuals to afford risk-based premiums under the [NFIP] through targeted assistance rather than generally subsidized rates, including means-tested vouchers; and (4) the implications for the [NFIP] and the Federal budget of using each such method". See Id. at (a)(3)-(4). Under subsection (b), to inform the Administrator in the conduct of the study under subsection (a)'s study, Congress directed the Administrator "to enter into a contract under which the National Academy of Sciences, in consultation with the Comptroller of the United States, shall conduct and submit to the Administrator an economic analysis of the costs and benefits to the Federal Government of a flood insurance program with full risk-based premiums, combined with means-tested Federal assistance to aid individuals who cannot afford coverage, through an insurance voucher program. The analysis shall compare the costs of a program of risk-based rates and means-tested assistance to the current system of subsidized flood insurance rates and federally funded disaster relief for people without coverage." The Administrator was directed to report to the Committee on Banking, Housing, and Urban Affairs of the Senate and the Committee on Financial Services of the House of Representatives.

[7] Section 9 of the Homeowner Flood Insurance Affordability Act of 2014, Pub. L. 113-89 (Mar. 21, 2014) directed the Administrator to prepare a draft affordability framework and to submit the draft affordability framework to the full Committee on Banking, Housing, and Urban Affairs and the Committee on Appropriations of the Senate and the full Committee on Financial Services and the full Committee on Appropriations of the House of Representatives.

[8]Our data indicates that when prices of insurance increase, participation in the NFIP will decline, regardless of whether this price change is because of map updates or premium and fee increases.

NFIP will be able to determine risk-based premiums more accurately; some rates will increase while others may decrease. Rate increases could cause additional affordability challenges for policyholders who are already burdened by the cost of their flood insurance, as well as for potential policyholders. These challenges will be particularly salient for policyholders who currently receive discounts and subsidies.

Considerations in Developing the Affordability Framework

Affordability is a general concept used to address the concern that policyholders may not be able to afford their flood insurance premiums from rate increases—neither BW-12 nor HFIAA provided a definition of flood insurance affordability.[9] Thus, in the process of developing the affordability framework, FEMA solicited guidance from other federal agencies in late 2016 on how to define affordability in the flood insurance context. Based upon their feedback and our insights, we defined the concept of affordability from a cost burden or "ability to pay" perspective. Therefore, households applying for assistance face a means test to determine whether they qualify for benefits.

We included several other key considerations as we developed the framework:

- Flood insurance is the best way for a household to recover from a flood. Insured survivors recover more quickly and more fully than uninsured survivors, who often rely on federal disaster assistance and charity in order to recover;
- Targeting potential policyholders in addition to current policyholders for assistance could increase the number of property owners who want to purchase Federal flood insurance;
- Price is one of the best signals of risk that a consumer receives; any affordability assistance should be delivered with communication of the policyholder's full-risk, non-discounted rate;
- Any affordability program developed based on the framework that is funded by NFIP's current premiums and fees reduce the NFIP's ability to cover the cost of certain flood events, while creating additional affordability challenges, and work counter to our goal of creating a sound financial framework; and
- We discuss the definitions of affordability emerging from our work at the end of Chapter 2, and apply those definitions in developing the options in Chapter 3, and quantitatively illustrate the impacts of those options on affordability in Chapter 4.

[9]HFIAA suggests that premiums are unaffordable if the premium exceeds 1 percent of the policy coverage limit. However, the premium-to-coverage ratio has no means test associated with it. For example, a $100,000 property with $100,000 of coverage paying $1,000 for insurance would appear to be equally burdened as a $1,000,000 property with $250,000 of coverage paying $2,500 for insurance. The latter property owner may not face a cost burden when deciding whether to purchase $2,500 per year flood insurance policy.

Approach to Developing the Affordability Framework

In responding to HFIAA to develop an affordability framework, we took a multi-prong approach involving both qualitative and quantitative analyses. In essence, FEMA completed three tasks:

1. **To better understand affordability of the current NFIP portfolio as a baseline to understand the impact of changes going forward.** There has been very little nationwide analysis of flood insurance affordability because policyholder data on incomes and incomes of households in high-risk flood zones were largely unavailable. To fill this gap, FEMA developed an agreement with the U.S. Census Bureau (Census) to use data from the American Community Survey (ACS) to assess the incomes and housing expenses of NFIP policyholders and non-policyholders.[10] Specifically, FEMA conducted a series of analyses at the Census including the following: (1) FEMA analyzed how ACS respondents intersect with the Special Flood Hazard Area (SFHA) using the National Flood Hazard Layer (NFHL) to determine whether there were differences in incomes between those who live in the SFHA and those who live out of the SFHA; (2) FEMA and Census matched the NFIP policyholder data with ACS respondent data; and (3) FEMA conducted a number of exploratory analyses to build an understanding of the differences between policyholders and potential policyholders, differentiating by flood risk, income, and mortgage status.

2. **To begin the process of building the affordability framework.** This first part of that process was qualitative and involved engaging a broader policy community including academia, and other government agencies during two workshops to develop a series of flood insurance affordability program design options including administrative and funding options for the design options. The second part of the process was to qualitatively assess the administrative and funding options.

3. **To complete the process of building an affordability framework.** To complete this task, FEMA used Census data from Task 1 and the different affordability options that emerged from Task 2 to illustrate the scope and cost of the affordability options. The goal included showing *how* policymakers could use the framework and provide cost estimates of different program options. We selected parameters to show the range of possibilities to complete the examples. The examples we produce are merely illustrative. To use the framework to estimate actual costs and impacts on affordability, policymakers need to select specific parameters to model.

[10]The U.S. Census Bureau (Census) is part of the U.S. Department of Commerce (Commerce) and is overseen by the Economics and Statistics Administration (ESA) within Commerce. The Economics and Statistics Administration provides high-quality economic analysis and fosters the missions of the Census and the Bureau of Economic Analysis. FEMA's use of Census data is subject to this agreement, the Freedom of Information Act (FOIA), 5 U.S.C. § 552, as amended, the Privacy Act of 1974, 5 U.S.C. § 552a and any other applicable laws or regulations. The ACS is a household survey developed by Census to replace the long form of the decennial census program. The ACS is a large demographic survey collected throughout the year using mailed questionnaires, telephone interviews, and visits from Census field representatives to about 3.5 million household addresses annually. Starting in 2005, the ACS produced social, housing, and economic characteristic data for demographic groups in areas with populations of 65,000 or more. (Prior to 2005, the estimates were produced for areas with 250,000 or more population.) The ACS also accumulates sample over 5-year intervals to produce estimates for smaller geographic areas, including census tracts and block groups.

We include more discussion of the qualitative and quantitative approaches in the chapters that follow.

Framework Organization

We have organized the framework around the three tasks discussed above:

- Chapter 2 provides the results from Task 1 of using the quantitative data we developed to provide context and a baseline of affordability across the NFIP portfolio.
- Chapter 3 provides the results of the qualitative process of using workshops to develop a series of flood insurance affordability program design options that FEMA deemed feasible (first part of Task 2).
- Chapter 4 provides the results of using the quantitative data to illustrate examples of the scope and cost of the affordability options described in Chapter 3 (Task 3).
- Chapter 5 provides the workshop results for the administrative and funding options for the affordability designs (second part of Task 2).
- Finally, Chapter 6 provides conclusions based on our findings.

The framework also includes four appendices:

- Appendix A: Data and Statistical Methods;
- Appendix B: Additional Information on Flood Insurance Cost and Household Income;
- Appendix C: Methods Used to Develop Costs of Program Design Options;
- Appendix D: The Person Identification Validation System (PVS): Applying the Center for Administrative Records Research and Applications' (CARRA) Record Linkage Software.

II. Background on the Cost and Affordability of Flood Insurance Policies

Previous discussions around flood insurance affordability have been largely anecdotal rather than driven by data analysis because FEMA does not collect data on policyholder incomes required to analyze this topic sufficiently. We initiated the agreement with Census that allowed us to begin analyzing the issues around flood insurance affordability concerns using data-driven approaches. Based on matched analysis of the Census and FEMA data, we found the following:

- Policyholders tend to have higher incomes than non-policyholders, especially in the highest risk areas. This suggests that policymakers should pay particular attention to the affordability of flood insurance for households that currently do not have flood insurance but face flood risk.
- About 26 percent of NFIP residential policyholder households inside Special Flood Hazard Areas (SFHAs) are low income and 51 percent of non-policyholder households in SFHAs are low income, as defined by the United States Department of Housing and Urban Development, (HUD).
- Flood insurance premiums tend to be lower and household incomes tend to be higher outside SFHAs, so affordability is less of a problem outside SFHAs. However, there remain a significant number of households outside SFHAs for which affordability is an issue.
- About 49 percent of policyholders in SFHAs who own their homes spend less than 1 percent of household income on flood insurance. Although this statistic is a useful reference point, there currently is no rational basis to determine when the purchase of flood insurance becomes burdensome based on the percentage of income spent on flood insurance.
- The ratio of mortgage principal and interest payments, property taxes, and insurance (including flood insurance), or Principal, Interest, Taxes, Insurance (PITI), to household income exceeds 0.4 for approximately 12 percent of homeowners with flood insurance policies in SFHAs. The lending industry typically considers housing to be unaffordable when the PITI ratio exceeds 0.4. The PITI ratio provides a basis for defining when flood insurance becomes unaffordable.
- Incomes of homeowners with mortgages are higher than incomes of homeowners without mortgages.
- The affordability of flood insurance represents a challenge for a greater number of households as FEMA moves closer to risk-based rates for currently discounted policies.

Please find the support for FEMA's findings in the remainder of this chapter.

Location of Policies

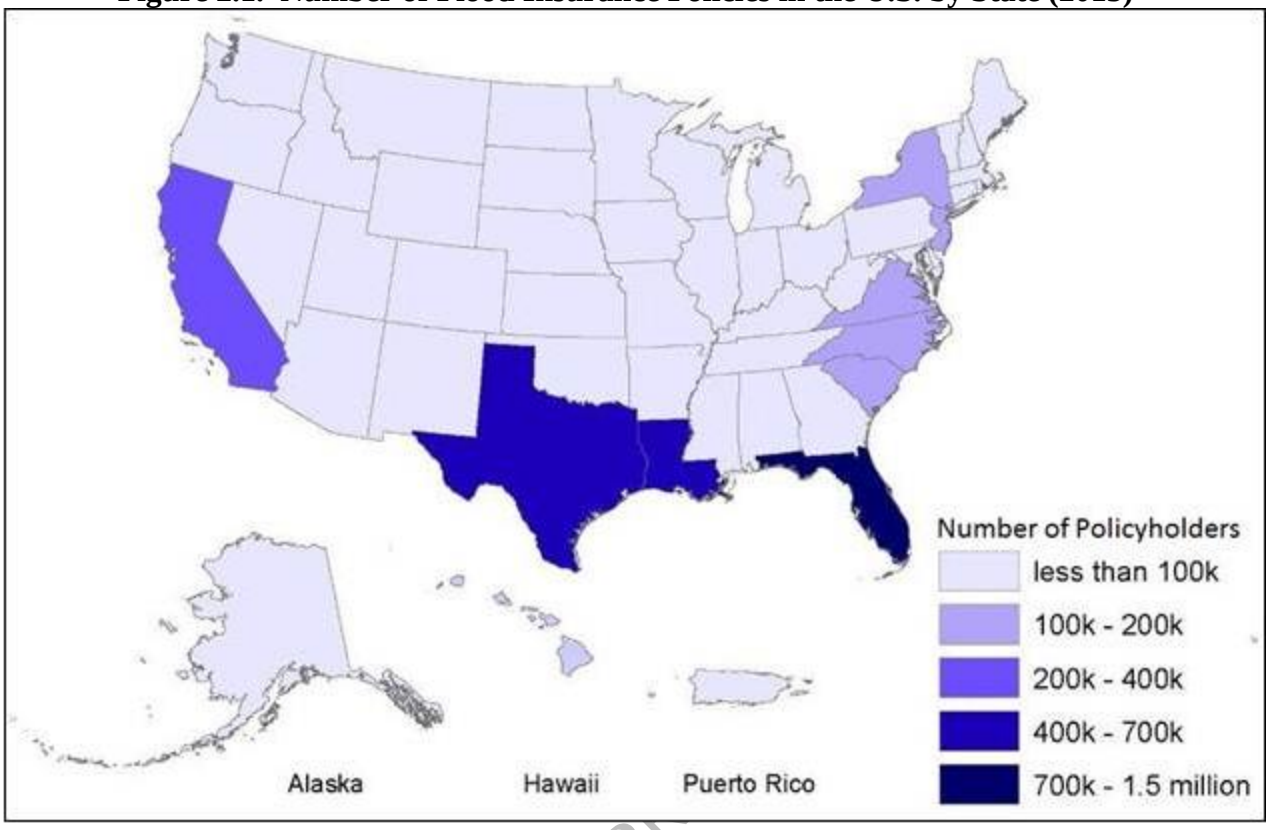
Figure 2.1 illustrates FEMA's analysis of 2015 NFIP data, highlighting the presence of policyholders in every state and emphasizing states with the highest counts of policyholders. FEMA included any policyholder with an active insurance contract in 2015 for this analysis;

yielding 4.8 million policies nationwide.[11] The majority of states had under 100,000 policyholders in 2015, while several Atlantic coast states had between 100,000 and 200,000 policyholders. California had around 300,000 policyholders, while Texas and Louisiana had a larger number of policyholders, ranging from 500,000 to 700,000. Florida had the largest number of policyholders at almost 1.5 million.

See Table B.1 in Appendix B for counts of NFIP policies by state in 2015.

[11]For this analysis, hereafter, NFIP "policies" actually refer to *contracts in force*. FEMA differentiates between contracts in force and policies in force for multi-unit structures. An insured structure counts as one contract in force, but if that structure has multiple units that are covered under one contract, each unit is counted as a policy. Therefore, a 100-unit condominium complex essentially counts as one contract but 100 policies. As explained in the appendices, FEMA does not keep a list of each policyholder in a multi-unit structure, only the name on the master policy for these structures. For FEMA and Census data matching purposes, FEMA used contracts rather than policies, but refers to them here as policies for simplicity.

Figure 2.1. Number of Flood Insurance Policies in the U.S. by State (2015)



SOURCE: FEMA analysis of NFIP policyholder data.

As shown in Table 2.1, the 4.5 million residential policies are nearly evenly divided between areas inside and outside the SFHA. However, the majority of nonresidential policies reside inside the SFHA. FEMA identifies those parts of the county comprising high-risk floods zones as Special Flood Hazards Areas (SFHA). The SFHA are those areas where there is a 1 percent annual chance of flooding. From the perspective of flood insurance affordability, SFHAs are relevant because flood insurance premiums are generally higher in SFHAs and flood insurance is mandatory for households with federally backed loans in these areas.

Table 2.1. Flood Insurance Policies by Extent and Source of Flood Risk

	Residential		Non-Residential	
Flood Zone	**Number (thousands)**	**Percent**	**Number (thousands)**	**Percent**
In SFHA	2,359	52%	227	69%
Outside SFHA	2,150	48%	104	31%
Total	**4,508**	**100%**	**330**	**100%**

SOURCE: FEMA analysis of NFIP policyholder data.

NOTE: In this table and throughout the framework, FEMA calculated column totals based on the raw data from each row, then rounded each individual row and the column total for ease of reporting. As a result, column totals may differ slightly from the sum of each reported row.

Policyholder Costs

Approximately 90 percent of the 4.5 million residential NFIP policies are for single-family homes. Table 2.2 summarizes the distribution of policy costs for these 4.1 million properties. Costs include both premiums and fees, with fees on average accounting for about 15 percent of total policyholder costs.[12] As can be seen in Table 2.2, policyholder costs for single-family homes average $1,098 inside SFHAs and $492, or less than half, outside the SFHA. The amounts policyholders paid varied considerably, with one-quarter of households in SFHAs paying more than $1,376 (the 75th percentile) and one-quarter paying less than $496 (the 25th percentile). These costs reflect the amount of coverage and the deductibles selected by policyholders. They also reflect any premium reductions because of grandfathering, pre-flood insurance rate map (pre-FIRM) subsidies, and Community Rating System (CRS) discounts.[13] The median and average can both represent the typical cost for a policyholder; however, extreme values do not affect the median as much as the average. The median, used throughout this framework, represents the middle value of the distribution of costs—roughly half of policyholders pay more than the median and half pay less than the median. Median policy costs are $738 and $439, respectively, inside and outside the SFHAs.

Table B.2 in Appendix B contains more information on the breakdown of policyholder costs into premiums and fees.

[12]HFIAA was passed in March 2014 and FEMA did not begin collecting the HFIAA surcharge until April 2015. As a result, some policies used in this analysis pre-dated FEMA collecting the HFIAA surcharge. For the purposes of this analysis, to ensure that policy costs more accurately reflect the NFIP's current fee structure, we added the HFIAA surcharge ($25 for primary residences and $250 for non-primary residences) to those policies missing a HFIAA surcharge.

[13]Grandfathering is a discount that allows properties constructed prior to being identified and mapped into a higher flood risk zone to keep their previous rates. Pre-FIRM subsidies are a discount provided to properties that were constructed or had substantial improvement on or before December 31, 1974 or before the effective date of an initial flood insurance rate map (FIRM). The CRS is a program developed by FEMA to provide incentives for those communities in the program that have gone beyond the minimum floodplain management requirements to develop extra measures to provide protection from flooding. CRS communities are eligible for certain flood insurance rate discounts.

Table 2.2. Policyholder Costs for Single-Family Homes (for policies in effect in 2015)

	In SFHA	Outside SFHA	Total
5th Percentile	$329	$296	$308
25th percentile	$496	$415	$437
Median	$738	$439	$485
Average	$1,098	$492	$800
75th percentile	$1,376	$485	$822
95th percentile	$2,922	$738	$2,328
Number of policies	**2,062,274**	**2,000,729**	**4,063,003**

SOURCE: FEMA analysis of NFIP policyholder data.
NOTE: Includes premiums and fees and all single-family homes whether owner-occupied or not.

Income of Policyholders and Non-policyholders

In this section, we provide an overview of the incomes of policyholders and non-policyholders inside and outside the SFHAs across the country. Policyholder income is an essential input into characterizing flood insurance affordability as we have defined it.

Methods

FEMA worked with Census to determine the income for a sample of NFIP policyholders and non-policyholders inside and outside the SFHA. We matched information from the ACS on 1.9 million households to NFIP policy data using the identity of the policyholder (which includes name and Social Security Number) and location. Overall, we found matches for approximately 65,000 of the 4.5 million residential policyholders. We used the ACS sampling rates to extrapolate findings for the matches back to the overall population of NFIP policyholders. We based our analysis in this chapter and Chapter 4 on the resulting 3.7 million NFIP policyholder households and 104.4 million non-policyholder households.

The technique we used to match NFIP policyholder and Census data yielded around 3.7 million policyholders for the analysis, which is less than the 4.5 million residential policyholders shown in Table 2.1. There are a number of reasons for this difference. For example, NFIP-insured, renter-occupied properties would not likely be represented in the data; if the landlord is a business, it would not be included in our sample because businesses are not included in the ACS. Also, if the landlord is an individual, the property would not be included because the landlord (who might be in the ACS) is not the same as the occupant. Similarly, Residential Condominium Building Association Policies (RCBAP) would not be included because the policyholder is typically a condominium association, which is not a part of the ACS.[14] Conversely, contents-only policies purchased by renters and policies purchased by individual condominium unit owners would be represented. Table A.1 in Appendix A summarizes how the weighted sample of NFIP

[14]An RCBAP can cover the structure of an entire residential condominium building (and all the individual units in it).

policyholders we used in this analysis compares with the full set of NFIP residential policies. In addition to having a lower number of single-family homes in the weighted sample than in the NFIP policy database, the percentage of single-family homes is higher in the sample than for the full set of NFIP policies. Please see Appendix A for additional information.

We believe this properly interpreted data are sufficient to analyze options for an NFIP affordability program, illustrated in Chapter 4.

Household Income Inside and Outside High-Risk Areas

Generally, incomes are higher outside the SFHA than they are inside the SFHA, as shown in the last row of Table 2.3. The combination of higher premiums and lower incomes in the SFHA creates affordability pressure on households. There is some variation across states in the relation between income inside and outside the SFHA, and as shown in Figure B.1 in Appendix B, median income is higher inside the SFHA than outside of it for a handful of states.
The results become more clear and stark when we separate policyholders and non-policyholders in the sample. As shown in the first row of the Table 2.3, the median household income for residential policyholders is $82,000, although it is substantially lower in the SFHA than outside the SFHA. Similarly, median income for households that do not have flood insurance is lower in the SFHA than outside the SFHA.

Table 2.3. Median Household Income of Policyholders and Non-policyholders (number of households in parentheses)

	In SFHA	Outside SFHA	Total
Policyholders	$77,000 (1.8 M)	$88,000 (1.9 M)	**$82,000 (3.7 M)**
Non-policyholders	$40,000 (3.3 M)	$56,000 (101.1 M)	**$55,000 (104.4 M)**
All Households	$50,000 (5.1 M)	$57,000 (103.0 M)	**$56,000 (108.1 M)**

SOURCE: FEMA analysis of NFIP policyholder data and Census ACS data.
NOTE: Data weighted using ACS sample weights. Median income rounded to nearest $1,000; number of households rounded to nearest 100,000; M = millions

As can be seen by comparing the first two rows in Table 2.3, policyholders tend to have higher incomes than households that currently do not carry flood insurance. In particular, the median income of the 3.3 million non-policyholders in the SFHA ($40,000) is $37,000 less than the median income of the 1.8 million policyholders in the SFHA ($77,000). Complicating the comparison of policyholder and non-policyholder incomes are the difference in the percentages of renters in each group. However, as we will see when we discuss the results in Table 2.6 below, the incomes of policyholders remain higher than that of non-policyholders even when comparing only homeowners or only renters, respectively.

Median income is higher for policyholders than non-policyholders in all states and the magnitude of the difference varies considerably (Figures B.3 and B.4 in Appendix B). Findings on the relative incomes of policyholder and non-policyholders suggest that policymakers should pay

particular attention to the affordability of flood insurance for households that currently do not have flood insurance.

Percentage of Households That Are Low Income

Low-income households endure the greatest difficulty affording flood insurance, and we use income categories based on Area Median Income (AMI) to identify the number of low-income policyholders and non-policyholders (Table 2.4). The advantage of basing income categories on AMI, as opposed to the Federal Poverty Level (FPL), is the AMI accounts for incomes and the considerable cost of living variance across the country, while the FPL does not.[15] HUD defines low-income households as those with income less than 80 percent of AMI, with three categories within low income shown in bold Table 2.4.[16] Cutoffs for higher-income groups vary, and for illustration, we use values from New York City housing assistance programs.[17] We refer to households with income over 165 percent AMI as *higher income households.*

Table 2.4. Income Categories

Household Income Cutoff	Terminology
<= 30% of AMI	**Extremely low income**
>= 30% and < 50% of AMI	**Very low income**
>= 50% and < 80% of AMI	**Low Income**
>= 80% and < 120% of AMI	Moderate income
>= 120% and <= 165% of AMI	Middle income
> 165% AMI	Higher income

NOTE: Low income categories highlighted in bold; AMI = Area Median Income.

Summing the first three rows of Table 2.5 reveals that around:

- 26 percent of policyholders inside the SFHA are low income.
- 21 percent of policyholders outside the SFHA are low income.
- 51 percent of non-policyholders in the SFHA are low income.
- 41 percent of non-policyholders outside the SFHA are low income.

[15]For example, AMI for a four-person household in Tampa, Florida, is $41,000. In New York City the AMI for a four-person household is $71,000.

[16]See https://www.huduser.gov/portal/datasets/il.html

[17]The 120 percent of AMI is the maximum income threshold for New York City's Housing Development Fund Corporation cooperative program (see https://www1.nyc.gov/assets/hpd/downloads/pdf/Owners/hdfc-coop-regulatory-agreement-faq.pdf). The 165 percent of AMI is the maximum income threshold for a handful of other programs targeted at moderate-to middle-income households (see http://www.nychdc.com/pages/Income-Eligibility.html).

Table 2.5. Distribution of Income for Policyholders and Non-policyholders

	Policyholders		Non-policyholders		
	In SFHA	Outside SFHA	In SFHA	Outside SFHA	All Households
Extremely low income (<= 30% AMI)	6%	4%	16%	12%	12%
Very low income (31 to 50% AMI)	7%	6%	16%	12%	12%
Low income (50 to 80% AMI)	13%	11%	19%	17%	17%
Moderate income (81 to 120% AMI)	18%	16%	19%	19%	19%
Middle income (121 to 165% AMI)	17%	16%	12%	16%	15%
Higher income (> 165% AMI)	39%	47%	17%	24%	25%
Total	100% (1.8 M)	100% (1.9 M)	100% (3.3 M)	100% (101.1 M)	100 (108.1 M)

SOURCE: FEMA analysis of NFIP policyholder data and Census ACS data.
NOTE: Data weighted using ACS sample weights. Number of households in parentheses; M = millions

Household Income by Mortgage Status and Source of Flood Risk

One of the flood insurance affordability measures developed in this framework considers total housing costs, including mortgage and interest payments, relative to income (PITI). Using the PITI approach would result in policyholders or prospective policyholders with mortgages being more likely eligible for assistance programs based on this measure of affordability. As a result, the incomes for households with and without mortgages are relevant to evaluating different affordability programs.

As can be seen from the top two rows in Table 2.6, the median income of homeowners with mortgages is substantially higher than for homeowners without mortgages. The pattern holds whether inside or outside the SFHA and for both policyholders and non-policyholders. Policyholders have higher median incomes than non-policyholders, regardless of their homeownership status. In the SFHA, more non-policyholders own their homes outright than have mortgages; they also reflect significantly lower incomes than their policyholder counterparts. This finding supports our extensive anecdotal evidence that there is a significant population in the SFHA of lower-income families who have either inherited their homes or are retirees who are particularly sensitive to the financial burden of flood insurance.

Table 2.6. Income by Housing Tenure and Mortgage Status (number of households in parentheses)

	Policyholders		Non-policyholders	
	In SFHA	Outside SFHA	In SFHA	Outside SFHA
Homeowners				
Homeowner has mortgage	$85,000 (1.1 M)	$104,000 (1.0 M)	$66,000 (661,000)	$83,000 (41.5 M)
Homeowner does not have mortgage	$70,000 (388,000)	$74,000 (657,000)	$40,000 (1.0 M)	$49,000 (23.8 M)
Renters				
Renters who pay rent	$52,000 (253,000)	$61,000 (191,000)	$34,000 (1.5 M)	$36,000 (33.8 M)
Space occupied without rent	$36,000 (22,000)	$40,000 (20,000)	$25,000 (103,000)	$28,000 (1.9 M)
Total Households	**(1.76 M)**	**(1.89 M)**	**(3.26 M)**	**(101.0 M)**

SOURCE: FEMA analysis of NFIP policyholder data and Census ACS data.
NOTE: Data weighted using ACS sample weights; median income rounded to nearest $1,000; number of households rounded to nearest 100,000; M = millions

As can be seen in the table above, renters have different income and expense profiles than homeowners. Comparing the number of households in the first two rows and second two rows of Table 2.6 show that 13 percent of policyholders are renters and 36 percent of non-policyholders are renters.[18] As expected, the results show that the median household incomes of renters are considerably lower than that for homeowners. In addition, in keeping with our previous analyses and findings in Tables 2.3 and 2.6, the incomes of policyholders tend to be higher than the incomes of non-policyholders, even when controlling for housing tenure (homeowner versus renter).

To determine whether incomes are higher in areas subject to coastal flooding—an issue often considered by FEMA—we classified the source of flood risk facing households in SFHAs as either coastal or noncoastal using our matched NFIP and Census data.[19] Table 2.7 compares median household income of the two groups and shows median income is higher for policyholders and non-policyholders exposed to coastal risk for both homeowners and renters. However, the income differences by source of flood risk are not sizeable compared, for example, to the differences in income between mortgage holders, outright homeowners, and renters.

[18]For reference, 1.3 million of the 1.9 million responses in the ACS (raw data), or 68 percent, are homeowners. We used the calculation of (0.49M/3.63M) for policyholder renters and (37.3/104.3M) for non-policyholder renters.

[19]Mark Crowell, Kevin Coulton, Cheryl Johnson, Jonathan Westcott, Doug Bellomo, Scott Edelman, and Emily Hirsch (*2010*) An Estimate of the U.S. Population Living in 100-Year Coastal Flood Hazard Areas. *Journal of Coastal Research*: Volume 26, Issue 2: pp. 201 – 211.

Table 2.7. Weighted Median Income by Source of Flood Risk in the SFHA

	Policyholders	Non-policyholders
Homeowners		
Coastal	$85,000 (876,000)	$51,000 (671,000)
Riverine	$78,000 (623,000)	$48,000 (1.0 M)
Renters		
Coastal	$52,000 (156,000)	$36,000 (758,000)
Riverine	$48,000 (119,000)	$31,000 (869,000)

SOURCE: FEMA analysis of NFIP policyholder data and Census ACS data.
NOTE: Data weighted using ACS sample weights; number of households in parentheses; Median income rounded to nearest $1,000; number of households rounded to nearest 100,000.

Flood Insurance Affordability

To develop the affordability framework, FEMA solicited guidance from other federal agencies on how to define affordability in a flood insurance context (please see Chapter 3). Based on this feedback and our own insights, we considered three different concepts of affordability in developing our flood insurance affordability framework:

1. **Flood insurance is considered unaffordable based strictly on household income.** A program based on this definition of affordability provides a benefit if a household's income falls below a certain threshold. Several federal programs use income as a measure for means-tested social assistance in their applications.
2. **Flood insurance is considered unaffordable when the cost of insurance exceeds a specified percentage of household income.** For example, flood insurance might be considered unaffordable when the household needs to spend more than 1 percent of its income on flood insurance.
3. **Flood insurance is considered unaffordable if the housing burden (including flood insurance) is more than a specified percentage of income.** HUD uses the concept of housing burden based on income in its rental assistance programs. For homeowners, housing burden consists of mortgage principal and interest (PI), property taxes (T), and insurance (including flood insurance—I), or PITI. For renters the housing burden is defined as of the ratio of rent plus insurance (typically contents insurance) to household income. FEMA would consider flood insurance unaffordable if flood insurance causes the ratio of PITI to income to exceed 0.30 to 0.40—cutoffs that are taken from both HUD and private mortgage industry standards.
4. The following two tables provide an overview of how current NFIP policyholders rank according to the second and third measures:

The top half of Table 2.8 shows the percent of income that homeowners that purchase flood insurance spend on it. Around 49 percent of policyholders inside SFHA spend less than 1 percent of their annual income on flood insurance, including both premiums and fees. The remaining 51 percent of homeowners inside SFHA spend more than 1 percent of household income on flood insurance, with 24 percent spending between 1 and 2 percent of their incomes, and 7 percent spending more than 5 percent of their incomes. Because household income tends to be higher and flood insurance premiums tend to be lower outside the SFHA, homeowners outside the SFHA tend to spend a lower proportion of their income on flood insurance. As such, around 80 percent of policyholders outside the SFHA spend less than 1 percent of income on flood insurance.

Table 2.8. Flood Insurance Costs as Percentage of Weighted Household Income for Residential Policyholders

Flood Insurance Cost as Percentage of Income	In SFHA		Outside SFHA		Total	
	Number (in 000s)	Percent of Total	Number (in 000s)	Percent of Total	Number (in 000s)	Percent of Total
Policyholders Who Own Their Residence						
<= 1%	741	49%	1,322	80%	2,063	65%
>1% and <=2%	366	24%	222	13%	588	19%
>2% and <=3%	163	11%	56	3%	219	7%
>3% and <=4%	76	5%	22	1%	98	3%
>4% and <=5%	44	3%	11	1%	55	2%
>5%	109	7%	28	2%	138	4%
Total	**1,499**	**100%**	**1,663**	**100%**	**3,162**	**100%**
Policyholders Who Are Renters						
<= 1%	86	31%	120	57%	206	42%
>1% and <=2%	69	25%	49	23%	118	24%
>2% and <=3%	34	12%	18	9%	52	11%
>3% and <=4%	26	9%	9	4%	35	7%
>4% and <=5%	14	5%	3	2%	17	3%
>5%	46	17%	11	5%	58	12%
Total	**275**	**100%**	**211**	**100%**	**486**	**100%**

SOURCE: FEMA analysis of NFIP policyholder data and Census ACS data.
NOTE: Data weighted using ACS sample weights; policyholder costs includes premium and fees.

Renters with flood insurance tend to spend a higher percentage of income on flood insurance (bottom half of Table 2.8). Inside the SFHA, 31 percent of renters spend less than 1 percent of income on flood insurance while 17 percent of renters spend more than 5 percent of their incomes on it. Similarly, the percentage of income spent on flood insurance is lower outside of the SFHA. Renters typically purchase contents only insurance as they do not own the property in which they reside.

We cannot calculate flood insurance costs as a percentage of income for non-policyholders. However, as household income for non-policyholders is lower than for policyholders and if the

cost of a policy ends up being similar for non-policyholders and policyholders, then premium as a percentage of income would be higher for non-policyholders.

The numbers in Table 2.8 provide information on the burden households' face when purchasing flood insurance under a percentage of income measurement. For comparison, national data from the Bureau of Labor Statistics show that, on average, households in owner-occupied housing units across the country (the vast majority of which do not purchase flood insurance) spend 1.8 percent of their income on home insurance, maintenance, repair, and other housing expenses, excluding mortgage payment and property taxes.[20] While the percentage of household income spent on home insurance, maintenance, repair, and other housing expenses, excluding mortgage payments and property taxes, provides some point of reference, it is not particularly useful in determining when the purchase of flood insurance becomes burdensome. Currently, there is no reasonable basis for determining at what point (in terms of the percent of household income) flood insurance cost becomes burdensome.

Table 2.9 provides a measure of affordability based on housing-burden which is an alternative affordability measure that accounts for a household's total housing cost. The top area of the table shows the PITI ratio for homeowners with flood insurance. Around 12 percent of homeowners have a PITI ratio greater than 0.4—they are considered to be burdened and at a level above which few lenders would be willing to make loans to them.[21] These households consider the cost of flood insurance burdensome and difficult to afford. Household incomes are higher and flood insurance premiums are lower outside the SFHA, and consequently a lower percentage of homeowners outside SFHAs have a PITI ratio in excess of 0.4 (7 percent as opposed to 12 percent).

[20]U.S. Bureau of Labor Statistics, "Consumer Expenditures in 2009," new release, USDL-10-1390, October 2010.

[21]"Section F. Borrower Qualifying Ratios (4155.1)," hud.gov website, March 1, 2011b. As of March 20, 2017: https://portal.hud.gov/hudportal/documents/huddoc?id=4155-1_4_secF.pdf

Table 2.9. Ratio of PITI to Household Income for Residential Policyholders

Housing Burden (PITI Ratio)	In SFHA		Outside SFHA		Total	
	Number (000s)	Percent of Total	Number (000s)	Percent of Total	Number (000s)	Percent of Total
Policyholders Who Own Their Residence						
<= 0.3	1,213	81%	1,461	88%	2,673	85%
>0.3 and <=0.4	106	7%	77	5%	183	6%
>0.4 and <=0.7	104	7%	75	4%	179	6%
>0.7	76	5%	51	3%	127	4%
Total	**1,499**	**100%**	**1,663**	**100%**	3,162	100%
Policyholders Who Are Renters						
<= 0.3	133	53%	112	59%	245	55%
>0.3 and <=0.4	35	14%	26	14%	61	14%
>0.4 and <=0.7	52	20%	33	17%	85	19%
>0.7	33	13%	20	10%	53	12%
Total	**253**	**100%**	**191**	**100%**	444	100%

SOURCE: FEMA analysis of NFIP policyholder data and Census ACS data.
NOTE: Data weighted using ACS sample weights; table excludes ACS respondents who occupy a rental property without payment of rent.

PITI ratios for renters (defined as gross rent over income) are higher than those for homeowners with 33 percent of renters inside SFHAs and 27 percent of renters outside SFHAs maintaining a PITI ratio over 0.4. By this definition, flood insurance is unaffordable for a substantial percentage of renters.

Table 2.10 repeats the analysis for non-policyholders. As reported above, non-policyholders tend to have lower incomes than policyholders, but they presumably also tend to have lower PITI costs because they do not purchase flood insurance. Thus, it is not obvious how the PITI ratios for non-policyholders will compare to those of policyholders. Comparing Tables 2.9 and 2.10 reveals that no consistent relationship exists between the PITI ratios of the two groups. For example, 7 percent of homeowners in SFHAs without flood insurance have a PITI ratio over 0.4 compared to 12 percent of policyholders. In contrast, 34 percent of renters outside SFHAs without flood insurance have a PITI ratio over 0.4 compared to 27 percent of policyholders. Overall, there is no strong relationship between the PITI ratios of policyholders and non-policyholders.

Table 2.10. Ratio of PITI to Household Income for Residential Non-policyholders

Housing Burden (PITI Ratio)	In SFHA		Outside SFHA		Total	
	Number (000s)	Percent of Total	Number (000s)	Percent of Total	Number (000s)	Percent of Total
Non-policyholders Who Own Their Residence						
<= 0.3	1,498	88%	57,012	87%	58,511	87%
>0.3 and <=0.4	70	4%	3,384	5%	3,454	5%
>0.4 and <=0.7	69	4%	2,982	5%	3,051	5%
>0.7	59	3%	1,996	3%	2,055	3%
Total	1,696	100%	65,374	100%	67,071	100%
Non-policyholders Who Are Renters						
<= 0.3	719	47%	17,118	51%	17,118	51%
>0.3 and <=0.4	244	16%	5,197	15%	5,197	15%
>0.4 and <=0.7	308	20%	6,390	19%	6,390	19%
>0.7	254	17%	5,140	15%	5,140	15%
Total	1,524	100%	33,845	100%	33,845	100%

SOURCE: FEMA analysis of NFIP policyholder data and Census ACS data.
NOTE: Data weighted using ACS sample weights; table excludes ACS respondents who occupy a rental property without payment of rent.

While the PITI-based measure of affordability takes a more holistic view of household finances than the ratio of the premium to income, it has its own drawbacks. We detail those drawbacks in Chapter 3 when we discuss the advantages and disadvantages of a flood insurance affordability program based on the PITI-based measure of affordability.

III. Affordability Framework: Developing a Feasible Set of Flood Insurance Affordability Program Design Options

In order to develop a set of flood insurance affordability program design options, we conducted background research. As part of that research, FEMA obtained the assistance of the National Academies of Science, Engineering, and Medicine (NAS) for two workshops. The workshops convened by NAS included attendees from other federal agencies that administer assistance programs, academics who have studied flood insurance affordability, and other professionals with relevant expertise related to insurance. FEMA's goals for the workshops were to learn more about these other assistance programs and to obtain information to enhance FEMA's ability to develop a set of potential flood insurance affordability options. We discuss the flood insurance affordability design options that emerged from this effort and qualitatively assess them based on our own expertise and in light of expertise of the workshop attendees. The workshops also generated some ideas for administrative and funding options, which we discuss in Chapter 5.

Based on the background research and workshops, we identify four program design options:

1. **Income-based premium sharing:** Lower-income households would be responsible for paying for a portion of the premium amount and FEMA covers the remainder of the premium amount. In this option, as household income levels rise, the portion of the premium that would be covered by FEMA decreases.
2. **The premium burden-based benefit:** Lower-income households would be responsible for paying for a portion of their income for flood insurance. If the required proportion of income is not sufficient to cover the insurance premiums, FEMA would cover the remainder of the premium amount.
3. **The housing burden-based benefit**: Lower-income households that spend more than a specified amount of their income on housing-related expenses, such as mortgage amount, taxes and insurance would receive assistance.
4. **Mitigation grants or loans:** This approach would complement the other program designs. Under this design option, the government would provide financial assistance to fund structure-specific mitigation activities that lead to reduced risk. The assistance would be a grant for lower-income households and a loan for more moderate-income households.

The four design options vary in terms of their design characteristics, advantages and disadvantages, and risk communication implications, which are described below.

What We Did

Background Research on Developing Flood Insurance Affordability Design Options

FEMA began developing the program options proposed in this document by reviewing two congressionally mandated NAS reports—*Affordability of National Flood Insurance Program Premiums, Reports 1 and 2*.[22] The first report documented the history of the NFIP pricing practices, reviewed the literature on the demand for flood insurance, and identified questions for consideration when designing an assistance program. The second report identified criteria for evaluating potential affordability policy options and highlighted the absence of data necessary to determine what would be affordable for NFIP policyholders.

To further our understanding of the challenges associated with developing an affordability program, we reviewed an overview by the U.S. Government Accountability Office (GAO) of 80 public assistance programs released in 2015. Based on that report, we developed a spreadsheet to track characteristics across programs, such as the program goal, the eligibility criteria, the assistance provided, and the costs.[23] Based on that information, we classified the programs into different categories of assistance programs and then selected representative programs within those categories that were most applicable to a potential flood insurance affordability program. We focused on three characteristics of those programs: type of assistance provided, how it is provided, and who receives assistance. In addition, we asked NAS to convene two workshops for this study, with the goals of learning more about these assistance programs and soliciting individual perspectives useful for developing a set of potential flood insurance affordability options.

These workshops focused on the following questions:

1. Who will receive a benefit?
2. What is the level of benefit for different household types?
3. By who and how will the program be administered?
4. How will the program be funded?

First Workshop

We identified six federal agencies that implemented ten different benefit assistance programs and invited them to a workshop to discuss how those programs operated. The federal agencies that attended the workshop included the Department of Energy (DOE), Department Health and Human Services (HHS), Department of Agriculture (USDA), Department of Housing and Urban

[22]*Affordability of National Flood Insurance Program Premiums – Report 1*, 2015. Committee on the Affordability of National Flood Insurance Premiums, Water Science and Technology Board, the National Research Council of the National Academies, Washington D.C. *Affordability of National Flood Insurance Program Premiums – Report 2*, 2016. Committee on the Affordability of National Flood Insurance Premiums, Water Science and Technology Board, the National Research Council of the National Academies, Washington D.C.

[23]Federal Low-Income Programs: Multiple Programs Target Diverse Populations and Needs. July 2015. U.S. Government Accountability Office, GAO-15-516.

Development (HUD), the Federal Housing Finance Agency (FHFA), and the Small Business Administration (SBA). Representatives from the Office of Management and Budget (OMB) and the GAO also attended. In addition, several academics who study the issues of flood insurance affordability participated.

The programs discussed in the first workshop (along with their relevant agencies) were as follows:

1. Home Weatherization (DOE);
2. Disaster Supplemental Nutrition Assistance Program (USDA);
3. Supplemental Nutrition Assistance Program (USDA);
4. Small Business Disaster Loan Program (SBA);
5. Head Start (HHS);
6. Health Insurance Exchange (HHS);
7. Medicare Nursing Home Program (HHS);
8. Housing Choice Vouchers (HUD);
9. Public Housing Program (HUD); and
10. Low Income Home Energy Assistance Program (HHS).

The workshop provided insight into a wide variety of assistance programs within the Federal Government, and into how each defines affordability and how the responsible agency administers its program. We learned from the first workshop that most of the program characteristics, such as the definition of affordability and how the program is administered, arose from the program's initiating legislation. As a result, agency views on the concept of affordability differ based on whether their programs grew out of a congressional mandate or through interpreting rules that govern these programs.

For many programs, there are specific proportions of income that individuals have to contribute for the good or service being provided before receiving the benefit. For example, HUD's Public Housing Program is based on a formula that requires that the household to spend a certain percentage of its income on housing with the Federal Government subsidizing the remainder of the payment.[24] This method defines affordability through an acceptable level of cost burden placed on a household. By contrast, other programs define affordability through income levels. For example, the Supplemental Nutrition Assistance Program (SNAP) program states that given an income level, a specific benefit is defined for the household that may only be expended on food. We believe this concept is similar to the idea of cost burden, because policymakers believe there is a specific amount of money that is reasonable for a household to spend on certain basic living expenses.

The information we learned from this initial workshop helped us to frame our options for an affordability framework that is grounded in structure and experience of existing federal

[24]The formula used to determine rent for a public housing resident is the highest of the following, rounded to the nearest dollar: (1) 30 percent of the monthly adjusted income. (Monthly Adjusted Income is annual income less deductions allowed by the regulations); (2) 10 percent of monthly income; (3) welfare rent, if applicable; or (4) a $25 minimum rent or higher amount (up to $50) set by a Housing Authority.

programs. While those programs are generally aimed at providing affordable goods or services, which is inherently different to flood insurance, the catalog of different programs and critical program information from the agencies administering them provided us with insights and analysis to develop alternative program designs for (1) who would be eligible, (2) how benefits could be calculated, and (3) how the program could be administered.

Second Workshop

Based on a review of existing assistance programs and the input received at the first workshop, FEMA developed a set of six potential program design options, as well as four potential methods of administering a flood insurance affordability program. For the second workshop, the NAS reconvened most of the federal agencies that attended the first workshop, several academics, and two loaned executives from insurers that participate in the Write Your Own (WYO) program to participate in a second workshop. The federal agencies in attendance included the DOE, HHS, USDA, HUD, FHFA, and SBA. Additionally, an individual associated with the District of Columbia's Health Insurance Exchange (DC Health Link) participated. Representatives from OMB and the GAO also attended the workshop.

FEMA presented a set of potential design and administration options to the full group at the second workshop. We divided the attendees into three groups to examine, discuss, and provide concrete feedback on both the design and administration options for a potential affordability program. FEMA's goal was to obtain opinions on specific program design and administration options, including a discussion of the pros and cons of each option, from knowledgeable parties having experience designing or administering social assistance programs or those familiar with NFIP's mechanics. For these meetings, participants were asked not to consider any issues related to the cost of the program.

Feedback from the second workshop was valuable, and, allowed FEMA to narrow the set of options to those that address the issue of affordability and that we could implement most efficiently and with the least amount of complexity. (See Chapter 5 for discussion of the administration options.)

Design Options for a Flood Insurance Affordability Program

Based on the second affordability workshop, FEMA developed a set of design options for an affordability program:

1. Income-based premium sharing
2. Premium burden-based benefit
3. Housing burden-based benefit
4. Mitigation grants and loans

Program Design 1: Income-Based Premium Sharing

The income-based premium sharing design is based on cost sharing between lower-income households and the Federal Government to reduce a household's policy cost.[25] The Federal Government pays a part of the policy cost that represents the benefit, and the household pays a portion, depending on its income. To prevent marked changes in benefits at certain income thresholds, we assume that the benefit will decrease gradually as income increases.

Design Characteristics

Under the income-based premium option, benefits to households decrease as household income rises. This approach requires establishing the relationship between income and benefits. As previously discussed, two common income thresholds are used to determine eligibility for program benefits: (1) FPL, and (2) AMI. While the FPL is constant across the entire country, the AMI is specific to a location—most commonly at the county, municipality, or metropolitan area level—because it incorporates the cost of living in that location. Because FEMA is developing a nationwide framework and incomes and cost of living vary by states and regions, we concluded that AMI is the most appropriate measure for our purposes. We used the income categories listed in Table 2.4 of Chapter 2.

This design option has three main parameters:

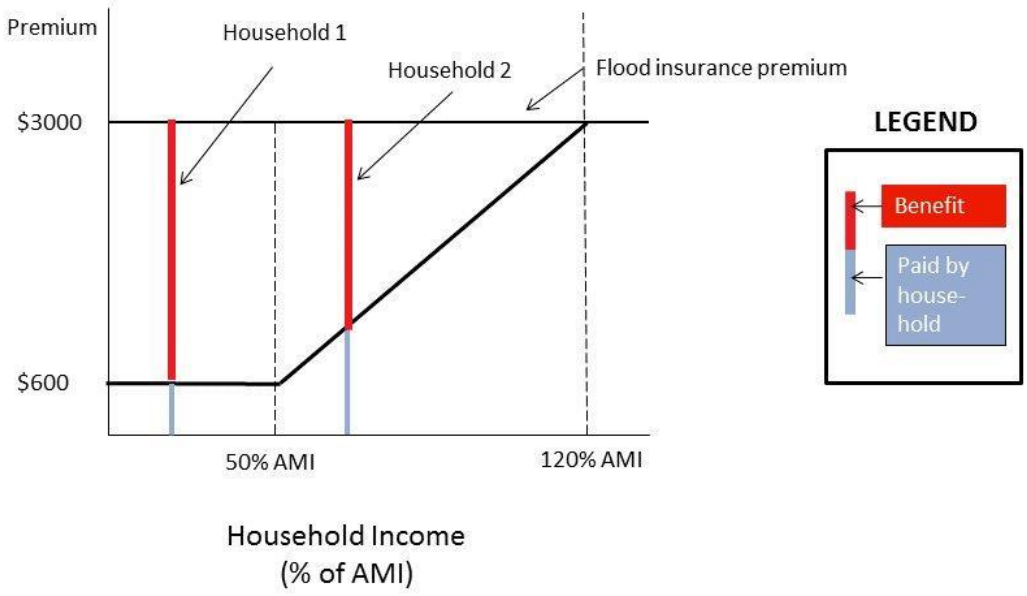
1. What level of benefit is provided to the lowest-income households?
2. What is the income cutoff, measured with respect to AMI, for households receiving that level of benefit?
3. What is the income cutoff for receiving any benefit from the Federal Government?

Figure 3.1 provides an illustrative example of how the program could operate. In this example, the benefit for the lowest income household is 80 percent of the premium, meaning even the lowest-income households would still be responsible for paying at least 20 percent of their policy cost. Households earning less than 50 percent of AMI would receive 80 percent of the premium and the portion of the premium covered would gradually fall to zero as income approaches 120 percent of AMI.

Assuming a premium of $3,000, Figure 3.1 shows what the benefit would be for each income level. The blue bars show the amounts paid by two different households, and the red bars show what the benefits would be for the same two different households. Household 1 has an income less than 50 percent of AMI; thus, FEMA covers 80 percent of its premium. It would thus pay $600 of the $3,000 premium, with the program picking up the rest ($2,400). Household 2 has income between 50 percent and 120 percent of AMI and receives a benefit, but the benefit would be less than 80 percent of the premium.

[25]This option is motivated by HHS's Low Income Home Energy Assistance Program (LIHEAP). As administered by some states, the program provides beneficiaries with payments that partially offset their home energy costs; the amount received varies by the household's income category.

Figure 3.1. Illustrative Example of Income-Based Premium Sharing



Source: Adapted from Dixon, Clancy, Miller, et al., The Cost and Affordability of Flood Insurance: Economic Impacts of Rising Premiums and Policy Options for One- to Four-Family Homes, RAND, RR-1776, 2017, p. 91.

As with the other options described below, the number of households eligible and the overall cost of the program will depend on the design parameters. This approach could offer benefits to moderate-income households (those with incomes between 80 and 120 percent of AMI); although benefits decline as incomes rise. Reducing the income cutoff for receiving any government assistance (120 percent of AMI, in this case) will reduce the number of households eligible for the program. Reducing the maximum percent of the flood insurance premium covered (80 percent of the premium, in this case) or the income cutoff for receiving the maximum benefit (50 percent of AMI, in this case) will reduce the cost of the program given the number of households eligible.

The program cost will also depend on a number of program features, specifically on whether:

- Only the current owner of the home is eligible or whether future buyers of the property are also eligible to receive the benefit;
- The program should sunset after a certain number of years;
- Only homes in high-risk flood areas (where flood insurance is mandatory for homeowners with federally regulated mortgages) are eligible, or whether all homes in the United States are eligible; and
- The program is for primary homeowners or also for renters, those with second homes, and businesses.

These decisions will have important impacts on the number of households assisted, the extent of the assistance provided, and the overall cost of the program. Ultimately, policymakers will make the tradeoff decisions between program cost and the impact of the program on flood insurance affordability.

Advantages and Disadvantages

In the second NAS workshop, participants noted that one of the primary advantages of the income-based premium sharing approach is that it would be relatively straightforward to implement. The information that FEMA would collect from the household and use (income level and flood insurance premium) is obtainable and relatively straightforward to use. Additionally, because benefits vary continuously across incomes, households do not risk experiencing a situation where earning an additional dollar results in them receiving drastically fewer, or no benefits. By using AMI, the amount of benefits received adjusts to local costs of living. The FPL does not adjust to the local cost of living.

The primary disadvantages to this approach are that there is no existing legislative guidance or precedent for how to select the parameters of the benefit structure. Second, while this option targets households with low to moderate levels of income, some households for whom flood insurance is not difficult to afford may receive benefits. For example, a low-income household that spends only a very small proportion of its income on flood insurance would still receive benefits in this program design. Another disadvantage of basing eligibility on income alone is that households with low income but high net worth could receive assistance. FEMA could administer an asset test to prevent such an occurrence, but the administrative burden of conducting such a test on every household applying for assistance could be high. Previous work suggests that the percent of households with low incomes and high net worth is low, so a simpler solution is to require that households that receive the benefit certify that their net worth is not above a specified threshold, informing applicants that forms are subject to audit.[26] The program could then audit a small number of the eligibility applications each year to deter false statements.

Risk Communication

FEMA would notify NFIP participants of their full-risk rate before providing the program benefit to signal their risk is greater than their premium price reflects. Even though households pay a part of their premium if they receive a benefit, they do not pay for the full risk of living in a flood-prone area and may not expend resources to avoid or mitigate risk in high-risk areas. However, when households receive benefits as a percentage of the premium and premiums rise with risk, they would have some incentives to avoid riskier or costlier areas.

Program Design 2: Premium Burden-Based Benefit

Requiring households to pay a certain percentage of their incomes toward flood insurance is a second approach to developing an income-based option. If the required proportion of income is not sufficient to cover flood insurance premiums, the Federal Government would pay for the remainder of the cost.

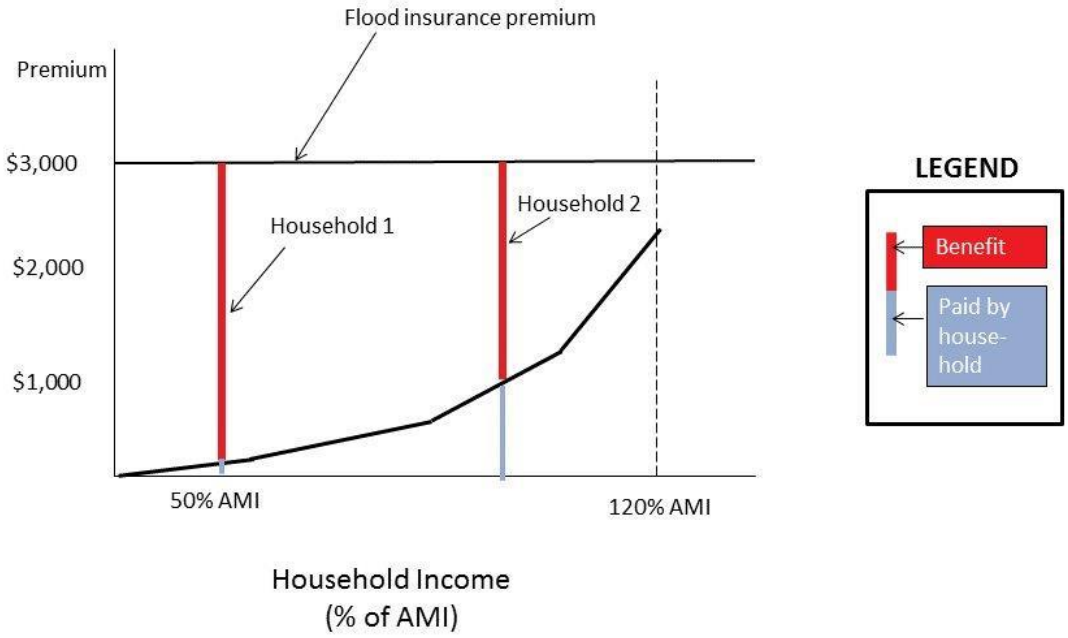
[26]For example, previous work has shown that 4.2 percent of households with incomes less than $24,000 have net worth (including the equity in their home) of $500,000 or more (Dixon, Clancy, Miller et al., *The Cost and Affordability of Flood Insurance: Economic Impacts of Rising Premiums and Policy Options for One- to Four-Family Homes*, RAND, RR-1776, 2017, p. 149).

Design Characteristics

In this design option, households would be required to spend a proportion of their income on flood insurance before collecting any benefit. As in Design 1, percentages of AMI are cutoff points for the portion of income households are expected to spend on flood insurance. For example, a household with income less than 50 percent of AMI might be required to spend up to 0.5 percent of their income on its flood insurance premiums (excluding fees), while a household with income between 50 and 80 percent of AMI might be required to spend 1.5 percent of their income. The Federal Government would pay the remainder of the premium cost.

Figure 3.2 provides a visualization of the potential design option. The blue lines show the amount that two hypothetical households would be required to pay toward the flood insurance premium, while the red lines show the program benefit received. Household 1 receives a larger benefit than Household 2, because Household 1 has a lower income.

Figure 3.2. Illustrative Example of Premium Burden-Based Benefit



Source: FEMA Analysis

Advantages and Disadvantages

As with Design 1, there are several advantages to consider. First, the information to be collected from the household and used by program administrators is relatively straightforward to obtain. Second, from the policyholder perspective, there is a cap on the maximum amount a household must pay on flood insurance, meaning premium increases do not create new affordability

concerns for the beneficiaries.[27] Third, the government can base required household contribution on a reasonable expectation of how much a household should pay toward flood insurance.

One of the disadvantages of this approach is that households no longer share in any potential premium increases because of increased flood risk at the property; as such, they do not have an incentive to avoid additional flood risk. Second, there are few precedents for policymakers to follow in setting a reasonable percentage of income that policyholders should spend on flood insurance, and third, potentially, households with low income but high net worth could receive assistance because income, and not household wealth is considered. The same solution to this drawback could apply here, which is to have households that receive the benefit certify their net worth is not above a specified threshold, informing applicants their forms are subject to audit.

A final disadvantage of this approach is that it uses an affordability metric that is based on just one component of overall housing cost; thus, this option could result in providing a benefit to a household with high flood insurance cost but a low mortgage payment (and who thus could afford flood insurance according to a PITI-based measure of affordability). For example, a low-income household may be able to afford a property with a very high flood insurance premium and an associated low property value. A low-income household may be able to afford the property if the Federal Government was largely subsidizing the premium. In that instance, the Federal Government would be facilitating the purchase of high-risk properties by low-income households.[28]

Risk Communication

FEMA would notify program participants of their full-risk rate before providing the program benefit to signal that their risk is greater than their premium price reflects. However, the pricing for this option does not communicate risks to the household, because all households with a given income will pay the same, regardless of the risk they face and regardless of whether their risk increases. Thus, this provides no incentive for the household to avoid or mitigate flood risk in high-risk areas.

Designs 1 and 2 with Income Bins Rather Than Continuous Subsidy

For certain income ranges in both Designs 1 and 2, the subsidy declines gradually as income increases, but it is possible to recast the subsidy in terms of income "bins" for both options. That is, all households within certain AMI categories would receive the same benefit. For example, with Option 1, the program could be designed such that for an AMI less than 50 percent, households would receive 80 percent of the premium as a benefit. Between 50 percent and 100 percent of AMI, they would receive 60 percent of the premium. Finally, between 100 percent and 120 percent, they would receive 40 percent of the premium. Several federal workshop participants stated that this binned approach may be more streamlined from an administrative

[27] A higher flood insurance premium shifts up the horizontal flood insurance premium line above $3000 in Figure 3.2, but the curve indicating the required amount of household contribution does not change.

[28] See Dixon, Clancy, Miller, et al., RAND, RR-1776, 2017, p. 78-82 for a discussion of the relationship between flood insurance cost and property value.

perspective, but it comes at a potential cost. Strict cutoff points penalize households that are just above the cutoff points—in particular, households might see a considerable drop in benefits if the household slightly increases its income.[29] This cut off could potentially discourage low-income households from seeking additional income.

Program Design 3: Housing Burden-Based Benefit

In this design option, benefits are targeted at households that are both income and housing-burdened—that is, the percent of income spent on homeownership exceeds a specified threshold.[30] As in previous options, benefits are limited to households with incomes below a fixed AMI threshold, but only available to households that are spending more than a fixed percentage of their income on housing. Benefits are no larger than the household's flood insurance premium. This design benefits households that spend a larger portion of their income on housing; referred to as PITI (principal, interest, taxes, and insurance) and is a standard lending industry metric.

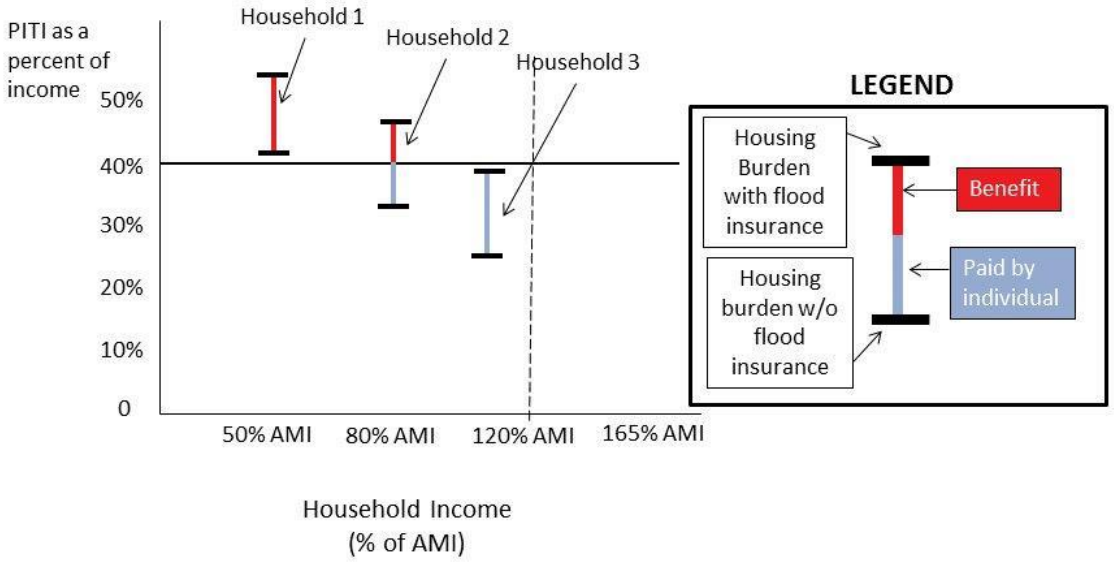
Design Characteristics

The benefit would cover that part of the flood insurance premium that, when added to PITI, would cause the percentage of income spent on housing to rise further above the specified threshold. Two main parameters determine eligibility in this design: first, the benefit is only available for households below a fixed AMI threshold. Figure 3.3 sets the income eligibility threshold at 120 percent of AMI, as in the previous examples. Second, the benefit is only available to households that spend more than a fixed percentage of their income on housing. Based on practices in the lending industry, our example in Figure 3.3 assists households with housing burden above 40 percent of their income—these households are housing-burdened and, would likely have difficulty affording flood insurance.

[29]Saez, E. (2010). Do taxpayers bunch at kink points? *American Economic Journal: Economic Policy*, *2*(3), 180-212.

[30]This is a similar model to HUD's Section 8 rental housing assistance program, where participants are expected to pay 30 percent of their monthly income toward their housing costs. The HUD subsidy covers the remaining amount up to a capped market rental amount.

Figure 3.3. Illustrative Example of Housing Burden-Based Premium Benefit



Source: Adapted from Dixon, Clancy, Miller, et al., The Cost and Affordability of Flood Insurance: Economic Impacts of Rising Premiums and Policy Options for One- to Four-Family Homes, RAND, RR-1776, 2017, p. 98.

In Figure 3.3 above, Household 1 spends more than 40 percent of its income on housing costs, with or without flood insurance costs and therefore would receive a benefit for its entire flood insurance premium. In contrast, Household 3's PITI-to-income ratio is less than 40 percent, with or without flood insurance, and as such, it would receive no benefit. Household 2, spends more than 40 percent of its income on housing when flood insurance is included, and therefore would receive a benefit for that portion of the flood insurance premium that causes its housing burden to increase above 40 percent of income—which is the difference between the PITI with flood insurance and 40 percent of income. For example, suppose that Household 2 has an income of $62,000 and currently spends $23,500 on PITI excluding flood insurance (38 percent of income). The household faces a flood insurance premium of $3,000, moving its housing costs to $26,500 and its PITI ratio to about 43 percent of income. A 40 percent PITI-to-income ratio would be $24,800; consequently, Household 2 would be responsible for $1,300 of its flood insurance premium ($24,800–$23,500) and would receive a benefit of $1,700.

Advantages and Disadvantages

A principal advantage of this approach is that it bases program parameters (particularly the PITI ratio cutoff) on best practices in the lending industry. Lenders typically will not make loans that cause housing burden to exceed 30 to 40 percent, and this cutoff can be used as the definition of affordability underlying the program. Previous work has shown the PITI ratio is highly

correlated with household income, so a program basing eligibility on the PITI ratio will target lower-income households.[31]

Despite these advantages, the PITI approach also has several disadvantages. First, because this is a more complex design, FEMA would have to gather more information about the recipient to implement this option. In addition to the household's income, this design requires FEMA to gather information on household's mortgage payments, property taxes, and insurance payments. Although this approach would require FEMA to obtain additional information, the components that make up PITI can be mostly or entirely observed using mortgage data and administrative data.[32] Second, the approach also creates some potentially perverse incentives by providing larger benefits for households potentially overextended on housing costs and smaller benefits for households who were more frugal in making their choices. As a consequence, this approach could encourage low-income households to spend a larger portion of their income on housing than those households would have spent without the benefit.

Third, households that are already receiving benefits do not share in any policy cost increases—all premium increases for such households are picked up by the program. This concern could be mitigated by modifying the design to only compensate a maximum percentage of insurance costs, similar to the 80 percent maximum benefit in Design 1. Fourth, this design option may also steer benefits away from low-income policyholders, even though some of these households may able to afford flood insurance using housing burden-based measure of affordability. For example, a family who lives in an inherited home may have low income but also a low PITI ratio and thus would be ineligible for assistance under this program. Conversely, a higher-income household might have a large mortgage and high PITI ratio and thus be eligible for assistance.[33] Using an asset test and the income eligibility cutoff in addition to the PITI ratio could eliminate benefits to more affluent households; however, low-income households without mortgages would still be less likely to qualify for assistance than households with access to credit.
A final potential concern with a PITI-based approach is that households in regions with high costs of living (and mortgage payments) might be more likely to benefit from the program than households in regions with lower costs of living.[34]

Risk Communication

FEMA would notify program participants of their full-risk rate before providing the program benefit to signal that their risk is greater than their premium price reflects. However, price is one of the best signals of risk there is, and households that are above the housing burden eligibility requirement prior to considering flood insurance will not pay anything toward their premiums. Additionally, those households that are just over the housing burden threshold because of higher flood insurance premiums would pay part of their premiums but would not pay their full risk

[31]See Dixon, Clancy, Miller, et al., The Cost and Affordability of Flood Insurance: Economic Impacts of Rising Premiums and Policy Options for One- to Four-Family Homes, RAND, RR-1776, 2017, p. 97.

[32]Other measures include debt-to-income ratios, which often include other loans, such as auto loans, and revolving debt such as credit card debt.

[33]Note, however, that the income cutoff for program eligibility will prevent benefits to households over the income cutoff.

[34]Lenders will presumably not make loans even in the high-cost areas that cause PITI to rise beyond 40 percent of income, which should limit the variation in PITI by geographic region.

premium. These households would have limited incentives to reduce or mitigate their risks, because they do not pay full risk rates.

Program Design 4: Mitigation Grants and Loans

Design Options 1, 2, and 3 subsidize policyholders' premiums by providing financial assistance to an individual unable to pay for insurance. While those design options will reduce the burden of flood insurance premiums for policyholders, they will not change the physical flood risk borne by individual policyholders or the nation as a whole.

Reducing flood risk is an important policy goal for the NFIP, and FEMA offers premium discounts to policyholders and communities where policyholders live if they undertake certain mitigation activities. While structure-specific mitigation activities lead to reduced risk, the resulting premium discounts are often insufficient to cover the cost of mitigation activities. In addition, the premium reduction that policyholders could realize after mitigating their properties depends on their prior flood risk and the type of mitigation effort they undertook. Even with the potential benefits of reduced future losses and decreased current premiums, households are often required to make large upfront costs to achieve the risk reduction that comes from mitigation activities. As households must pay upfront for this stream of benefits, and low-to moderate-income households may be unable to afford those costs.

Design 4 focuses on providing mitigation grants or loans to cover these large upfront costs to achieve cost-effective flood mitigation. In addition to benefiting the policyholder, mitigation can also benefit the Federal Government by reducing the cost of an affordability program and by reducing disaster relief costs.[35] Mitigation grants or loans could be a stand-alone program, or they could be added to any one of the previously discussed designs. Because mitigation measures may not be feasible or cost effective for many homes, mitigation grants and loans alone would not be an adequate affordability program. Therefore, we recommend that mitigation grants and loans could be an add-on to Designs 1, 2, and 3 instead of being the primary delivery mechanism for an affordability program.

Design Characteristics

Low-income households (for example, those with less than 80 percent of AMI) would receive funding for mitigation activities through a grant as these households will likely not qualify for a loan and would have difficulty repaying one. In contrast, moderate-income households (for example, those with 80 to 120 percent of AMI) would receive low-interest loans. Site and structure-specific mitigation measures that currently result in reduced flood insurance premiums include elevating the structure, filling in the basement, elevating utilities so they are less likely to be damaged in a flood, and installing flood vents. To be a feasible option, chosen mitigation measures would need to pass a cost-benefit analysis. Specifically, the reductions in the

[35]By reducing the flood insurance premium, mitigation can reduce the need for assistance and thus the cost of an affordability program. For illustrations of how mitigation can reduce program costs, see Kousky and Kunreuther, *Addressing Affordability in the National Flood Insurance Program*, Resources for the Future and the Wharton School of Business, Issue Brief, 13-02, August 2013; and Dixon, Clancy, Miller, et al., *The Cost and Affordability of Flood Insurance: Economic Impacts of Rising Premiums and Policy Options for One- to Four-Family Homes*, RAND, RR-1776, 2017, pp. 113–114.

discounted present value of flood insurance premiums over time would need to be greater than the cost of the mitigation activity. Community-level mitigation measures, such as green infrastructure and acquisition of NFIP insured structures with the requirement to keep the property as open-space in perpetuity, could also be a viable option to reduce flood risk and premiums. Incorporating such measures into an affordability program is outside the scope of our analysis.

We considered requiring households receiving benefits through Designs 1, 2, or 3 to take cost-effective mitigation measures, however, some mitigation measures may not be feasible or desirable to policyholders that need premium discounts. For example, households may experience considerable inconvenience (such as moving out of the house) while the structure is being modified. Therefore, we recommend that households be encouraged but not be required to take cost-effective mitigation measures if they received benefits through an affordability program.

Advantages and Disadvantages

A mitigation activity provides advantages if the effort reduces flood risk to the home and thereby reduces the discounted present value of flood insurance premiums by more than the cost of the mitigation measure. That is, the funds spent on the mitigation measures are less than the amount the policyholder receives in premium discounts over the life of the policy. In this case, a mitigation measure would have a clear benefit to the policyholder and the Federal Government. Given that mitigation measures may not be feasible or pass a cost-benefit test, such a program may apply to very few policyholders, and provide limited affordability assistance to households relative to the cost of administering the program. Participants in the affordability workshops also noted that mitigation on a property-by-property basis may or may not be to be as cost effective as community-wide mitigation activities in reducing overall risk to flooding.

Mitigation grants and loans have several current and potential challenges associated with them. First, for mitigation grants, structural mitigation may reduce the value of the structure being mitigated and the timeline for implementing grant programs can be lengthy. For example, a filled-in basement could lead to a loss of rental income or make the property less desirable or marketable because of a reduction in usable space. Finally, implementing a loan program would also be complex and would rely on other agencies to conduct cost–benefit analyses on mitigation activities and to administer a loan program.

Risk Communication

In this option, there is full risk communication to the household because mitigation measures that directly reduce individual flood risk also reduce flood insurance premiums. In addition, a household would only receive the grant or loan if it lowered its risk by enough so the premium reductions were greater than the cost of the investment.

Comparison of Program Design Options

Table 3.1 summarizes the key advantages and disadvantages of the designs considered. It also provides a short description of what is required to participate in the program from the policyholder's perspective.

Table 3.1. Comparison of Flood Insurance Affordability Program Design Options

	1. Income-Based Premium Sharing	2. Premium Burden-Based Benefit	3. Housing-Burden Based Benefit	4. Mitigation Grant and Loan Add-On
Overview	Household pays a percentage of flood insurance premium, with percentage rising as income rises Only households below a specified income cutoff are eligible	Household pays up to a specified percentage of income on flood insurance premium, and program pays the rest of the premium Only households below a specified income cutoff are eligible	Benefits provided to low- and moderate-income households with high housing costs relative to income Only households below a specified income cutoff are eligible	Grants and low interest loans for structure-specific mitigation Mitigation must pass a cost–benefit test Program is a voluntary add-on to Designs 1, 2, and 3
Advantages	Simple eligibility criteria Household shares cost of increasing premiums	Simple eligibility criteria Required household contribution can be based on expectations for how much household should pay toward flood insurance	Household responsible for spending a certain percentage on homeownership before receiving benefits Definition of housing burden can be based on lending industry practices	Enables mitigation that makes sense from a cost-benefit basis Reduces need for premium subsidies
Disadvantages	May provide benefits to households for which flood insurance is not unaffordable based on a premium burden or housing burden test No obvious basis for setting program parameters May provide benefits to those who have low income but substantial assets	May provide benefits to households who can afford flood insurance based on a housing burden test Benefits based on only one component of an interrelated basket of housing costs Household does not share cost of increasing premiums May provide benefits to those who have low income but substantial assets	Need to collect more detailed household financial information Rewards households that have taken on too much debt relative to income Can limit benefits delivered to low-income households that own homes outright	Mitigation measures may not be feasible or pass a cost-benefit test for many structures May be little interest in the program Administratively complex
Customer Experience	Household must provide income documentation	Household must provide income documentation	Household must provide information on income, mortgage, property taxes, and insurance	Structure must be evaluated; household may have to move out of structure

Table 3.1. Comparison of Flood Insurance Affordability Program Design Options

	1. Income-Based Premium Sharing	2. Premium Burden-Based Benefit	3. Housing-Burden Based Benefit	4. Mitigation Grant and Loan Add-On
			payments	during mitigation process

IV. How Policymakers Can Assess the Design Options

In this chapter, we estimate the number of beneficiaries, annual benefit cost, and program impact for various versions of each design using the matched ACS/NFIP data. The analysis in this chapter does not provide the cost for or recommend a particular affordability program, rather it uses the simulation models we built to illustrate the magnitude of costs that might be involved and the tradeoffs among different designs. This framework is intended to inform affordability proposals that the Administration and Congress may consider advancing. To that end, we simulate a number of different scenarios for each design, with the scenarios reflecting a wide range of eligibility cutoffs and benefit levels. These cutoffs and benefit levels are not realistic examples of how to implement a particular design but rather illustrate a range of possibilities.

Based on our modeling and analysis using the simulation model we built, we found the following:

- The number of enrollees and costs of each program design largely depend on the program eligibility cutoffs and benefit levels chosen;
- Program Design 1, Income-Based Premium Sharing, provides benefits to all households with income below the specified income cutoffs;
- Program Design 2, Premium Burden-Based Premium Sharing, focuses benefits on households who spend a high percentage of income on flood insurance;
- Program Design 3, Housing Burden-Based Benefit, focuses benefits on households with high PITI ratios. These households also spend a considerable proportion of household income on flood insurance, but not as much as in Design 2. Although this design effectively targets households for whom flood insurance is unaffordable based on a lending industry definition of affordability, it misses certain households where policyholders believe they need assistance to purchase flood insurance. For example, retirees with low incomes who have paid off their mortgages, or households with low incomes and inherited homes may not receive assistance;
- The design options can substantially reduce the proportion of household income spent on flood insurance (addressing one affordability metric), but they usually have limited impact on the PITI ratio because flood insurance is typically not a large part of PITI; and
- Further work is needed to understand the extent to which an affordability program would induce some non-policyholders to purchase coverage, thus increasing the number of households insured against flood losses but also potentially increasing the costs of the affordability program.

Underlying Assumptions for Interpreting Example Findings

In interpreting the findings, we note that the reader should keep in mind the following assumptions and limitations:

1. The estimates capture *annual*, as opposed to one-time, cost.
2. The estimates are for the benefits provided to program participants and do not include the program's set up or administrative costs.

3. The modeled program options deliver benefits only to current NFIP policyholders. FEMA chose to model only current policyholders because we have policy cost for these households.
4. Policymakers must decide whether the benefit should accrue to only current policyholders or accrue to non-policyholders as well.
5. An affordability program would likely induce some non-policyholders to purchase coverage, thus potentially increasing the costs from those estimated here.
6. There must be further work to estimate the magnitude of such an effect.
7. All policyholders eligible for a program will enroll in it. In reality, not all eligible households will enroll, thus reducing the costs from those estimated here.
8. The number of enrollees and the size of the benefits depend on premiums and fees currently paid by policyholders. Increases in premium and fees, other things being equal, will increase the number of policyholders eligible for the program and the benefit per participant.

We do not have sufficient data on the cost and premium impact of mitigation measures for each of our policies to analyze the potential costs of adding a mitigation grant and loan component on to each of the other design options.

In addition, we modified our simulation approach for all design options for modeling PITI. The ACS reports what households spend on insurance, but it does not report an amount for a household in some cases, even though it purchased flood insurance according to the NFIP policy database. Thus, for these simulations, we added the cost of the flood insurance policy to the PITI totals reported in the ACS. As a result, the PITI ratios for policyholders without the program are somewhat higher than they are in those that underlie the distribution of the PITI ratio in Table 2.9. Further work is needed to evaluate the sensitivity of the results to this practice.

We describe the methods we used to develop these estimates in Appendix C.

Design 1: Income-Based Premium Sharing

Table 4.1 presents simulations of four different versions of the income-based sharing design.

The first three rows list the three key parameter values assumed in each scenario:

1. What level of benefit is provided to the lowest income households?
2. What is the income cutoff, measured with respect to AMI, for households receiving maximum benefit?
3. What is the income cutoff for receiving any benefit from the Federal Government?

Recall that a household's benefit falls linearly from the maximum to zero as household income increases from the level chosen for parameter 2 to the level chosen for parameter 3.

Table 4.1. Illustrative Scenarios for an Income-Based Premium Sharing Program

	Scenario			
	A	**B**	**C**	**D**
Program Parameter				
1. Percent of policy cost paid by government for lowest income households	10%	10%	100%	100%
2. Income cutoff for maximum benefit (% of AMI)	10%	10%	70%	160%
3. Program eligibility cutoff (% AMI)	80%	165%	80%	165%
Program Outcomes if Program Only Available to Policyholders in SFHA				
Number of HH receiving benefit	469,000	1,090,000	469,000	1,090,000
Total benefit paid per year	$21 M	$53 M	$424 M	$995 M
Average Benefit per HH	$45	$49	$906	$912
Program Outcomes if Program Available to All Policyholders				
Number of HHs receiving Benefit	858,000	2,087,000	858,000	2,087,000
Total benefit paid per year	$29 M	$75 M	$593 M	$1,552 M
Average benefit per HH	$34	$36	$691	$744

SOURCE: FEMA analysis of NFIP policyholder data and Census ACS data.
AMI: Area Median Income; HH: households.

Scenarios A and B are very minimal versions of the program. The government pays a maximum of 10 percent of the policy cost (premium plus fees), and this maximum benefit is available only to households with incomes less than 10 percent of AMI. Only low-income households are eligible to participate in Scenario A (households with income less than 80 percent of AMI), while Scenario B extends the program to moderate- and middle-income households (households with incomes less than 165 percent of AMI). Scenarios C and D provide much more substantial benefits to the same sets of households: The government pays the full policy cost for households below 70 percent and 160 percent of AMI, respectively.

Table 4.1 presents estimates first assuming the program is only open to current policyholders in SFHAs (the middle of the table) and second assuming that program is open to all policyholders (the bottom of the table). The results for Scenario A show that 469,000 policyholders in SFHAs are eligible for the program and that annual benefit payments total $21 million, translating into $45 per beneficiary on average. The number of beneficiaries is around 28 percent of the 1.8 million policyholders in SFHAs represented in the analysis. The number of beneficiaries rises to 1.090 million in Scenario D when the program is restricted to SFHAs.

Total benefits for this much more expansive program come in at slightly less than $1 billion per year, with average benefits around $912 per enrollee. The number of beneficiaries and total

program costs increase considerably when the program is open to all 3.7 million residential policyholders, whether they live inside or outside of SFHAs (bottom rows of Table 4.1). Table 4.2 provides some initial information on the impact of the program on the amount paid for flood insurance by program participants and on our two affordability metrics. In Scenario D, for example, the average policy cost (premium plus fees) falls from $755 to $12. Because nearly all program beneficiaries in Scenario D receive the full policy cost, the median share of household income spent of flood insurance for program participants falls from 1.2 percent without the program to none with the program. Even with the large benefits, the median PITI ratio falls only slightly from 0.22 to 0.21 because flood insurance premiums are often not large compared with the other components of PITI. This design provides benefits to households that typically spend a low percentage of income on flood insurance and have modest PITI ratios. The median percentage of income spent of flood insurance without the program varies from 1.2 percent to 2.2 percent, and the median PITI ratio without the program varies from 0.22 to 0.37.

Table 4.2. Illustrative Scenarios for Program Available Both Inside and Outside SFHA (for program beneficiaries only)

Outcome	Scenario			
	A	B	C	D
Average policyholder cost				
With program	$727	$719	$71	$12
Without program	$762	$755	$762	$755
Median policyholder cost				
With program	$474	$463	$0	$0
Without program	$490	$485	$490	$485
Median of percentage of income spent on flood insurance				
With program	2.1%	1.1%	0%	0%
Without program	2.2%	1.2%	2.2%	1.2%
Median PITI ratio*				
With program	0.37	0.22	0.35	0.21
Without program	0.37	0.22	0.37	0.22
Number of program participants	858,000	2,087,000	858,000	2,087,000

SOURCE: FEMA analysis of NFIP policyholder data and Census ACS data.
*Excludes households occupying a residence they do not own without payment of rent.

Design 2: Premium Burden-Based Benefit

Table 4.3 presents simulations of three different versions of the income-based sharing design. The key parameters for a program based on premium burden are:

1. The required contribution to flood insurance costs by income category;
2. The income cutoff for receiving any benefit from the Federal Government.

In Scenario A of Table 4.3, policyholders within incomes below 165 percent of AMI (which covers low, moderate, and middle-income households) would not pay any part of the policy costs. In Scenario B, which is more similar to how health care subsidies are structured, the eligible households would be required to increase their contribution with increased household income, reaching 4 percent for middle-income households. Scenario C provides the least generous benefit, with all eligible households required to contribute 4 percent of income toward the flood insurance premium before receiving the federal benefit.[36]

Table 4.3. Illustrative Scenarios for a Premium Burden-Based Benefit Program

	Scenario		
	A	**B**	**C**
Program Parameter			
1. Required flood insurance payment as a percentage of household income			
Extremely low income	0%	0%	4%
Very low income	0%	1%	4%
Low income	0%	2%	4%
Moderate income	0%	3%	4%
Middle income	0%	4%	4%
2. Program eligibility cutoff (% of AMI)	165%	165%	165%
Program Outcomes if Program Only Available to Policyholders in SFHA			
Number of HH receiving benefit	1,090,000	408,000	211,000
Total benefit paid per year	$1,101 M	$326 M	$154 M
Average benefit per HH	$1,010	$799	$731
Program Outcomes if Program Available to All Policyholders			
Number of HH receiving benefit	2,087,000	614,000	265,000
Total benefit paid per year	$1,576 M	$398 M	$170 M
Average benefit per HH	$755	$649	$642

SOURCE: FEMA analysis of NFIP policyholder data and Census ACS data.
AMI: Area Median Income; HH: households.

As shown in the bottom portion of Table 4.3, the number of beneficiaries varies by around a factor of five across these very different parameterizations of the program. When households are not required to contribute to flood insurance, annual benefit payments inside the SFHA total just over $1.1 billion, and approximately 60 percent (1.090 million) of the 1.8 million policyholders in SFHA participate. The results for Scenario C lie at the other end of the spectrum, with 12 percent (211,000) of policyholders participating. Extending the program to policyholders outside

[36]Households paying 4 percent of income toward premiums would end up paying a higher percentage of income to premiums plus fees.

the SFHA increases the number of beneficiaries and program cost. However, the costs and number of beneficiaries do not increase significantly in Scenario C because flood insurance premiums are lower outside SFHAs and, households are required to contribute a substantial share of income toward the premium before receiving a subsidy.
As shown in Table 4.4, the average amount paid for flood insurance by program beneficiaries falls substantially. In Scenario B, for example, the average amount paid falls from $1,068 without the program to $419 with the program.

Table 4.4. Illustrative Scenarios for Burden-Based Benefit Program When Program Available Both Inside and Outside SFHA (for program beneficiaries only)

Outcome	Scenario		
	A	B	C
Average policyholder cost			
With program	$0	$419	$861
Without program	$755	$1,068	$1,503
Median policyholder cost			
With program	$0	$220	$649
Without program	$485	$710	$1,309
Median of percentage of income spent on flood insurance			
With program	0%	1.0%	4.0%
Without program	1.2%	3.6%	6.5%
Median PITI ratio*			
With program	0.21	0.42	0.56
Without program	0.22	0.44	0.60
Number of program participants	2,087,000	614,000	265,000

SOURCE: FEMA analysis of NFIP policyholder data and Census ACS data.
*Excludes households occupying a residence they do not own without payment of rent.

Scenarios B and C target households who spend a substantial portion of income on flood insurance and have high PITI ratios. For example, the median percentage of income spent on flood insurance is 3.6 percent for program participants without the program in Scenario B and 6.5 percent in Scenario C. The median PITI ratios without the program are 0.44 and 0.60 for Scenarios B and C, respectively. By contrast, 50 percent of the beneficiaries in Scenario A spend less than 1.2 percent of income on flood insurance without the program and 50 percent have a PITI ratio of less than 0.22 without the program. The reason for these outcomes is that Scenario A opens the program to all policyholders with incomes less than or equal to 165 percent of AMI.

Design 3: Housing Burden-Based Benefit

The key parameters for the program based on housing burden are:

1. The PITI ratio cutoff; and
2. The income cutoff for receiving any benefit from the Federal Government.

As discussed in Chapter 3, one of the advantages of this program is that program eligibility is based on practices in the lending industry. We determined the PITI ratio cutoff of 0.40 in Scenario B because of the mortgage industry standard in which loans that result in the PITI to exceed 40 percent of income are seldom approved. When the PITI cutoff is set to 0.40, households receive assistance for that part of the policy costs that causes their PITI ratio to exceed 0.40. Scenarios A and C allow the PITI ratio cutoff to range from 0.10 to 0.70. We previously stated that a small percentage of homeowners (4 percent) have PITI ratios above 0.70. In all scenarios, low, moderate, and middle-income households are eligible to participate in the program.

Table 4.5. Illustrative Scenarios for a Housing Burden-Based Benefit Program

	Scenario		
	A	B	C
Program Parameter			
1. PITI ratio cutoff	0.10	0.40	0.70
2. Program eligibility cutoff (% of AMI)	165%	165%	165%
Program Outcomes if Program Only Available to Policyholders in SFHA			
Number of HH receiving benefit	926,000	288,000	120,000
Total benefit paid per year	$943 M	$302 M	$126 M
Average benefit per HH	$1,018	$1,048	$1,045
Program Outcomes if Program Available to All Policyholders			
Number of HH receiving benefit	1,654,000	472,000	194,000
Total benefit paid per year	$1,293 M	$394 M	$162 M
Average benefit per HH	$782	$834	$835

SOURCE: FEMA analysis of NFIP policyholder data and Census ACS data.
NOTE: Program not open to households occupying a residence they do not own without payment of rent; AMI = Area Median Income.

In Design 3, 288,000 policyholders in the SFHA receive benefits under Scenario B (middle portion of Table 4.5). Annual benefits come to $302 million per year, with an average benefit per participating household of $1,048. The number of beneficiaries and annual benefit payments vary substantially as the program parameters change.

As expected, the program targets households with high PITI ratios, except when the PITI ratio cutoff is low. For example, the median PITI ratio without the program is 0.62 in Scenario B and 1.10 in Scenario C ("median PITI ratio" rows of Table 4.6). Scenario B and C also end up providing benefits to households that spend a substantial share of household income on flood insurance. The design reduces the PITI ratio somewhat for participating households and reduces the percentage of household income spent on flood insurance a great deal. Although this

program provides benefits to households that tend to have a high PITI ratio and spend a substantial share of income on flood insurance, it does miss some households that may be of concern to policymakers. In particular, it may not provide benefits to retirees who have paid off their mortgages or low-income households that have inherited the property, mortgage-free.

Table 4.6. Illustrative Scenarios for Impact of Housing Burden-Based Benefit When Program Is Available to All Policyholders (for program beneficiaries only)

Outcome	Scenario		
	A	B	C
Average policyholder cost			
With program	$20	$60	$60
Without program	$801	$894	$895
Median policyholder cost*			
With program	$0	$0	$0
Without program	$509	$562	$561
Median of percentage of income spent on flood insurance*			
With program	0%	0%	0%
Without program	1.3%	2.8%	5.1%
Median PITI ratio*			
With program	0.25	0.58	1.04
Without program	0.27	0.62	1.10
Number of program participants	1,654,000	472,000	194,000

SOURCE: FEMA analysis of NFIP policyholder data and Census ACS data.
*Excludes households occupying a residence they do not own without payment of rent.

V. Administrative and Funding Options

FEMA administers the NFIP with private insurers participating in the Write Your Own (WYO) program. The NFIP has not been structured to assess eligibility for benefit assistance and deliver those benefits. As a result, we also worked with the workshop participants to develop both administration and funding options that could be used to implement the flood insurance affordability program design options. We asked workshop participants to consider (1) how to administer an assistance program inside and outside the NFIP, and (2) how to fund such an assistance program. This chapter discusses three administration options and three funding options that emerged from prior research and workshop discussions.

The three administration options include the following:

- FEMA partners with the Internal Revenue Service (IRS) to administer the benefit assistance by either providing the benefit as a deduction on policyholders' taxes or by providing the benefit directly to the insurance company at the time of purchase.
- FEMA's NFIP Direct administers the benefit program, treating those eligible as a separate special handling group. The NFIP Direct could enlist the help of another federal agency to conduct the eligibility determinations.
- States and the Small Business Administration (SBA) administer the part of the benefit program providing mitigation grants and loans.

The three funding options include the following:

- Internally funding the program by collecting additional fees and further cross-subsidization.
- Externally by receiving an appropriation from Congress that funds the assistance program.
- A hybrid approach by securing funding from a combination of internal and external sources.

We describe the various options and discuss their advantages and disadvantages in the remainder of this chapter.

Administration Options

Based on research and discussion at the workshops, we developed three administration options for implementing the four design options described in Chapter 3 and illustrated in the simulation model in Chapter 4. All of these options would benefit from communicating a policyholder's full-risk rate even if policyholders are not required to pay the full-risk premium. Because price is an excellent signal of risk, communicating a policyholder's full-risk rate could play an important role in communicating flood risk.

Option 1: IRS Tax Credit

There are two main approaches for how to engage the IRS to implement an affordability design option subsidy. The most straightforward approach would be for the policyholder to pay the flood insurance premium and then receive the appropriate tax credit when they file income taxes in April of each year (version 1).[37]

Advantages and Disadvantages

The advantage of this administration option would be that once it is set up, is that the IRS would administer the program allowing FEMA to play an oversight role in executing this process. However, this approach faces certain disadvantages; first, households may not be able to pay the full premium at the start of the premium year and second, this approach is complicated and requires that the insurance company or FEMA determine how much to collect from the policyholder and reconcile payments received from the IRS. In addition, the process of IRS agreeing to implement this programs—and FEMA entering into an agreement with IRS, setting up and implementing the procedures—would be a complex and extensive effort. In addition, as workshop participants noted, it may be difficult to obtain data from the IRS necessary to monitor the program and track the total amount of tax credits issued.

Option 2: FEMA through NFIP Direct

In this administrative option, policyholders eligible for an affordability program could be added to FEMA's NFIP Direct as another special handling group, whose policies are written and serviced through NFIP Direct. NFIP Direct writes and services approximately 13 percent of the NFIP's flood insurance contracts and services groups of policies such as Severe Repetitive Loss Properties and Group Flood Insurance Policies. Under an NFIP Direct administered program, FEMA would establish the eligibility criteria using one of the design options previously described and establish a relationship with another agency (like HUD or IRS) to determine eligibility. The policy declaration page could show the full flood premium, could show the amount of the affordability credit (the subsidy), and then could show the reduced amount owed by the policyholder. The WYO program allows participating private insurers to write and service the standard flood insurance policy (the NFIP policy), because private insurers largely do not offer residential flood insurance. For insurance companies not participating in the WYO program, their agents can still obtain NFIP flood insurance coverage for their policyholders by working through NFIP Direct on their clients' behalf.

After discussions held at the second workshop, we determined that NFIP Direct presents a more feasible option for FEMA to administer an affordability program than using WYOs. There are several reasons that make it difficult for private insurers to administer an affordability program. First, WYOs are required to receive the full flood insurance premium payment at the beginning of a household's policy year. In order for a WYO to change how they accept payments from policyholders and accept a benefit payment, FEMA would have to write new rules to accept partial payments. The WYOs also have concerns related to collecting personal information such as income. Agents also do not want to be a conduit for the handling of money between the

[37]This would be a similar model to the first-time homebuyer's tax credit.

government and potential policyholders. As a result, administering the program through the WYOs would be much more challenging to navigate than if FEMA administered the program through NFIP Direct.

Advantages and Disadvantages

The primary advantage that FEMA's NFIP Direct administration option has over the IRS option is that FEMA would have increased control of the program. Thus, if modifications had to be made to increase or decrease eligibility criteria or subsidy amount, the adjustments could be made through the NFIP Direct. Making adjustments or mid-course corrections could prove difficult when utilizing the IRS. Another advantage in administering the program through the NFIP Direct is lowered expense because of internally contained processes and logistics as opposed to a WYO administered program. The program could be relatively seamless from the policyholder's perspective because NFIP Direct would need to interact with the agency conducting eligibility determinations. NFIP Direct and the assisting agency determines eligibility and the insurance agents continue their relationships with policyholders to help them with their insurance needs without taking on affordability program legal and financial risks. Agents would be able to place their business with NFIP Direct and policyholders applying for assistance could retain a relationship with their trusted agents.

A disadvantage of the NFIP Direct option is the NFIP Direct would rely on another agency for the eligibility determinations and FEMA would require additional resources and staff with relevant skill sets to work with the assisting agency on eligibility determinations. In addition, policyholders would face an additional administrative burden to provide required documents, unless eligibility was determined through other means.

Option 3: FEMA with Assistance from States and Small Business Administration

While the other administration options are applicable to any of the design options, this administration option is restricted to the mitigation grant or loan add-on design option. As noted earlier, mitigation is an important tool to reduce the risk and the impact of flooding; however, NFIP does not provide loans for mitigation assistance directly to policyholders. Operating such a program could require FEMA to devote significant resources to implement the program, in addition to any grants or loans it might provide. For example, FEMA would be required to increase its resources available to evaluate mitigation activities for cost-effectiveness, hire staff to monitor implementation of mitigation projects, and certify the completion of that mitigation projects. FEMA would most likely need to partner with other agencies to execute the loan program.

The NFIP also lacks the authority to make loans to policyholders, although FEMA may be able to enlist the SBA to administer a mitigation loan program, akin to what the SBA already provides for post-disaster mitigation loans.[38] According to workshop participants, FEMA could

[38]Another federal agency option for administering a mitigation loan program could be HUD given that the Federal Housing Administration currently administers the 203K Rehabilitation Mortgage program that allows homeowners to borrow up to $35,000 to repair a home damaged from flooding or to mitigate against future flood damage.

model a mitigation program after DOE's Weatherization Program. This program works with state and local agencies to assess the viability of weatherization measures and to implement them. FEMA does not have those formal relationships established currently.

Advantages and Disadvantages

The main advantage of a mitigation program is relevant mitigation activities will help to reduce the risk of flooding, which lowers flood insurance premiums for policyholders and potential future payouts by the Federal Government in the form of subsidies or disaster relief. The main disadvantage for FEMA is that the agency would have to establish and rely on partnerships with other state and local agencies that have the capacity to manage a grant and loan program, as well as develop a mechanism to monitor their work on the affordability program. Another disadvantage is that the amount of funding may not cover demand, so the states will need a mechanism to prioritize who would receive assistance.

Comparison of Administration Options

Table 5.1 summarizes the different administration options, compares their advantages and disadvantages, and provides a short description of the policyholder's experience.

Table 5.1. Summary of Advantages and Disadvantages of Administration Options

	1A. IRS Tax Credit (Version 1)	1.B. IRS Tax Credit (Version 2)	2. FEMA Through NFIP Direct	3. FEMA with Assistance from States or SBA
Overview	Households receive program benefits when it files federal taxes in April of each year	Benefits provided to insurer by IRS as a pre-tax credit	NFIP Direct administers the program and enters into an agreement with another agency to determine eligibility and benefit levels	FEMA partners with state agencies and/or SBA to administer the program
Advantages	FEMA plays oversight, not administrative, role NFIP continues to price and sell policies as before	FEMA plays oversight, not administrative, role NFIP continues to price and sell policies as before Household does not have to front full premium	Relies on experienced agency to determine eligibility Providing benefits only through NFIP Direct (as opposed to WYOs) reduces administrative costs and simplifies process	Relies on experienced state and local agencies to evaluate and approve mitigation measures Employs existing SBA lending capabilities to make low-interest loans; definition of housing burden can be based on lending industry practices
Disadvantages	May be difficult to monitor IRS's implementation of the program and determine subsidy payments Households may not be able to front flood insurance premium	IRS or FEMA must determine how much to collect from a policyholder and reconcile payments made to insurers Complexity of IRS payments insurers	Lack of direct control over eligibility determination process Cost of coordinating with assisting agency	Coordination and oversight costs of partnering with state and local agencies and the SBA Amount of funding may not cover demand, requiring applicant prioritization
Customer Experience	Household files required forms with IRS at tax time	Household registers with IRS to determine benefit and then purchases policy from the NFIP Direct or WYOs	Agent refers household to the NFIP Direct; household registers with and provides documents to assisting agency	Household contacts state or local agency for mitigation grants and SBA for low-interest loans

Funding Options

Based on information provided by workshop participants, we identified three primary mechanisms to fund an affordability program for NFIP. First, an affordability program could be funded through policyholder premiums and fees. Second, Congress could provide an appropriation to fund the affordability program. Finally, Congress could enact a mix of these two options. We discuss these options below:

Option 1: NFIP Funded Affordability Program

A program funded through additional policyholder fees or through cross-subsidization has the advantage of being internally contained within NFIP. The program would be contained within a single agency and program from both a legislative and budgetary perspective. The downside to this option is that existing discounts and cross-subsidies are already creating undue pressure on the solvency and sustainability of the program. Currently, the NFIP receives its funding from the NFIP premiums and fees collected from policyholders. In addition, NFIP policyholders pay for mitigation grants and some mapping activities through their premiums and fees. If NFIP requires policyholders to fund an affordability program, annual rates or fees for at least some policyholders would have to increase further. Because some policyholders already consider these rates as unduly burdensome, the increases could create a more onerous affordability challenge for the program.

If premiums increased because Congress chooses to fund an affordability program from policyholder fees, the NFIP could eventually face a situation in which the cost of a low-risk policy would be greater than the benefit a policyholder gets from holding a policy—in these situations, many such policyholders would drop their coverage. This situation would result in adverse selection, because only those policyholders who consider the risk of flooding greater than or equal to their willingness to pay—i.e., they believe the coverage is worth the cost of increased insurance—would maintain flood insurance. Those who have lower risk would drop their coverage rather than pay for a policy they do not perceive as priced accordingly.
In addition to driving away low-risk policyholders, an affordability program funded fully by NFIP policyholders may entice high-risk, cost-burdened homeowners to enter the program without paying their full risk. While this would help the NFIP achieve the goal of greater participation in the program, it would serve to further geographically concentrate risk.

The unintended consequence of both non-mandatory purchase policyholders leaving the program and high-risk policyholders entering the program is that it could reduce the NFIP's ability to achieve a sound financial framework. FEMA assumes that most of the non-mandatory purchase structures are low-risk or non-SFHA and that the occupants of those structures would perceive their risk to be low from their experience, proximity to a flooding source, or FEMA's Flood Insurance Rate Maps. Further, an adverse selection spiral could result in a situation in which the average risk of the group increases as lower-risk properties opt out, leaving this group essentially underfunded by the premiums paid.

Option 2: Congressionally Appropriated Affordability Program

External funding could allow FEMA to implement an affordability program without further degrading the solvency and sustainability of the NFIP or creating additional affordability challenges for current or prospective policyholders. This option would require legislative action to appropriate funds to an affordability program and allow the program to expand and contract depending on political will and available funding. One challenge is that these annual appropriations may not be offered consistently, thereby leaving FEMA unable to consistently deliver a program to those who are eligible for assistance. Demand for assistance would also be inconsistent annually because the ability to afford flood insurance is counter-cyclical. Generally, when the macro economy is at its worst, demand for assistance programs are at their highest.

Demand for an affordability program is also likely to expand as premiums rise because of movement toward risk-based premiums and changes in physical risk.

Option 3. Hybrid Funded Affordability Program

A hybrid program whereby some of the funding comes directly from NFIP and another portion comes from other sources is an intermediary between an NFIP funded versus congressionally appropriated affordability program. This may be legislatively more attractive, because full funding of the program would not have to come from other sources. However, as was discussed in the NFIP funded program, there are already undue pressures on the solvency and affordability of the NFIP because of existing discounts and cross-subsidies. Having the NFIP fund an affordability program, even if only in part, will continue to create more affordability challenges.

VI. Conclusions

Congress faces important policy challenges as it considers options for legislating a program to make flood insurance accessible and affordable and to enable policyholders and communities to rebound more quickly and fully from flooding events than they would without flood insurance. Part of the challenge has been the lack of representative nationwide data to assess the extent of the affordability challenge and to quantify the costs and benefits of potential solutions. To address this shortcoming, FEMA developed a first-of-its-kind database that links NFIP policy information with household income and housing cost information from Census. These data provide policymakers with a much better understanding of the extent of the problem and of the particular population groups most in need of assistance to afford their flood insurance premiums. Because there is a no widely accepted definition of flood insurance affordability, we developed several alternatives based on a household's cost burden or ability to pay. Each alternative has advantages and disadvantages, with ramifications for determining who is and who is not in need of assistance. The data developed from this project is intended to provide helpful information on the number of households for whom flood insurance may be difficult to afford based on the various measures.

A number of different options exist for designing, administering, and paying for a flood insurance affordability program. The program design options vary in the types of households targeted and the amount of information program applicants are required to provide. Some program designs can provide benefits to a broad range of households, but they may also end up providing benefits to households for whom flood insurance is not financially burdensome according to other affordability metrics. Others can more narrowly target benefits, but they may do so at the cost of excluding households that some stakeholders may deem in need of assistance. We also found that flood insurance affordability programs have the ability to create perverse incentives. For example, some version of the designs examined rewards households who make risky financial decisions, such as taking out large mortgages that strain their ability to pay. Other design options end up encouraging lower-income households eligible for assistance to purchase properties in very risky areas, such as where flood insurance premiums are high, but where property values and mortgages are relatively low.

For all the program designs, we emphasize the importance of communicating information to policyholders about the full risks (and the associated full-risk flood insurance rates) of living in flood-prone areas. We also emphasize the importance of targeting potential policyholders in addition to current policyholders for assistance. A flood insurance affordability program will presumably make the purchase of flood insurance more feasible and more attractive for households that currently do not carry flood insurance and thus are not able to recover quickly or as fully after a flood event—often relying on a combination of charity and federal disaster assistance to partially recover. However, good public policy must balance the undeniable benefits of increased flood insurance take-up with increased program costs due to larger program enrollment of policyholders paying less than full-risk rates.

The options identified for administering and paying for the program also vary in terms of advantages and disadvantages. One of the underlying challenges for all the administrative options is each requires FEMA to expand the current number of NFIP staff and requires it to

obtain additional information, such as household income and housing burden, to determine who is eligible. As far as funding options are concerned, policymakers must weigh several factors when considering how to pay for an affordability program. With an appropriation, FEMA would develop an affordability program commensurate with the size of the appropriation. If NFIP provides the funding through a cross-subsidy, FEMA would have to raise rates on other policyholders to pay for the program. If NFIP provides the funding through a pure subsidy, it will not collect premiums and fees necessary to pay claims in the future. An affordability program funded by NFIP premiums and fees, regardless of mechanism, would reduce the NFIP's ability to cover the cost of certain flood events, create additional affordability challenges, and work against creating a sound financial framework for the NFIP.

In creating a flood insurance affordability framework, we developed a set of simulation models that we used to estimate program costs and outcomes for a variety of program designs and parameterizations. Such models are very useful for estimating the number of beneficiaries, program costs, and program outputs with respect to various affordability metrics. The estimates made during the study illustrate the wide variation in costs across different program designs and the sensitivity of program costs and outcomes conforming to the details of the program design. The estimates provide some insight into the types of households that do and do not receive benefits from the various alternatives, but there is a need for more detailed analyses to better understand the types that households that receive benefits and the types the designs overlook. In addition, for simplicity, our analysis focused on current policyholders, but there must be additional work to better understand the extent to which an affordability program would lead non-policyholders to sign up for flood insurance. The model simulations also do not specify important features of an affordability program, such as whether the program will be limited to properties in high-risk flood zones, to current residents of eligible properties, to primary residences, or to homeowners (as opposed to renters). These simulation models are useful for exploring specifications that are more detailed.

Flood insurance is a valuable tool that helps survivors recover in the aftermath of devastating flood events and reduces federal disaster expenditures on individual assistance. Some policyholders and non-policyholders at high risk of flooding face critical tradeoffs between purchasing flood insurance and basic necessities. FEMA believes that this framework will help members of Congress and other stakeholders think through the advantages and disadvantages of different options for a flood insurance affordability program.

Appendix A. Data and Statistical Methods

FEMA worked with the United States Census Bureau (Census) to learn about NFIP policyholder and potential policyholder incomes by matching NFIP policyholder data to the American Community Survey (ACS). The ACS is a relatively new survey conducted by Census that uses a series of monthly samples to produce annually updated estimates for the same small areas (census tracts and block groups) formerly surveyed via the decennial census long-form sample. Academics, policy communities, and government agencies utilize the ACS. Census conducts the ACS throughout the United States and in Puerto Rico, where it is called the Puerto Rico Community Survey (PRCS).[39] The ACS includes households that have residents who, as of the date of the interview, either are living or intend to live at the address for more than 2 months. For example, if someone had just recently moved into a new home they would be included in the sample address because they intend to live there for more than 2 months. Alternatively, if no one is living at the sampled address at the time of the interview, even if someone spends more than two months of the year there (such as a vacation home), the housing unit would be considered vacant.[40] ACS data allows FEMA to quantify affordability challenges and determine the scope of potential affordability programs.

FEMA's data included NFIP policyholders who had active flood insurance policies for any duration during calendar year 2015. FEMA's sample included more policyholders than are reported in any publicly available dataset because publicly available sources capture a snapshot of NFIP policies as of one specific day. We provided this additional data to maximize the chances of matching with the data Census collected throughout the year.

Please note that throughout the framework, FEMA calculated column totals based on the raw data from each row, then rounded each individual row and the column total for ease of reporting. As a result, column totals may differ slightly from the sum of each reported row, particularly in Chapter 2.

Limitations of ACS Data

Before we describe how we linked the NFIP and ACS data, we review some limitation of the ACS data for the purposes of analyzing flood insurance affordability.

All self-reported survey data, including the ACS, comes with a variety of limitations. For example, self-report income may differ from the administrative records of income that would be used to determine program eligibility. Respondents may incorrectly recall their income or may forget sources of income. Misreported income could cause ACS-based estimates of total program cost to be incorrect due to over- or under-estimating the number of individuals eligible for the program. Relative to the administrative records in the National Income and Product

[39] ACS Design and Methodology Report, U.S. Census Bureau, www.census.gov. https://www.census.gov/programs-surveys/acs/methodology/design-and-methodology.html (accessed, July 12, 2017.)

[40] Ibid.

Accounts, the ACS tends to under-estimate income.[41] This could cause our analysis to over-estimate program costs by over-estimating the number of eligible homeowners.
Similarly, self-reported survey data may over- or under-estimate the costs associated with home ownership. The direction of this bias is less certain, as many costs associated with homeownership do not have administrative data that can be compared to the reported survey data. For under-reporting, we do know that the estimated annual insurance costs for many households are lower than the household's annual flood insurance premium. Certain types of respondents may also be less likely to answer certain questions. Research on another Census survey found respondents with higher income tended to answer more survey questions in general.[42] If respondents, particularly lower income respondents, are less likely to report housing costs, we might under-estimate their housing burden and hence under-estimate their program eligibility.

FEMA and ACS also use different definitions of residency. NFIP focuses on the "primary resident" of a home. For the NFIP, a primary residence is one in which the occupant lives more than 50 percent of the year. For ACS, the Census Bureau uses the concept of "current residence" to determine residents of sample housing units. The basic idea behind "current resident" is that everyone who is currently living or staying at an address, or intends to live there, for more than two months is considered a current resident of that address. Persons away from their residence for two months or less, whether in the United States or overseas, on a vacation or on a business trip, are a "resident" at the address, and Census considers the unit as occupied and the residents eligible for inclusion in the survey. Persons away from their residence for more than two months at the time of the interview are not included (for example, college students away at school, persons with two homes who are living at the other home at the time of the interview).For the ACS, if no one is determined to be a current resident in the sampled housing unit, Census classifies it as "vacant."[43]

As noted above, Census does not define primary resident the same as FEMA.[44] For that reason, we did not discern between primary residence and non-primary residence, as defined by FEMA,

[41]Rothbaum, Jonathan L., "Comparing Income Aggregates: How do the CPS and ACS Match the National Income and Product Accounts, 2007-2012," SEHSD Working Paper 2015-01 (2015). Available at https://www.census.gov/content/dam/Census/library/working-papers/2015/demo/SEHSD-WP2015-01.pdf

[42]Hedengren, David, and Thomas Stratmann. "The Dog that Didn't Bark: What Item Nonresponse Shows about Cognitive and Non-Cognitive Ability." (2012).

[43]There are a few exceptions to the "two-month" rule: (1) Persons Without Another Place to Stay -- Anyone staying at a residence who does not have another place to stay, even if they are at the residence for two months or less, are current residents, (2) Children Away at School -- Children (below college age) who are away at boarding school or summer camp for more than two months are always current residents of their parents' homes. College students establish current residency using the two-month rule, (3) Children in Joint Custody -- Children who live under joint custody agreements and move between residences are current residents of the sampled housing unit where they are staying at the time of the interview, and (4) "Commuter Workers" -- People who stay at a residence close to work and return regularly to another residence to be with their family are always considered current residents of the family residence, not the work-related residence. For more information, see "Frequently Asked Questions," U.S. Census Bureau, accessed, August 10, 2017. https://ask.census.gov/prweb/PRServletCustom?pyActivity=pyMobileSnapStart&ArticleID=KCP-2892.

[44]For FEMA, a primary residence is one in which a resident lives at least 50 percent of the year.

in our analysis. Additionally, we excluded non-residential NFIP policyholders from our analysis. Furthermore, the costs and number of participants that we estimated in our models assume residential households are either primary or non-primary. Our model estimates may deliver benefits to non-primary residence households.

Matching NFIP and ACS Data

Census created two unique identification numbers for NFIP policyholders, one derived from NFIP policyholders' identity and the other by location.[45] FEMA used the unique identification numbers to match policyholder data with 2015 one-year estimates of ACS households. Using the matched data, FEMA identified ACS households that have flood insurance, whether they reside in or out of a Special Flood Hazard Area (SFHA), and whether they reside in a coastal or riverine area. FEMA included NFIP residential policyholders and removed business policyholders and other non-residential policyholders. FEMA excluded ACS residential households that did not have a reported adjusted household annual income or had an adjusted household annual income of zero or below from our analysis.

The NFIP sells contracts for multi-family residences called RCBAP or Residential Condominium Building Association Policy policies. These policies cover damage to the condominium complex, not damage to units owned by individuals. The policy for the building association and the individual unit owner policies are at the same geographic location and appear as one contract in the NFIP database. For the purposes of our analysis, we assumed that in multi-family residences with an NFIP policy which responded to the ACS the respondent was representative of all households living in that residence. However, we did not modify the premium paid by this household to reflect their proportion of the premium.

There were 4,508,338 residential NFIP policyholders who had active policies at any time during calendar year 2015.[46] We matched 64,703 of the 4.5 million policyholders to ACS households by identity, not location. Using Census assigned weights; we calculated a weighted sample of 3,674,866 NFIP policyholders, which underestimates the true population size.

Representativeness of Matched Sample

To understand whether our weighted sample is an accurate representation of the actual number of NFIP residential policyholders, we show the comparisons of different sub-categories between our populations and the estimated population. We also tested whether the weighted sample of NFIP policyholders represented the same types of households as the actual population of policyholders. We selected several subsets of the weighted sample and compared their proportions with the proportions of the actual population. Furthermore, we calculated the standard error of the different subsets to estimate a 90 percent confidence interval to determine if

[45] See "The Person Identification Validation System (PVS): Applying the Center for Administrative Records Research and Applications' (CARRA) Record Linkage Software", Accessed August 22, 2017. https://www2.census.gov/programs-surveys/acs/tech_docs/pums/accuracy/2015AccuracyPUMS.pdf.

[46] We used the term policies to mean Contracts in Force.

the weighted sample is different from the actual population of NFIP policyholders.[47] In other words, we determined whether our weighted sample is representative of the actual population of NFIP policyholders.[48]

In testing the applicability of ACS weighting to the NFIP policyholder base, we approximated the standard error (SE) for NFIP premiums by using 80 different replicate weights values following the formula Census provided.

Formula A.1: Approximating the Standard Error

The standard error of X can be approximated after the replicate estimates X_1 through X_{80} are computed. The standard error is estimated using the sum of squared differences between each replicate estimate X_r and the full sample estimate X. The standard error formula is:

$$SE(X) = \sqrt{\frac{4}{80}\sum_{r=1}^{80}(X_r - X)^2}$$

Using the SE, we calculated the Margin of Error (ME) to estimate the 90 percent confidence interval by

1. multiplying the SE by 1.645 and
2. adding or subtracting the ME to or from the weighted variable of interest (See formula X above).

Formula A.2: Calculating the 90% Confidence Interval

LB=Lower bound = X - 1.645*SE(X)
UB=Upper bound = X + 1.645*SE(X)

The 90% confidence interval is the interval (LB, UB).

Then, we compared the upper and lower bounds of the 90 percent confidence interval to the NFIP premiums of different subgroups in the population.

Table A.1 reports the results of our analysis. The weighted sample underestimates the number of single-family homes and overestimates the proportion of single-family homes (relative to 2-4 family homes and other family homes). The weighted sample also slightly over-represents policyholders outside of the SFHA (relative to policyholders within the SFHA).

[47] Standard error is a statistical term that measures the accuracy with which a sample represents a population. Confidence interval indicates a range of values that is likely to encompass the true value. More formally, you calculate the confidence interval around your sample statistic so that it has a specified chance of surrounding (or containing) the value of the corresponding population parameter.

[48] See pdf attached PUMS Accuracy of the Data (2015)

Table A.1. Comparison of NFIP Portfolio of Residential Policies with Weighted Sample of Policyholders

	NFIP Residential Policyholders		Weighted Sample of NFIP Residential Policyholders	
	Number (in thousands)	Percentage of Total	Number (in thousands)	Percentage of Total
Total	4,508,338	100%	3,647,866	
In SFHA	2,358,552	52%	1,774,642	49%
Outside SFHA	2,149,786	48%	1,873,224	51%
Single family homes	4,063,009	90%	3,480,076	95%
2-4 family homes	218,432	5%	102,507	3%
Other family home	226,897	5%	65,283	2%
California	285,947	6%	266,793	7%
Florida	1,389,670	31%	1,084,605	30%
Louisiana	461,382	10%	391,264	11%
South Carolina	162,952	4%	109,032	3%
Texas	621,650	14%	550,286	15%

SOURCE: FEMA analysis of NFIP policyholder data and Census ACS data.

For the majority of sub-categories in Table A.1, the difference is one or two percentage points. The biggest difference is for single family homes at 5 percentage points, and these results suggest that our weighted sample over-estimates single family homes and under-estimates 2-4 family and other family homes. This over-estimation of single family homes is likely due to single family homes being more easily matched to a corresponding ACS sample. In the case of multi-family homes, it is more likely that the identity associated with the policy does not match the identity of the resident. Because our sample has fewer multi-family units than the population, we underestimate the proportion of multi-family residences and overestimate the proportion of single family residences.

Income Differences between Different Groups

FEMA compared incomes across a number of different subgroups (Table A.2). First, FEMA compared incomes between ACS households that are in and out of the SFHA by state. Second, FEMA compared incomes between ACS households with and without flood insurance by state. Third, FEMA compared incomes between ACS households with and without flood insurance that reside in an SFHA by state and households that reside outside an SFHA by state. Additionally, FEMA compared incomes between ACS households with and without a mortgage and between flood risk areas, i.e. coastal and riverine.

Table A.2. ACS Household Groups

Category	Sub-Groups
Homeowner	With mortgage
	Without mortgage
Flood Risk Area	Coastal
	Riverine
Special Flood Hazard Area	Inside
	Outside
Policyholder	Yes
	No

FEMA ran multiple tests to determine if the difference in income between groups was statistically significant. FEMA used the Wilcoxon Rank Sum Test, a non-parametric statistical test, to test for a difference in household income between specified groups at both the national and state level. FEMA chose the Wilcoxon Rank Sum Test because of the small sample sizes of the groups tested and the presence of extreme values, outliers, and a non-normal distribution of income values. We note that states with small sample sizes, which require cautious interpretation. Similarly, we used median income instead of average income in statistical tests because of outliers and non-normal distribution in the raw data.

After weighting the sample, we estimated the median income for 1.9 million ACS households residing in or out of the SFHA. We then calculated the differences in median incomes between those groups for 52 states or territories regardless of whether the household had flood insurance (Table A.2). At the national level, incomes of households residing in the SFHA were significantly different than income of households outside the SFHA ($W = 22.642$, $p < 2.2e\text{-}16$, n=1.9million). We identified nine states where the weighted household median income was greater inside an SFHA than outside and 43 states where weighted household median income was greater outside an SFHA. Further, statistical tests determined that 16 of the 52 states did not have a statistically significant difference in weighted income between those in and out of the SFHA. For these 16 states, despite having a difference in median income between households inside and outside the SFHA, the difference in income between groups was not statistically significant at a 95 percent confidence level (Table A.3).

Table A.3. Wilcoxon Rank Sum Tests Results, Difference in Income for Households Residing Inside versus Outside an SFHA (2015), using 95% Confidence Level

State	SFHA	Number of households (weighted)	Household Median Income (weighted)	Household Median Income Difference	W	p-values	Statistically Significant
Alabama	in	48,880	$41,334	$4,623	2.17	0.03	yes

Table A.3. Wilcoxon Rank Sum Tests Results, Difference in Income for Households Residing Inside versus Outside an SFHA (2015), using 95% Confidence Level

State	SFHA	Number of households (weighted)	Household Median Income (weighted)	Household Median Income Difference	W	p-values	Statistically Significant
	out	1,573,134	$45,957				
Alaska	in	5,694	$97,095	-$20,589	-2.63	0.01	yes
	out	199,823	$76,506				
Arizona	in	54,624	$49,585	$2,765	0.93	0.35	no
	out	2,199,916	$52,350				
Arkansas	in	46,085	$34,246	$9,213	5.51	< .001	yes
	out	979,720	$43,459				
California	in	329,551	$54,243	$10,998	9.72	< .001	yes
	out	11,368,417	$65,241				
Colorado	in	31,050	$50,069	$14,998	5.42	< .001	yes
	out	1,891,845	$65,067				
Connecticut	in	45,231	$70,042	$5,041	0.32	0.75	no
	out	1,168,545	$75,083				
Delaware	in	13,607	$69,142	-$7,070	-1.75	0.08	no
	out	319,079	$62,072				
District of Columbia	in	4,036	$46,789	$28,330	1.32	0.19	no
	out	245,977	$75,119				
Florida	in	1,593,892	$53,396	-$3,338	-9.53	< .001	yes
	out	5,021,391	$50,058				
Georgia	in	86,814	$52,069	$212	-0.11	0.91	no
	out	3,089,037	$52,281				
Hawaii	in	30,011	$55,075	$21,355	4.75	< .001	yes
	out	353,425	$76,430				
Idaho	in	17,849	$48,855	$126	-0.92	0.36	no
	out	511,049	$48,981				
Illinois	in	76,366	$52,090	$8,014	6.35	< .001	yes
	out	4,368,360	$60,104				
Indiana	in	60,107	$45,046	$6,129	3.63	< .001	yes
	out	2,288,900	$51,175				
Iowa	in	23,476	$49,434	$5,677	2.30	0.02	yes
	out	1,075,133	$55,111				
Kansas	in	26,725	$45,071	$10,002	4.96	< .001	yes
	out	984,245	$55,073				
Kentucky	in	60,474	$35,756	$10,361	6.16	< .001	yes

Table A.3. Wilcoxon Rank Sum Tests Results, Difference in Income for Households Residing Inside versus Outside an SFHA (2015), using 95% Confidence Level

State	SFHA	Number of households (weighted)	Household Median Income (weighted)	Household Median Income Difference	W	p-values	Statistically Significant
	out	1,506,151	$46,117				
Louisiana	in	462,149	$48,064	-$1,594	-2.55	0.01	yes
	out	1,099,310	$46,470				
Maine	in	9,832	$60,906	-$8,825	-1.97	0.05	yes
	out	499,616	$52,081				
Maryland	in	35,009	$70,096	$6,902	1.34	0.18	no
	out	1,951,314	$76,997				
Massachusetts	in	63,457	$62,773	$9,420	4.09	< .001	yes
	out	2,276,794	$72,193				
Michigan	in	60,817	$48,885	$3,187	1.79	0.07	no
	out	3,339,435	$52,072				
Minnesota	in	22,608	$62,525	$2,514	0.49	0.62	no
	out	1,970,646	$65,039				
Mississippi	in	61,593	$37,049	$4,996	2.54	0.01	yes
	out	927,354	$42,045				
Missouri	in	45,511	$38,039	$13,314	8.28	< .001	yes
	out	2,126,079	$51,353				
Montana	in	10,669	$43,737	$6,329	1.18	0.24	no
	out	353,459	$50,066				
Nebraska	in	21,928	$44,988	$10,802	4.30	< .001	yes
	out	657,148	$55,789				
Nevada	in	18,833	$54,693	-$297	0.12	0.91	no
	out	939,988	$54,397				
New Hampshire	in	11,344	$61,152	$10,212	1.50	0.13	no
	out	474,134	$71,364				
New Jersey	in	181,439	$70,085	$5,009	2.87	< .001	yes
	out	2,698,667	$75,093				
New Mexico	in	44,679	$40,045	$7,005	4.15	< .001	yes
	out	651,141	$47,049				
New York	in	194,970	$63,410	-$1,333	-1.38	0.17	no
	out	6,458,261	$62,077				
North Carolina	in	92,977	$47,003	$1,457	0.84	0.40	no
	out	3,321,140	$48,460				
North Dakota	in	8,863	$70,735	-$11,839	-2.09	0.04	yes

Table A.3. Wilcoxon Rank Sum Tests Results, Difference in Income for Households Residing Inside versus Outside an SFHA (2015), using 95% Confidence Level

State	SFHA	Number of households (weighted)	Household Median Income (weighted)	Household Median Income Difference	W	p-values	Statistically Significant
	out	251,556	$58,896				
Ohio	in	75,622	$43,244	$8,817	5.51	< .001	yes
	out	4,273,072	$52,061				
Oklahoma	in	32,636	$43,979	$5,087	1.99	0.05	yes
	out	1,262,413	$49,066				
Oregon	in	43,132	$49,047	$6,785	3.64	< .001	yes
	out	1,386,760	$55,832				
Pennsylvania	in	99,252	$46,053	$10,672	7.76	< .001	yes
	out	4,486,578	$56,725				
Puerto Rico	in	100,639	$20,023	$890	1.52	0.13	no
	out	949,754	$20,913				
Rhode Island	in	11,896	$59,738	$1,581	-0.65	0.52	no
	out	360,235	$61,318				
South Carolina	in	111,177	$68,226	-$20,986	-12.60	< .001	yes
	out	1,540,051	$47,240				
South Dakota	in	13,437	$37,698	$17,134	6.21	< .001	yes
	out	285,043	$54,832				
Tennessee	in	37,681	$43,418	$4,648	1.98	0.05	yes
	out	2,240,759	$48,066				
Texas	in	429,168	$47,399	$9,287	10.27	< .001	yes
	out	7,961,226	$56,686				
Utah	in	8,022	$56,289	$7,800	2.00	0.05	yes
	out	842,804	$64,089				
Vermont	in	7,412	$49,033	$9,445	2.25	0.02	yes
	out	232,255	$58,479				
Virginia	in	87,411	$56,376	$11,246	5.51	< .001	yes
	out	2,811,390	$67,622				
Washington	in	51,164	$49,059	$14,918	8.06	< .001	yes
	out	2,284,761	$63,977				
West Virginia	in	55,002	$38,613	$4,952	4.49	< .001	yes
	out	629,752	$43,565				
Wisconsin	in	27,969	$51,071	$5,503	2.92	< .001	yes
	out	2,128,382	$56,575				
Wyoming	in	5,064	$50,373	$9,715	2.40	0.02	yes

Table A.3. Wilcoxon Rank Sum Tests Results, Difference in Income for Households Residing Inside versus Outside an SFHA (2015), using 95% Confidence Level

State	SFHA	Number of households (weighted)	Household Median Income (weighted)	Household Median Income Difference	W	p-values	Statistically Significant
	out	200,161	$60,088				

SOURCE: FEMA analysis of NFIP policyholder data and Census ACS data.
Note: Data weighted using ACS sample weights.

Using the Wilcoxon Ranked Sum test at a 95 percent confidence level, we selected only those households who reside inside the SFHA, and tested the difference in weighted household income between policyholders and non-policyholders. Our test results determined that the difference between incomes were statistically significant for 46 of the 50 states tested. For households residing in the SFHA there is a true difference in household income between policyholders and non-policyholders: policyholders have a higher income than non-policyholders Test results did not show a significant difference in income between policyholders and non-policyholders residing in the SFHA in Montana, Puerto Rico, Vermont, and Wyoming (Table A.4).

Table A.4. Wilcoxon Rank Sum Tests Results, Difference in Income between Policyholders and Non-policyholders within an SFHA (2015), using 95% Confidence Level

State	SFHA	Number of households (weighted)	Household Median Income (weighted)	Household Median Income Difference	W	p-values	Statistically Significant
Alabama	non-policyholder	33,921	$32,342	$31,710	7.53	< .001	yes
	policyholder	14,959	$64,051				
Alaska	non-policyholder	4,349	$83,912	$35,934	3.12	< .001	yes
	policyholder	1,345	$119,846				
Arizona	non-policyholder	36,963	$39,216	$40,889	10.32	< .001	yes
	policyholder	17,661	$80,105				
Arkansas	non-policyholder	36,092	$32,039	$14,601	2.72	0.01	yes
	policyholder	9,993	$46,640				
California	non-policyholder	233,614	$46,491	$36,051	16.53	< .001	yes
	policyholder	95,937	$82,542				
Colorado	non-policyholder	23,366	$45,056	$23,555	4.70	< .001	yes
	policyholder	7,684	$68,611				
Connecticut	non-policyholder	29,204	$56,469	$66,702	9.06	< .001	yes
	policyholder	16,027	$123,171				
Delaware	non-policyholder	7,337	$55,632	$23,739	2.36	0.02	yes
	policyholder	6,270	$79,371				
	non-policyholder	3,561	$45,045	$208,012	*__	*__	*__

Table A.4. Wilcoxon Rank Sum Tests Results, Difference in Income between Policyholders and Non-policyholders within an SFHA (2015), using 95% Confidence Level

State	SFHA	Number of households (weighted)	Household Median Income (weighted)	Household Median Income Difference	W	p-values	Statistically Significant
District of Columbia	policyholder	475	$253,057				
Florida	non-policyholder	976,673	$43,462	$32,650	39.44	< .001	yes
	policyholder	617,219	$76,112				
Georgia	non-policyholder	49,829	$40,048	$35,060	11.31	< .001	yes
	policyholder	36,985	$75,109				
Hawaii	non-policyholder	25,605	$55,056	$36,806	3.00	< .001	yes
	policyholder	4,406	$91,862				
Idaho	non-policyholder	15,377	$47,499	$37,591	2.32	0.02	yes
	policyholder	2,472	$85,089				
Illinois	non-policyholder	54,561	$45,060	$25,028	8.16	< .001	yes
	policyholder	21,805	$70,088				
Indiana	non-policyholder	44,147	$41,031	$17,002	3.60	< .001	yes
	policyholder	15,960	$58,033				
Iowa	non-policyholder	17,452	$42,423	$18,054	3.81	< .001	yes
	policyholder	6,024	$60,477				
Kansas	non-policyholder	21,074	$40,764	$23,307	5.28	< .001	yes
	policyholder	5,651	$64,071				
Kentucky	non-policyholder	46,341	$30,044	$33,043	9.46	< .001	yes
	policyholder	14,133	$63,087				
Louisiana	non-policyholder	239,954	$32,769	$39,945	28.83	< .001	yes
	policyholder	222,195	$72,714				
Maine	non-policyholder	7,514	$56,046	$20,148	2.28	0.02	yes
	policyholder	2,318	$76,195				
Maryland	non-policyholder	16,080	$53,378	$45,173	7.99	< .001	yes
	policyholder	18,929	$98,551				
Massachusetts	non-policyholder	44,558	$53,314	$32,225	6.39	< .001	yes
	policyholder	18,899	$85,539				
Michigan	non-policyholder	46,720	$42,779	$29,830	8.75	< .001	yes
	policyholder	14,097	$72,609				
Minnesota	non-policyholder	17,540	$56,100	$20,348	4.01	< .001	yes
	policyholder	5,068	$76,448				
Mississippi	non-policyholder	42,945	$28,954	$34,903	9.59	< .001	yes
	policyholder	18,648	$63,858				
Missouri	non-policyholder	33,945	$34,643	$15,417	4.81	< .001	yes
	policyholder	11,566	$50,060				

Table A.4. Wilcoxon Rank Sum Tests Results, Difference in Income between Policyholders and Non-policyholders within an SFHA (2015), using 95% Confidence Level

State	SFHA	Number of households (weighted)	Household Median Income (weighted)	Household Median Income Difference	W	p-values	Statistically Significant
Montana	non-policyholder	8,635	$40,325	$17,344	1.45	0.15	no
	policyholder	2,034	$57,669				
Nebraska	non-policyholder	17,408	$37,109	$30,982	6.55	< .001	yes
	policyholder	4,520	$68,091				
Nevada	non-policyholder	12,076	$46,057	$29,210	3.75	< .001	yes
	policyholder	6,757	$75,268				
New Hampshire	non-policyholder	9,027	$50,468	$34,320	3.30	< .001	yes
	policyholder	2,317	$84,789				
New Jersey	non-policyholder	107,170	$56,444	$30,070	11.17	< .001	yes
	policyholder	74,269	$86,514				
New Mexico	non-policyholder	35,365	$36,878	$21,219	5.65	< .001	yes
	policyholder	9,314	$58,097				
New York	non-policyholder	132,977	$48,126	$53,924	18.00	< .001	yes
	policyholder	61,993	$102,050				
North Carolina	non-policyholder	53,273	$34,404	$38,648	13.72	< .001	yes
	policyholder	39,704	$73,051				
North Dakota	non-policyholder	6,595	$56,638	$45,397	4.56	< .001	yes
	policyholder	2,268	$102,035				
Ohio	non-policyholder	54,206	$38,007	$30,071	10.07	< .001	yes
	policyholder	21,416	$68,078				
Oklahoma	non-policyholder	25,387	$38,786	$22,275	4.54	< .001	yes
	policyholder	7,249	$61,060				
Oregon	non-policyholder	29,926	$39,492	$34,184	6.78	< .001	yes
	policyholder	13,206	$73,676				
Pennsylvania	non-policyholder	70,025	$39,264	$31,855	13.62	< .001	yes
	policyholder	29,227	$71,119				
Puerto Rico	non-policyholder	99,506	$19,848	$7,355	1.91	0.06	no
	policyholder	1,133	$27,203				
Rhode Island	non-policyholder	7,002	$49,614	$34,819	2.81	0.01	yes
	policyholder	4,894	$84,433				
South Carolina	non-policyholder	47,328	$43,051	$50,020	14.17	< .001	yes
	policyholder	63,849	$93,071				
South Dakota	non-policyholder	11,844	$35,047	$15,491	4.02	< .001	yes
	policyholder	1,593	$50,538				
Tennessee	non-policyholder	25,737	$34,215	$28,239	5.20	< .001	yes
	policyholder	11,944	$62,454				

Table A.4. Wilcoxon Rank Sum Tests Results, Difference in Income between Policyholders and Non-policyholders within an SFHA (2015), using 95% Confidence Level

State	SFHA	Number of households (weighted)	Household Median Income (weighted)	Household Median Income Difference	W	p-values	Statistically Significant
Texas	non-policyholder	295,351	$36,056	$46,128	26.90	< .001	yes
	policyholder	133,817	$82,184				
Utah	non-policyholder	7,258	$53,253	$38,575	*__	*__	*__
	policyholder	764	$91,829				
Vermont	non-policyholder	5,635	$47,041	$6,353	1.13	0.26	no
	policyholder	1,777	$53,394				
Virginia	non-policyholder	47,192	$41,734	$38,362	11.10	< .001	yes
	policyholder	40,219	$80,096				
Washington	non-policyholder	34,259	$42,396	$21,249	6.01	< .001	yes
	policyholder	16,905	$63,645				
West Virginia	non-policyholder	46,123	$36,056	$21,677	6.77	< .001	yes
	policyholder	8,879	$57,733				
Wisconsin	non-policyholder	21,269	$45,330	$26,306	5.40	< .001	yes
	policyholder	6,700	$71,636				
Wyoming	non-policyholder	3,896	$49,017	$9,432	0.34	0.74	no
	policyholder	1,168	$58,450				

SOURCE: FEMA analysis of NFIP policyholder data and Census ACS data.

Note: Data weighted using ACS sample weights.

*__ Statistical test not performed due to small sample sizes in unweighted values.

Using the Wilcoxon Ranked Sum test at a 95 percent confidence level, we selected only those households who reside outside the SFHA, and tested the difference in household income between policyholders and non-policyholders. Our test results determined that the difference between incomes were statistically significant for 48 of the 50 states tested. For those households residing outside the SFHA there is a difference in household income between policyholders and non-policyholders: policyholders have a higher income than non-policyholders. Test results did not show a significant difference in income between policyholders and non-policyholders residing outside the SFHA in Delaware and Idaho (Table A.5).

Table A.5. Wilcoxon Rank Sum Tests Results, Difference in Income between Policyholders and Non-policyholders Outside an SFHA (2015), using 95% Confidence Level

State	SFHA	Number of households (weighted)	Household Median Income (weighted)	Household Median Income Difference	W	p-values	Statistically Significant
Alabama	non-policyholder	1,557,577	$45,557	$31,035	8.58	< .001	yes
	policyholder	15,557	$76,592				
Alaska	non-policyholder	198,590	$76,223	$49,424	3.78	< .001	yes
	policyholder	1,233	$125,647				

Table A.5. Wilcoxon Rank Sum Tests Results, Difference in Income between Policyholders and Non-policyholders Outside an SFHA (2015), using 95% Confidence Level

State	SFHA	Number of households (weighted)	Household Median Income (weighted)	Household Median Income Difference	W	p-values	Statistically Significant
Arizona	non-policyholder	2,186,293	$52,183	$16,947	4.59	< .001	yes
	policyholder	13,623	$69,130				
Arkansas	non-policyholder	975,529	$43,311	$29,510	4.23	< .001	yes
	policyholder	4,191	$72,821				
California	non-policyholder	11,197,561	$65,086	$31,369	22.55	< .001	yes
	policyholder	170,856	$96,455				
Colorado	non-policyholder	1,878,634	$64,742	$41,172	10.64	< .001	yes
	policyholder	13,211	$105,914				
Connecticut	non-policyholder	1,154,593	$74,329	$50,450	8.12	< .001	yes
	policyholder	13,952	$124,779				
Delaware	non-policyholder	314,176	$61,708	$14,958	1.83	0.07	no
	policyholder	4,903	$76,666				
District of Columbia	non-policyholder	244,488	$75,110	$38,885	4.30	< .001	yes
	policyholder	1,489	$113,995				
Florida	non-policyholder	4,554,005	$47,571	$33,487	43.66	< .001	yes
	policyholder	467,386	$81,058				
Georgia	non-policyholder	3,053,771	$52,076	$30,049	10.74	< .001	yes
	policyholder	35,266	$82,125				
Hawaii	non-policyholder	349,520	$76,103	$30,770	4.44	< .001	yes
	policyholder	3,905	$106,873				
Idaho	non-policyholder	508,765	$48,962	$21,740	1.81	0.07	no
	policyholder	2,284	$70,702				
Illinois	non-policyholder	4,351,949	$60,095	$35,233	5.80	< .001	yes
	policyholder	16,411	$95,328				
Indiana	non-policyholder	2,277,896	$51,075	$17,381	7.00	< .001	yes
	policyholder	11,004	$68,456				
Iowa	non-policyholder	1,069,268	$55,081	$13,190	4.33	< .001	yes
	policyholder	5,865	$68,271				
Kansas	non-policyholder	978,901	$55,065	$13,035	2.88	< .001	yes
	policyholder	5,344	$68,100				
Kentucky	non-policyholder	1,499,978	$46,072	$19,010	4.18	< .001	yes
	policyholder	6,173	$65,082				
Louisiana	non-policyholder	930,241	$42,065	$30,208	22.77	< .001	yes
	policyholder	169,069	$72,273				
Maine	non-policyholder	497,645	$52,076	$37,907	5.35	< .001	Yes
	policyholder	1,971	$89,983				

Table A.5. Wilcoxon Rank Sum Tests Results, Difference in Income between Policyholders and Non-policyholders Outside an SFHA (2015), using 95% Confidence Level

State	SFHA	Number of households (weighted)	Household Median Income (weighted)	Household Median Income Difference	W	p-values	Statistically Significant
Maryland	non-policyholder	1,933,119	$76,601	$35,545	6.34	< .001	yes
	policyholder	18,195	$112,146				
Massachusetts	non-policyholder	2,256,403	$72,100	$34,380	6.21	< .001	yes
	policyholder	20,391	$106,480				
Michigan	non-policyholder	3,330,191	$52,069	$36,091	9.51	< .001	yes
	policyholder	9,244	$88,160				
Minnesota	non-policyholder	1,963,613	$64,971	$30,951	7.43	< .001	yes
	policyholder	7,033	$95,922				
Mississippi	non-policyholder	890,346	$41,053	$17,049	8.08	< .001	yes
	policyholder	37,008	$58,103				
Missouri	non-policyholder	2,119,720	$51,199	$21,974	5.69	< .001	yes
	policyholder	6,359	$73,173				
Montana	non-policyholder	350,174	$50,061	$17,408	3.45	< .001	yes
	policyholder	3,285	$67,469				
Nebraska	non-policyholder	654,360	$55,547	$31,313	2.25	0.02	yes
	policyholder	2,788	$86,860				
Nevada	non-policyholder	933,714	$54,108	$24,010	4.34	< .001	yes
	policyholder	6,274	$78,118				
New Hampshire	non-policyholder	471,676	$71,101	$23,716	2.99	< .001	yes
	policyholder	2,458	$94,817				
New Jersey	non-policyholder	2,665,030	$74,926	$41,244	10.95	< .001	yes
	policyholder	33,637	$116,170				
New Mexico	non-policyholder	645,419	$46,322	$45,946	8.41	< .001	yes
	policyholder	5,722	$92,268				
New York	non-policyholder	6,369,139	$61,431	$46,752	23.78	< .001	yes
	policyholder	89,122	$108,182				
North Carolina	non-policyholder	3,286,309	$48,075	$27,082	12.53	< .001	yes
	policyholder	34,831	$75,157				
North Dakota	non-policyholder	240,763	$57,090	$37,366	6.28	< .001	yes
	policyholder	10,793	$94,456				
Ohio	non-policyholder	4,255,819	$52,031	$33,069	10.67	< .001	yes
	policyholder	17,253	$85,100				
Oklahoma	non-policyholder	1,253,216	$49,056	$27,253	6.51	< .001	yes
	policyholder	9,197	$76,309				
Oregon	non-policyholder	1,377,267	$55,565	$18,412	4.09	< .001	yes
	policyholder	9,493	$73,977				

Table A.5. Wilcoxon Rank Sum Tests Results, Difference in Income between Policyholders and Non-policyholders Outside an SFHA (2015), using 95% Confidence Level

State	SFHA	Number of households (weighted)	Household Median Income (weighted)	Household Median Income Difference	W	p-values	Statistically Significant
Pennsylvania	non-policyholder	4,460,871	$56,536	$18,162	8.36	< .001	yes
	policyholder	25,707	$74,698				
Puerto Rico	non-policyholder	949,549	$20,910	$50,536	*__	*__	*__
	policyholder	205	$71,446				
Rhode Island	non-policyholder	354,151	$60,996	$31,549	4.98	< .001	yes
	policyholder	6,084	$92,545				
South Carolina	non-policyholder	1,494,868	$46,244	$32,880	16.49	< .001	yes
	policyholder	45,183	$79,124				
South Dakota	non-policyholder	283,102	$54,687	$19,689	3.46	< .001	yes
	policyholder	1,941	$74,377				
Tennessee	non-policyholder	2,226,501	$48,060	$34,401	8.32	< .001	yes
	policyholder	14,258	$82,461				
Texas	non-policyholder	7,544,757	$55,070	$44,990	42.25	< .001	yes
	policyholder	416,469	$100,060				
Utah	non-policyholder	840,791	$64,086	$29,395	3.05	< .001	yes
	policyholder	2,013	$93,481				
Vermont	non-policyholder	231,329	$58,319	$29,905	4.66	< .001	yes
	policyholder	926	$88,224				
Virginia	non-policyholder	2,762,778	$67,077	$41,046	16.44	< .001	yes
	policyholder	48,612	$108,123				
Washington	non-policyholder	2,270,236	$63,732	$27,002	5.31	< .001	yes
	policyholder	14,525	$90,734				
West Virginia	non-policyholder	627,117	$43,503	$34,803	3.23	< .001	yes
	policyholder	2,635	$78,305				
Wisconsin	non-policyholder	2,121,501	$56,386	$27,522	6.39	< .001	yes
	policyholder	6,881	$83,907				
Wyoming	non-policyholder	199,152	$60,081	$67,106	*__	*__	*__
	policyholder	1,009	$127,187				

SOURCE: FEMA analysis of NFIP policyholder data and Census ACS data.
Note: Data weighted using ACS sample weights.
*-- Statistical test not performed due to small sample sizes in unweighted values.

In addition, policyholders have higher median incomes than non-policyholders at the national level and for 50 out of 52 states tested. Test results did not show a significant difference in income between policyholders and non-policyholders in Vermont and Wyoming (Table A.6).

Table A.6. Wilcoxon Rank Sum Tests Results, Difference in Income between Policyholders and Non-policyholders Outside an SFHA (2015), using 95% Confidence Level

State	SFHA	Number of households (weighted)	Household Median Income (weighted)	Household Median Income Difference	W	p-values	Statistically Significant
Alabama	non-policyholder	1,557,577	$45,557	$31,035	8.58	< .001	yes
	policyholder	15,557	$76,592				
Alaska	non-policyholder	198,590	$76,223	$49,424	3.78	< .001	yes
	policyholder	1,233	$125,647				
Arizona	non-policyholder	2,186,293	$52,183	$16,947	4.59	< .001	yes
	policyholder	13,623	$69,130				
Arkansas	non-policyholder	975,529	$43,311	$29,510	4.23	< .001	yes
	policyholder	4,191	$72,821				
California	non-policyholder	11,197,561	$65,086	$31,369	22.55	< .001	yes
	policyholder	170,856	$96,455				
Colorado	non-policyholder	1,878,634	$64,742	$41,172	10.64	< .001	yes
	policyholder	13,211	$105,914				
Connecticut	non-policyholder	1,154,593	$74,329	$50,450	8.12	< .001	yes
	policyholder	13,952	$124,779				
Delaware	non-policyholder	314,176	$61,708	$14,958	1.83	0.07	no
	policyholder	4,903	$76,666				
District of Columbia	non-policyholder	244,488	$75,110	$38,885	4.30	< .001	yes
	policyholder	1,489	$113,995				
Florida	non-policyholder	4,554,005	$47,571	$33,487	43.66	< .001	yes
	policyholder	467,386	$81,058				
Georgia	non-policyholder	3,053,771	$52,076	$30,049	10.74	< .001	yes
	policyholder	35,266	$82,125				
Hawaii	non-policyholder	349,520	$76,103	$30,770	4.44	< .001	yes
	policyholder	3,905	$106,873				
Idaho	non-policyholder	508,765	$48,962	$21,740	1.81	0.07	no
	policyholder	2,284	$70,702				
Illinois	non-policyholder	4,351,949	$60,095	$35,233	5.80	< .001	yes
	policyholder	16,411	$95,328				
Indiana	non-policyholder	2,277,896	$51,075	$17,381	7.00	< .001	yes
	policyholder	11,004	$68,456				
Iowa	non-policyholder	1,069,268	$55,081	$13,190	4.33	< .001	yes
	policyholder	5,865	$68,271				
Kansas	non-policyholder	978,901	$55,065	$13,035	2.88	< .001	yes
	policyholder	5,344	$68,100				
Kentucky	non-policyholder	1,499,978	$46,072	$19,010	4.18	< .001	yes
	policyholder	6,173	$65,082				
Louisiana	non-policyholder	930,241	$42,065	$30,208	22.77	< .001	yes

Table A.6. Wilcoxon Rank Sum Tests Results, Difference in Income between Policyholders and Non-policyholders Outside an SFHA (2015), using 95% Confidence Level

State	SFHA	Number of households (weighted)	Household Median Income (weighted)	Household Median Income Difference	W	p-values	Statistically Significant
	policyholder	169,069	$72,273				
Maine	non-policyholder	497,645	$52,076	$37,907	5.35	< .001	yes
	policyholder	1,971	$89,983				
Maryland	non-policyholder	1,933,119	$76,601	$35,545	6.34	< .001	yes
	policyholder	18,195	$112,146				
Massachusetts	non-policyholder	2,256,403	$72,100	$34,380	6.21	< .001	yes
	policyholder	20,391	$106,480				
Michigan	non-policyholder	3,330,191	$52,069	$36,091	9.51	< .001	yes
	policyholder	9,244	$88,160				
Minnesota	non-policyholder	1,963,613	$64,971	$30,951	7.43	< .001	yes
	policyholder	7,033	$95,922				
Mississippi	non-policyholder	890,346	$41,053	$17,049	8.08	< .001	yes
	policyholder	37,008	$58,103				
Missouri	non-policyholder	2,119,720	$51,199	$21,974	5.69	< .001	yes
	policyholder	6,359	$73,173				
Montana	non-policyholder	350,174	$50,061	$17,408	3.45	< .001	yes
	policyholder	3,285	$67,469				
Nebraska	non-policyholder	654,360	$55,547	$31,313	2.25	0.02	yes
	policyholder	2,788	$86,860				
Nevada	non-policyholder	933,714	$54,108	$24,010	4.34	< .001	yes
	policyholder	6,274	$78,118				
New Hampshire	non-policyholder	471,676	$71,101	$23,716	2.99	< .001	yes
	policyholder	2,458	$94,817				
New Jersey	non-policyholder	2,665,030	$74,926	$41,244	10.95	< .001	yes
	policyholder	33,637	$116,170				
New Mexico	non-policyholder	645,419	$46,322	$45,946	8.41	< .001	yes
	policyholder	5,722	$92,268				
New York	non-policyholder	6,369,139	$61,431	$46,752	23.78	< .001	yes
	policyholder	89,122	$108,182				
North Carolina	non-policyholder	3,286,309	$48,075	$27,082	12.53	< .001	yes
	policyholder	34,831	$75,157				
North Dakota	non-policyholder	240,763	$57,090	$37,366	6.28	< .001	yes
	policyholder	10,793	$94,456				
Ohio	non-policyholder	4,255,819	$52,031	$33,069	10.67	< .001	yes
	policyholder	17,253	$85,100				
Oklahoma	non-policyholder	1,253,216	$49,056	$27,253	6.51	< .001	yes
	policyholder	9,197	$76,309				

Table A.6. Wilcoxon Rank Sum Tests Results, Difference in Income between Policyholders and Non-policyholders Outside an SFHA (2015), using 95% Confidence Level

State	SFHA	Number of households (weighted)	Household Median Income (weighted)	Household Median Income Difference	W	p-values	Statistically Significant
Oregon	non-policyholder	1,377,267	$55,565	$18,412	4.09	< .001	yes
	policyholder	9,493	$73,977				
Pennsylvania	non-policyholder	4,460,871	$56,536	$18,162	8.36	< .001	yes
	policyholder	25,707	$74,698				
Puerto Rico	non-policyholder	949,549	$20,910	$50,536	*__	*__	*__
	policyholder	205	$71,446				
Rhode Island	non-policyholder	354,151	$60,996	$31,549	4.98	< .001	yes
	policyholder	6,084	$92,545				
South Carolina	non-policyholder	1,494,868	$46,244	$32,880	16.49	< .001	yes
	policyholder	45,183	$79,124				
South Dakota	non-policyholder	283,102	$54,687	$19,689	3.46	< .001	yes
	policyholder	1,941	$74,377				
Tennessee	non-policyholder	2,226,501	$48,060	$34,401	8.32	< .001	yes
	policyholder	14,258	$82,461				
Texas	non-policyholder	7,544,757	$55,070	$44,990	42.25	< .001	yes
	policyholder	416,469	$100,060				
Utah	non-policyholder	840,791	$64,086	$29,395	3.05	< .001	yes
	policyholder	2,013	$93,481				
Vermont	non-policyholder	231,329	$58,319	$29,905	4.66	< .001	yes
	policyholder	926	$88,224				
Virginia	non-policyholder	2,762,778	$67,077	$41,046	16.44	< .001	yes
	policyholder	48,612	$108,123				
Washington	non-policyholder	2,270,236	$63,732	$27,002	5.31	< .001	yes
	policyholder	14,525	$90,734				
West Virginia	non-policyholder	627,117	$43,503	$34,803	3.23	< .001	yes
	policyholder	2,635	$78,305				
Wisconsin	non-policyholder	2,121,501	$56,386	$27,522	6.39	< .001	yes
	policyholder	6,881	$83,907				
Wyoming	non-policyholder	199,152	$60,081	$67,106	*__	*__	*__
	policyholder	1,009	$127,187				

SOURCE: FEMA analysis of NFIP policyholder data and Census ACS data.
Note: Data weighted using ACS sample weights.
*-- Statistical test not performed due to small sample sizes in unweighted values.

Appendix B: Additional Information on Flood Insurance Cost and Household Income

This appendix provides additional detail on several of the findings presented in Chapter 2.

Number of NFIP Policyholders and Cost of NFIP Policies

Table B.1 shows the number of NFIP residential policies by state. This tabulation includes all 4.8 million policies in effect at some point in 2015 before we merged ACS income data with NFIP data.

Table B.1. Number of NFIP Policyholders by State, including residential, non-residential business, and other non-residential (2015)

State	Number of Policyholders	State	Number of Policyholders
Florida	1,461,218	West Virginia	21,094
Texas	657,052	Hawaii	19,429
Louisiana	497,060	Oklahoma	19,178
California	307,436	Wisconsin	17,171
New Jersey	196,657	Iowa	16,917
New York	189,818	New Mexico	16,705
South Carolina	170,357	Rhode Island	15,232
North Carolina	131,753	North Dakota	14,769
Virginia	109,551	Nevada	14,508
Georgia	96,850	Kansas	13,550
Mississippi	76,308	Nebraska	13,198
Pennsylvania	72,680	Minnesota	13,189
Massachusetts	58,358	Puerto Rico	10,243
Maryland	53,061	Maine	10,074
Alabama	47,175	New Hampshire	8,535
Illinois	46,532	Idaho	6,756
Washington	45,129	Montana	6,671
Ohio	43,011	South Dakota	5,731
Connecticut	39,805	Vermont	4,806
Arizona	39,614	Utah	4,203
Tennessee	34,184	Alaska	3,206
Oregon	33,149	Wyoming	2,549
Indiana	31,824	District of Columbia	1,910
Missouri	27,342		
Michigan	26,826	Total	4,848,829
Kentucky	26,445		
Colorado	24,701		
Delaware	22,900		
Arkansas	22,409		

Table B.2 reports the average premium, fees, and total policy cost for single-family homes for policies in effect in 2015. Single-family homes account for 4.06 million of the 4.8 million policies in Table B.1. We added the HFIAA to policies that did not report a HFIAA surcharge. Policyholder costs average $800 overall, and are substantially higher in the SFHA than outside the SFHA. Fees account for about 17 percent of policyholder costs, with little variation inside and outside the SFHA.

Table B.2. Average Premiums and Fees for Single Family Homes (for policies in effect in 2015)

	Inside SFHA	Outside SFHA	Total
Premiums	$913	$405	**$663**
Fees	$185	$88	**$137**
Policyholder costs	$1,098	$492	**$800**
Number of policies	2,062,274	2,000,729	**4,063,003**

SOURCE: FEMA analysis of NFIP policyholder data.
NOTE: Includes premiums and fees.

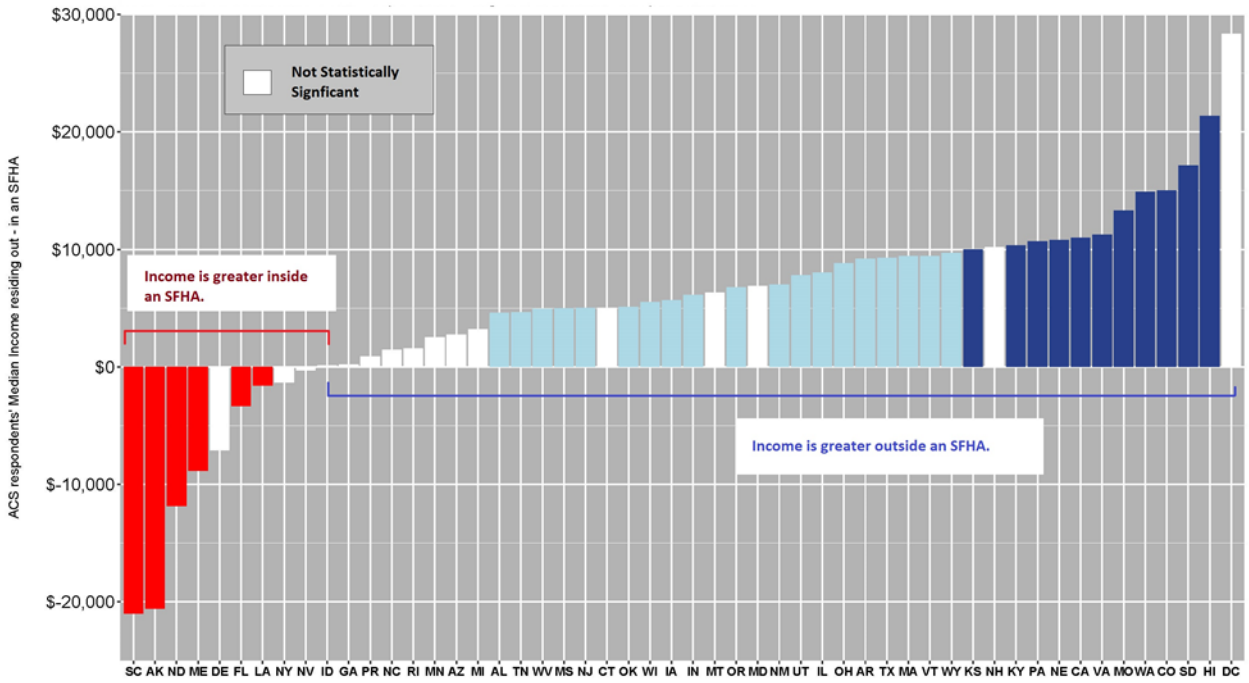
Policyholder and Non-policyholder Income

Figure B.1 shows the difference in median income between households inside and outside the SFHA by state.

- The red boxes represent states where median household income in the SFHA is higher and statistically different from income outside the SFHA.
- The blue boxes represent states where median household income in the SFHA is lower and statistically different from income outside the SFHA.
- The white boxes represent states where there is not a statistically significant difference between median incomes in the two regions.

Overall, households (policyholders and non-policyholders) inside the SFHA have lower median income than households outside the SFHA; however, there is considerable variation by state. For some states, income is higher in the SFHA than outside the SFHA. However, in most states, income is lower in the SFHA than outside the SFHA. Figure B.1 and subsequent figures use the weighted medians from the merged NFIP and ACS data.

Figure B.1. Difference in Median Incomes of ACS Respondents Living In and Out of the SFHA (2014)

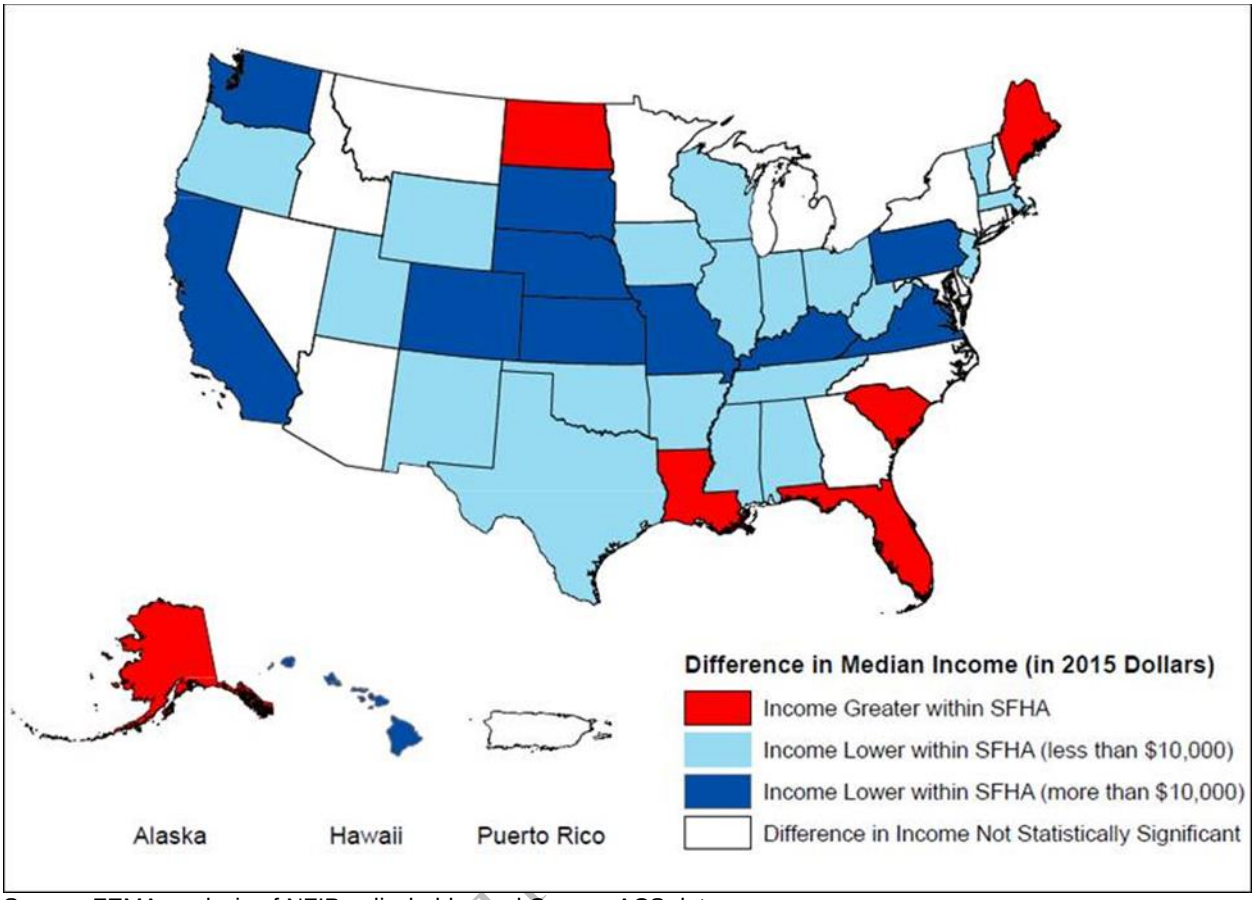


SOURCE: FEMA analysis of U.S. Census Bureau ACS data.

NOTE: Data weighted using ACS sample weights.

In Figure B.2 below, we present the same differences shown in Figure B.1 on a map. Red represents higher median incomes in the SFHA, blue represents lower median income in the SFHA, and white represents non-statistically significant differences.

Figure B.2. Map of Difference in Median Incomes of ACS Respondents Living In or Out of the SFHA (2015)

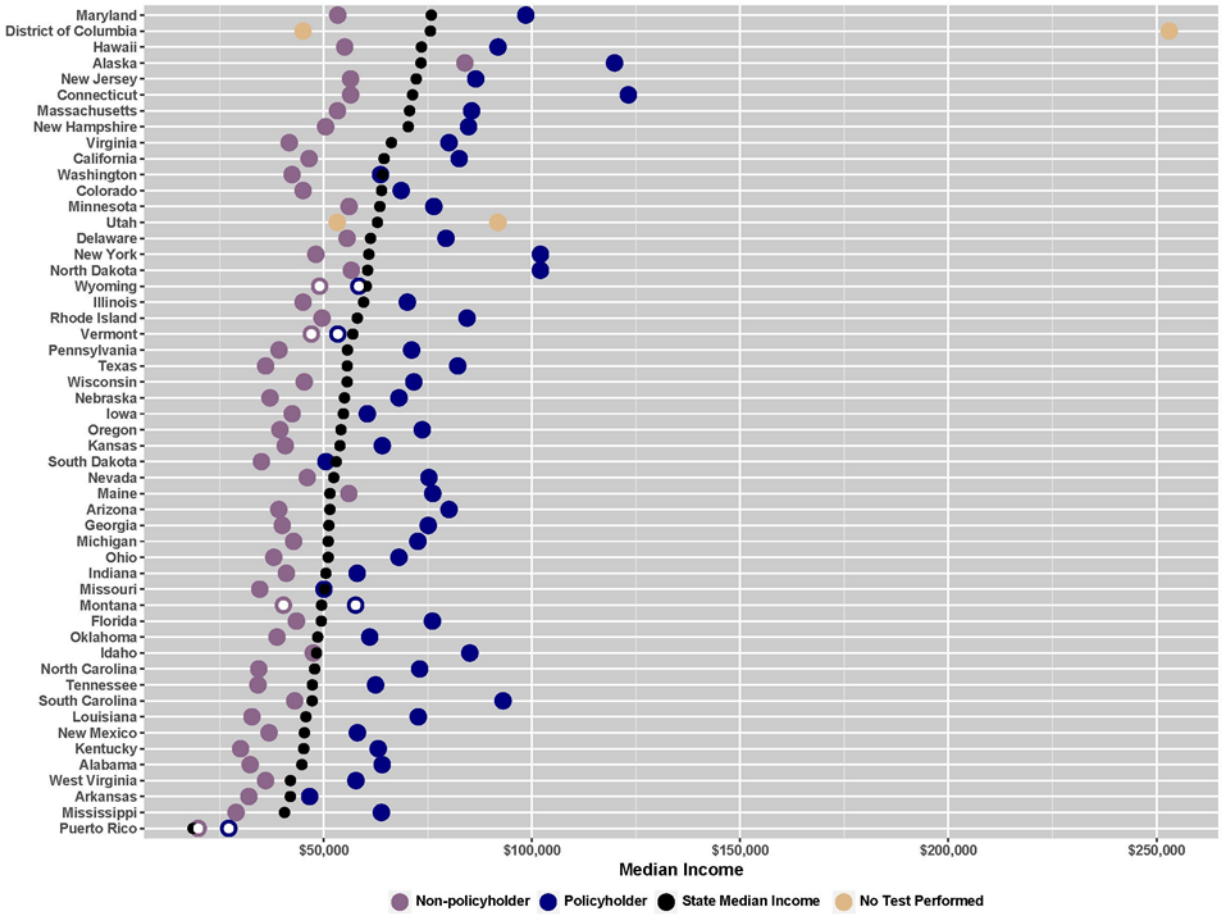


Source: FEMA analysis of NFIP policyholder and Census ACS data

Figures B.3 and B.4 show the difference in policyholder and non-policyholder income by state inside and outside the SFHA, respectively. The dark blue dots represent median policyholder income, purple dots represent non-policyholder median income, and black dots represent median income of all households in the states regardless of whether they are policyholders and whether they live inside an SFHA). Colorless dots represent states where there is no statistically significant difference between median incomes.[49] Policyholder income is higher than non-policyholder income in all states, but the magnitude of the difference varies.

[49]We did not test several states statistical significance for several states because of small sample sizes.

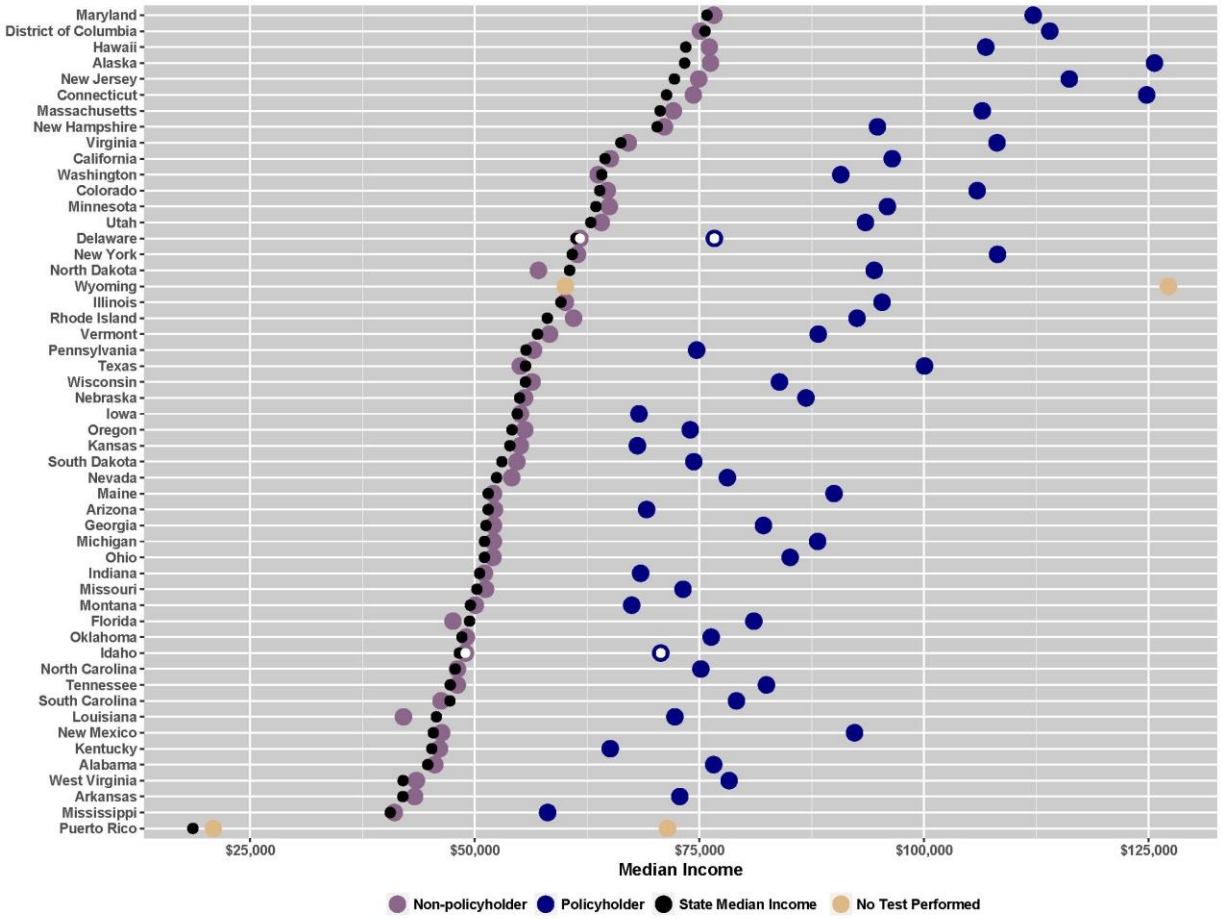
Figure B.3. Household Incomes for ACS Respondents Living inside the SFHA (2015, includes both policyholders and non-policyholders)



Note: We did not perform tests with unwieghted sample size below 15 households. Data points with white circles indicate that the difference in income is not statistically significant. State Median Incomes were reported from Census. Policyholder and Non-policyholder median values are weighted.

SOURCE: FEMA analysis of NFIP Policyholder and Census ACS data.
NOTE: Data weighted using ACS sample weights.

Figure B.4. Household Income for ACS Respondents Living outside the SFHA (2015, includes both policyholders and non-policyholders)



Note: We did not perform tests with unwieghted sample size below 15 households. Data points with white circles indicate that the difference in income is not statistically significant. State Median Incomes were reported from Census. Policyholder and Non-policyholder median values are weighted.

SOURCE: FEMA analysis of NFIP Policyholder and ACS Census data.

NOTE: Data weighted using ACS sample weights.

Appendix C: Methods Used to Estimate Program Costs

In this appendix, we discuss the methods we used to develop the costs for the program design options discussed in Chapter 4.

For each design option, we estimate the number of participants and annual program cost for different sets of program scenarios. FEMA estimated the cost and number of participants for current policyholders. We did not estimate the number of non-policyholders who might become policyholders and eligible for the program as a result of an affordability program. FEMA calculated the number of participants in each scenario by summing the household weights for those eligible to participate in the program. FEMA calculated the cost of each design option by multiplying the household weight variable by the benefit (amount of premium reduction) for each participant and then summed the weighted values.
Below we provide the formulas used to calculate benefit for each of the program designs.

Cost of Income-Based Premium Sharing

Cost When Benefits Vary Continuously with Household Income

This program is designed around the notion that there is cost sharing between the household and the Federal Government (the benefit program). This implies that we need to know the unsubsidized premium, P that each household would face given the location of their primary residence to begin to calculate the cost of the program. There are three parameters in Design 1: the benefit that low-income households receive expressed as a fraction β, the income cut-off that defines low-income households $\underline{I}$, and the income cut-off for eligibility to the program $\bar{I}$. Incomes over $\bar{I}$ would receive no benefit. For each household eligible for the program, i.e., those with incomes below the income eligibility cut-off, we calculate the benefit that the household would receive. The formula for calculating the benefit that the household would receive is:

$$Benefit = \begin{cases} \beta P & ,if\ Income \le \underline{I} \\ \beta P \dfrac{Income - \underline{I}}{\bar{I} - \underline{I}} & if\ \underline{I} \le Income \le \bar{I} \end{cases}$$

To calculate the total cost of the program, we simply add the benefits that each household receives. Additionally, since not all eligible households will either purchase flood insurance or participate in the program, we would assume a take-up rate that would be a function of income. As β increases, every household receives a larger benefit and makes the program more costly. As $\underline{I}$ increases, more households are eligible for the full benefit, making the program more expensive. As $\bar{I}$ increases, more households are eligible for the program and the cost of the program increases but at a smaller rate than by increasing $\underline{I}$.

Cost When Benefits Based on Income Bins

With a binned approach, we have many more parameters to estimate. For each bin, we would have upper- and lower-income thresholds but the upper-income threshold for one bin is the

lower-income threshold for the next. Suppose that there are three income bins that are eligible for benefits with income cutoffs: $I_1, I_2,$ and I_3 that receive proportions $\beta_1, \beta_2,$ and β_3 from the program. The benefit that a household would receive is:

$$Benefit = \begin{cases} \beta_1 P & ,if\ Income \le I_1 \\ \beta_2 P & ,if\ I_1 < Income \le I_2 \\ \beta_3 P & ,if\ I_2 < Income \le I_3 \end{cases}$$

To estimate the total cost of the program, we add the benefits paid for each income bin and add up the total cost of each income bin.

The complication that arises when estimating the cost of this design option, whether continuous or binned, is that the income cutoffs are place-specific, because they are based on AMI. Thus, AMIs for all potential policyholders would need to be calculated.

Cost of Design 2: Premium Burden-Based Benefit

Cost When Benefits Vary Continuously with Household Income

This design option is very similar to the previous option, but instead of calculating the benefit as a proportion of the premium, P households must pay a portion of their income toward flood insurance. The benefit would never exceed the price of a policy. For this design, we assume that there are three income cutoffs, $I_1, I_2,$ and I_3, and three corresponding income burdens $\beta_1, \beta_2,$ and β_3. The benefit that the household receives is:

$$Benefit = \begin{cases} P - \beta_1 Income & if\ Income \le I_1 \\ P - \beta_2 Income & if I_1 < Income \le I_2 \\ P - \beta_3 Income & if\ I_2 < Income \le I_3 \end{cases}$$

To calculate the total cost of the program, we add the benefits received by each eligible household.

Cost When Benefits Based on Income Bins

To transfer Option 2 to a binned approach would require an additional simplification. In particular, we would assume a particular income and pay a fixed proportion amount of the premium. This design would be:

$$Benefit = \begin{cases} P & if\ Income \le I_1 \\ P - \beta_2 I_2 & if I_1 < Income \le I_2 \\ P - \beta_3 I_3 & if\ I_2 < Income \le I_3 \end{cases}$$

Cost of Design 3: Housing Burden-Based Benefit

To calculate the benefit in this design we must know the household's PITI with flood insurance $PITI_w$ and without flood insurance, $PITI_{wo}$ and the flood insurance premium, P. There are only two parameters to this design: the income cutoff for eligibility, $\bar{I}$ and the ratio of PITI to

household income above which a household is housing burdened (α). We determined the PITI benchmark for each household, $\overline{PITI}$, by multiplying α (e.g., 0.4) times household income. If $PITI_w$ is less than $\overline{PITI}$, the household is not eligible for the program. If $PITI_{wo}$ is greater than $\overline{PITI}$, then the household receives a benefit equal to the full amount of the premium. In the intermediate case when the PITI without flood insurance is less than $\overline{PITI}$ and PITI with flood insurance is greater than $\overline{PITI}$, the benefit is:

$$Benefit = PITI_w - \overline{PITI}$$

We can adjust the cost of the program by changing $\bar{I}$ and α.

Cost Design 4: Mitigation Grants and Loans Add-On

It is very difficult to estimate the cost of a mitigation grant and loan program as the number of eligible properties is difficult to estimate without identifying the properties and computing the cost of mitigation itself. The Small Business Administration (SBA) reported during the NAS workshops conducted for this study that a very small percentage of eligible households participated in their mitigation loan program. If a flood insurance mitigation grant or loan program had similar utilization rate, the potential costs of the program could be modest.

Appendix D: The Person Identification Validation System (PVS): Applying the Center for Administrative Records Research and Applications' (CARRA) Record Linkage Software

CARRA Working Paper Series

Working Paper #2014-01

The Person Identification Validation System (PVS): Applying the Center for Administrative Records Research and Applications' (CARRA) Record Linkage Software

Deborah Wagner
U.S. Census Bureau

Mary Layne
U.S. Census Bureau

Center for Administrative Records Research and Applications
U.S. Census Bureau
Washington, D.C. 2023

Report Issued: July 1, 2014

The Person Identification Validation System (PVS): Applying the Center for Administrative Records Research and Applications' (CARRA) Record Linkage Software

July 1, 2014

Deborah Wagner
U.S. Census Bureau

Mary Layne
U.S. Census Bureau

Abstract

The Census Bureau's Person Identification Validation System (PVS) assigns unique person identifiers to federal, commercial, census, and survey data to facilitate linkages across and within files. PVS uses probabilistic matching to assign a unique Census Bureau identifier for each person. The PVS matches incoming files to reference files created with data from the Social Security Administration (SSA) Numerical Identification file, and SSA data with addresses obtained from federal files. This paper describes the PVS methodology from editing input data to creating the final file.

TABLE OF CONTENTS

I. INTRODUCTION

The Census Bureau performs research with administrative records[1] files to investigate methods to improve the Census Bureau's statistical processes. Further, many projects at the Census involve matching persons across surveys and Federal data to enhance the understanding of participation in various Federal programs. Fundamental to this work is a method to ensure the same person is linked across multiple administrative files. The Census Bureau's Person Identification Validation System (PVS) is used to ascertain unique person and address identifiers.

The PVS uses probabilistic linking (Fellegi and Sunter, 1989) to match person data to a reference file. The reference files are derived from the Social Security Administration (SSA) Numerical Identification file (SSA Numident). The Numident contains all transactions recorded against one Social Security Number (SSN) and is reformatted to create the Census Numident. The Census Numident reference file contains one record for each SSN, keeping all variants of date of birth (DOB) and name data in separate files. The Census Numident is enhanced with address information from administrative records to create another reference file, the GeoBase.

The matched person record is assigned a unique person identifier called the protected identification key (PIK) and it is an anonymous identifier as unique as a SSN. Once assigned, the PIK serves as a person linkage key across all files that have been processed using PVS. The PIK also serves as a person unduplication key within files.

The Census Bureau developed the PVS in 1999 in collaboration with the SSA. The PVS was tested using the 1997 Current Population Survey (CPS) survey previously verified using the SSA's SSN validation process, the Enumeration Verification System (EVS). An independent evaluation led by the Census Bureau's Demographic Surveys Division (DSD) compared the results of the PVS and EVS and showed the PVS achieved match rates more than two percentage points higher than achieved by SSA's EVS. The use of address data in the PVS system proved valuable in achieving this higher SSN assignment rate. As part of our Memorandum of Understanding with SSA, the Census Bureau must validate SSNs using PVS any time data are linked to any SSA administrative data. The SSA has authorized the use of the PVS at the Census Bureau.

[1] As used here, "administrative records" are files from the IRS, Department of Housing and Urban Developments Tenant Rental Assistance Certification System and Public and Indian Housing Information Center (HUD-TRACS and HUD-PIC), Selective Service System (SSS), Indian Health Service (IHS), Medicare Enrollment (MEDB), and commercial data from various sources.

The PVS was also used within the Statistical Administrative Records System (StARS) (Farber and Leggieri, 2002) to un-duplicate person records from Federal administrative records data. The StARS was produced annually from 1999 through 2010, each year processing approximately 900 million person records through the PVS system. PVS has been used with the Census Bureau's demographic surveys and censuses since 2001.

The PVS is the cornerstone of the Census Match Study (O'Hara and Marshall, 2011). Following the 2010 Decennial Census, the Census Match Study was undertaken to compare the coverage of administrative records and commercial data to the 2010 Decennial Census enumeration. In addition to data from Federal administrative records, commercial sources were evaluated. Nine sources of commercial data were selected to provide additional addresses and information for people not found in administrative records as of the Census date of 4/1/2010. These commercial data sources were thought to contain more timely information on address data as of census day to supplement addresses not available on the Federal files.

To meet the challenge of employing the PVS on commercial data sources and new Federal files, The Center for Administrative Records Research & Applications (CARRA) reviewed its current PVS capability. Independent contractors (NORC, 2011a) evaluated the PVS and recommended further enhancements to improve the already sound methodology of PVS.

One of the key enhancements increased the coverage of the reference files by including records for persons with Individual Taxpayer Identification Numbers assigned by the Internal Revenue Service (ITINs) to the SSN-based Numident data. The PVS is an important tool at the Census Bureau as it continues to pursue research using the 2010 Decennial data and plan for administrative records use in the 2020 Decennial Census.

Record linkage requires human input throughout the process. Files need to be properly edited before they can be linked successfully. Parameters have to be set and links examined. In the subsequent sections, we describe the PVS. Section II provides an overview of record linkage and Section III details the methodology for the current PVS system and its uses at the Census Bureau. Section IV presents PVS results for Federal and commercial files. Section V discusses advantages and future improvements of the PVS. Finally, Section VI summarizes the salient points of the paper.

II. RECORD LINKAGE BACKGROUND

Record linkage is the process of bringing together two or more records relating to the same entity. In 1946, H. L. Dunn of the United States National Bureau of Statistics introduced the term in this way: "Each person in the world creates a Book of Life. This Book starts with birth and ends with death. Record linkage is the name of the process of assembling the pages of this Book into a volume" (Dunn, 1946). In 1959, computerized record linkage was first undertaken by the Canadian geneticist Howard Newcombe and his associates.

In 1969, Ivan Fellegi and Alan Sunter provided a mathematical framework to provide a viable theorem for linking two data files on common characteristics. This section briefly describes the general theory of record linkage. The paper by Felligi and Sunter (1969) provides the detailed theory.

Felligi and Sunter describe a *comparison space*, **A** x **B**, consisting of all comparisons of records from two files *a* and *b*. There are three outcomes for the comparison space **A** x **B**: links, possible links, and non-links. In order to classify record pairs, comparisons are made between the same fields in each of the files.

Denote the set of all comparison vectors, **Γ**, in **A** x **B** by:

$$\boldsymbol{\Gamma}[\boldsymbol{A}\ x\ \boldsymbol{B}] = \{\tau^1[(\alpha(a),\beta(b))],\{\tau^2[(\alpha(a),\beta(b))],\cdots,\{\tau^k[(\alpha(a),\beta(b))]\}$$

Two conditional probabilities are computed for each comparison pair.

m, is the probability of agreement for a given comparison, when the record is *in truth* a match. Because all matching variables are subject to data coding error (for example, typographical or scanning errors), this *m* probability is less then 1.0 .

u, is the probability of the comparison agreeing purely by chance for two records not belonging to the same individual.

The ratio of these two conditional probabilities, *R*, is an odds ratio defined as:

$$R = \frac{P(agree|\boldsymbol{m})}{P(agree|\boldsymbol{u})} = \frac{m(\tau)}{u(\tau)}$$

Probabilistic record linkage assigns a value of R, to a pair of records.[2] The analyst provides a lower and upper threshold for a match. The optimum linkage is the one

[2] In practice, the comparison space is far more complicated because matching occurs on multiple fields, e.g., first name, middle name, last name, street address, street name, date of birth, etc.

where records linked have a higher odds ratio than the upper threshold. All unlinked records have a odds ratio lower than the lower threshold. These two thresholds are set based on tolerance levels for the two types of error, linking unmatched records and failing to link matched ones.

If $R \geq Upper$, then the pair r, is a link. If $Lower < R < Upper$, the pair r is a potential match and assigned for analyst review. If $R \leq Lower$, the pair r is a non-match. The cut-off thresholds for *Lower* and *Upper* are determined before the linking is done. It is important to review records falling into each category.

Individual agreement weights, w_τ for the i^{th} fields of the r^{th} pair of records can be used as a substitute for the odds ratio, R:

$$w_\tau = log \frac{P(\tau \in \mathbf{\Gamma} | \boldsymbol{m})}{P(\tau \in \mathbf{\Gamma} | \boldsymbol{u})}$$

The analyst designates weight values for upper and lower cutoffs. Record pairs where weights are above upper cutoff are links, weights below the lower cutoff are non-links, and those weights between cutoffs are in the clerical review region.

Estimating *m* and *u* Probabilities

There are several methods for automatically or semi- automatically estimating the m and u probabilities (Winkler, 2002, Winkler and Yancey, 2006). Often, value specific or frequency-based estimates are employed. Another approach is using the EM algorithm, which involves finding the maximum likelihood estimate of parameters. The likelihood function is simplified by assuming the existence of, and values for, missing or hidden parameters (in this case, m and u) (Bilmes, 1998).

In practice, a single threshold value is often used (Winglee, 2005) and for PVS production purposes, CARRA sets a single threshold. In this case, the analyst decides links made below a threshold value are not valid links. Upper and lower bounds can also be set in PVS, and the system can provide links for analyst review.

III. PVS METHODOLOGY

The PVS software serves two primary functions: standard data editing and probabilistic matching. The PVS provides a documented, practical solution for processing many data files including census surveys (Current Population Study, American Community Survey), administrative records, and commercial files. The PVS probabilistically matches an incoming file to reference files in order to assign an anonymous PIK.

Optimized parameters, which provide information relevant to a probabilistic search,[3] are preset by analysts for each file type based on years of usage, research, and testing. The staff in CARRA have expertise in the parameter setting process from their research and bring this knowledge to production. When a new file arrives, different from all previous files, CARRA staff optimize the parameters for probabilistic linking. Record linkage is both an art and a science. It is a balancing act between link quality, processing speed, and setting optimal parameters.

3.1 Multi-Match Software

The PVS employs its probabilistic record linkage software, Multi-Match (Wagner 2012), as an integral part of the PVS. CARRA's Multi-Match is a generalized probabilistic matching routine at the core of all of CARRA's matching applications, and it is used heavily in the PVS. The Multi-Match is programmed as a SAS macro and is used by various record linkage applications. A user-defined parameter file provides information to Multi-Match, enabling it to perform the following functions

- Block records according to parameter file specifications
- Process records through passes,[4] as defined by the parameter file
- Perform comparisons between records in the incoming file and those in the reference file
- Create linked output.

Multi-Match can link *any* two files on *any* set of characteristics and can be used outside of the PVS construct for other matching applications.

Multi-Match, while an important part of the PVS, is not the only element in the PVS. The PVS performs many additional functions: data editing to prepare fields for matching, assigning Census geographical codes to incoming files, and data housekeeping functions.

3.2 Reference Files

Matching in the PVS requires a reference file to match against. Reference files are based on SSA's Numident and contain SSN, a CARRA assigned PIK associated with the SSN, date of birth, name, gender and any addresses where the person may have resided.

[3] Parameter files are discussed in-depth in Section 3.3.3.

[4] Each pass through the data defines differing matching strategies. See Section 3.3.3 for more details.

The SSA Numident file contains all transactions ever recorded against any single SSN. The Census Bureau builds a Census Numident on a regular basis from personal information derived from the SSA Numident file. All transactions related to a given SSN are resolved to produce a Census Numident file containing one data record for each SSN. All variants of name information for each SSN are retained in the Alternate Name Numident file, while all variants of date of birth data are retained in the Alternate DOB Numident. In addition to the Census Numident, PVS creates three other sets of reference files containing Numident data: the GeoBase Reference File, the Name Reference File, and the DOB Reference file.

The GeoBase Reference File appends addresses from administrative records attached to Numident data, including all possible combinations of alternate names and dates of birth for SSN. Addresses from administrative records are edited and processed through commercial software product to clean and standardize address data. ITIN data is also incorporated into the Geobase.

The Name and DOB Reference files are reformatted versions of the Census Numident and includes all possible combinations of alternate names and dates of birth, as well as ITIN data. All of the reference files contain SSN/ITIN and the corresponding PIK. When an input record is linked to a reference file, the corresponding PIK is assigned. Table 1 presents the number of observations in each of the reference files.

Table 1. Number of Observations in PVS Reference Files

Reference File	Observations
Census Numident	780 million
GeoBase	1.2+ billion
Name and Date of Birth	800 million

3.3 Preparing Incoming Files for PVS

The first step of the PVS process is to edit data fields to make them homogenous for comparisons between incoming and reference files. There are several standard edits, which insure maximum success during linkage.

3.3.1 Incoming File Edits

The first edits are parsing and standardizing - parsing separates fields into component parts, while standardizing guarantees key data elements are consistent (e.g., STREET, STR are both converted to ST). Name and address fields are parsed and

standardized as they are key linkage comparators. Figure 1 provides an example of how name and address are parsed into separate fields and then standardized.

During record linkage, a matching scheme might require that the first two letters of the first name must agree, the first four letters of the last name must agree, and the first two letters of the middle name agree, with similar rules for the address components. In Figures 1 and 2 the records are determined to be non-matches even though at first glance they may appear to be matches. This example demonstrates the importance of data parsing and standardization.

Figure 1. Data Parsing and Standardization In File Editing

Record #	Name	Address
1	Mr. Bob G. Smith, Jr.	2345 S. Main Street
2	Robert George Smith	2345 Main St.

Figure 2. Parsed and Standardized Through File Editing

Record #	Name					Address			
	Prefix	*First*	*Middle*	*Last*	*Suffix*	*House #*	*Pre-Directional*	*Street Name*	*Type*
1	Mr.	Robert	G	Smith	Junior	12345	South	Main	ST
2		Robert	George	Smith		12345		Main	ST

The data parsing and standardizing phase of PVS is extremely important for the matching algorithms to function properly. The PVS system incorporates a name standardizer (McGaughey, 1994), which is a C language subroutine called as a function within SAS. It performs name parsing and includes a nickname lookup table and outputs name variants (standardized variations of first and last names). For example, Bill becomes William, Chuck and Charlie becomes Charles, etc. The PVS keeps both the original name (Bill) and the converted name (William) for matching. PVS also has a fake name table to blank names such as “Queen of the House” or “Baby Girl.” The name data are parsed, checked for nicknames, and standardized.

The PVS editing process also incorporates an address parser and standardizer, written in the C language and called as a function within SAS (U.S. Census Bureau Geography Division, 1995). It performs parsing of address strings into individual output fields (see Figure 2), and standardizes the spelling of key components of the address such as street type. The PVS also incorporates use of a commercial product to update zip codes, and correct misspellings of address elements.

3.3.2 Assign Geographical Codes to Address

The PVS provides an additional address enhancement by matching records in the incoming file to Census Bureau's Master Address File (MAF) in order to assign a unique address identifier, the MAF Identifier (MAFID), and other Census geographical codes (e.g., Census tract and block). The MAFID is used in the PVS for search purposes and as a linkage key for administrative files. Then, addresses are matched to the Census Bureau's Topologically Integrated Geographic Encoding and Referencing Database (TIGER) to obtain Census geographical codes.

CARRA receives bi-annual deliveries of an extract of the Census Bureau's MAF. Prior to using the MAF in any search, the MAF records are processed through the same commercial product as used in file editing to update zip codes and correct any misspellings of address elements. The MAFMatch function in PVS matches input addresses to the MAF extract using the Multi-Match engine and attaches the MAFID and Census geographical codes (county, tract, and block). This process contains a series of blocking strategies:

1. Matching addresses using the full address, including the within structure unit number
2. Using rural route addresses
3. Using basic street address (BSA) without the within-structure unit number.

The full address pass normally blocks on house number and zip code, matching on house number, street name, street prefix and suffix type, directional, and within structure unit ID. The rural route address pass normally blocks on the zip code and box number, matching on the rural route ID. The BSA-level pass is similar to the full address pass, but ignores the within structure unit ID. Only geographical codes are retained from BSA level matches.

Next, addresses are matched to TIGER (Census Bureau, 2010), which is comprised of shapefiles[5], and includes information for the fifty states, the District of Columbia, Puerto Rico, and the Island areas (American Samoa, the Commonwealth of the Northern Mariana Islands, Guam, and the United States Virgin Islands). TIGER Census geographical codes are added to the address record. Matching to the TIGER is done through one five digit zip code blocking pass.

3.3.3 Parameter Files

PVS uses the same Multi-Match engine for each probabilistic search type. For each search module the analyst defines a parameter file, which is passed to Multi-Match. The parameter file includes threshold value(s) for the number of passes, blocking keys, and within each pass, the match variables, match comparison type, and matching weights. The pass number determines how many times a match will be attempted, with various configurations and rules for the matching variables. In general, each successive pass is less restrictive than the previous one.

Records must first match exactly on the blocking keys before any comparisons between the match variables are attempted. Each match variable is given an m and u probability, which is translated by MultiMatch as agreement and disagreement weights. The sum of all match variable comparison weights for a record pair is the composite weight. All record pairs with a composite weight greater than or equal to the threshold set in the parameter file are linked, and the records from the incoming file for these linked cases are excluded from all remaining passes. All Numident records are always available for linking in every pass. Any record missing data for any of the blocking fields for a pass skips that pass and moves to the next pass.

Because the PIK is intended to link the same person between files, the cutoff thresholds are set conservatively (i.e., higher than might normally be set in some record linkage applications).

3.3.4 Blocking Strategy

Potentially, each record in the incoming data can be compared against all records in the reference file. The number of comparisons grows quadratically with the number of records for matching (Baxter, 2003). Considering the size of incoming data processed for Census Bureau purposes (incoming files can have over 500 million records) this is computationally untenable. Therefore PVS incorporates blocking strategies, which are methods of reducing the search space. Similar records are grouped together using

[5] A popular geospatial vector data format for geographic information systems software.

information from record attributes to create a blocking key, which may be comprised of more than one data attribute. Analysts have to be thoughtful in the selection of blocking keys, because of the error characteristics of the attributes used. Fields prone to a high probability of error should not be used. The use of multiple passes can mitigate this concern if independent blocking fields are incorporated for each pass.

In the PVS, the blocking key for the first pass is highly restrictive and has a low probability of reporting error. (An example is an exact address match.) Subsequent blocking schemes can be less restrictive, but may produce weaker links. A feature of the PVS is its analysis program, which is run after any search program. This program provides a listing of links, by pass number, with the score and listing of the records matched from each file. This allows the analyst to examine links in a very compact and elucidating form. For example, if there are three passes, the analyst can review matches in the third (less restrictive) pass with the lowest score to determine if the parameter estimates are yielding sensible links.

The Multi-Match engine, when utilized for the PVS, is a one-to many matching system. The reference file is searched to find the best match for the input record. All reference file records are available for matching during each pass, but matched input records do not proceed to the next pass.[6]

3.3.5 Summary

The time required to determine the appropriate data cleaning edits should not be under estimated, as it can take days or even weeks. In 2011 CARRA received over two billion commercial records in nine different files. While preparing these data records for PVS was time-consuming, the editing programs are reusable for future vintages of the same file.

Once the comparison fields used in the matching modules have been cleaned, the file proceeds through the rest of the PVS modules. Running PVS modules from beginning to end (edit program to the very last program) is a fast process. The 2010 Census Unedited File (CUF), had 350 million records and processed through every PVS module, excluding MAF match and SSN verification, in 60 hours .

In the following sections, we discuss each module of the PVS application. The user can choose to employ each module for the incoming file, or exclude certain modules if data are not available. Research is ongoing about the impact of switching the order of

[6] The PVS system can also be run to produce multiple matches for each input record (matching with replacement).

the modules. Preliminary results indicate the current module ordering is optimal. As new modules are added, the research on module ordering will be repeated.

3.4 PVS Search Modules

The PVS consists of one exact match – SSN Verification - - and four search modules: GeoSearch, NameSearch, DOBSearch, and Household Composition Search (HHComp). Incoming records *cascade* through the PVS – and only records failing a particular matching module proceed on to the next module. If a record is assigned a PIK from the reference file in any of the modules, no further searches are conducted. Each module has its own set of user defined blocking factors and parameter thresholds.

3.4.1 Verification Module

If the input file has a SSN data field, it first goes through the verification process. The verification module matches the reported SSN from the incoming file to the Census Numident file, along with the Alternate Name and Alternate Date of Birth Numident files. If the SSN is located in the Census Numident, and the name and date of birth agree sufficiently, the SSN is considered verified and the corresponding PIK is assigned. The SSN verification module is an exact match to SSN, so no parameter file is required for this step.

3.4.2 GeoSearch Module

The GeoSearch attempts to find SSNs or ITINs for incoming records that failed the verification module or without reported SSNs. This module links records from the incoming file to the GeoBase through blocking passes defined in the parameter file.

The typical GeoSearch blocking strategy starts with blocking records at the household level, then broadens the geography for each successive pass and ends at blocking by the first three digits of the zip code. The typical match variables are first, middle, and last names; generational suffix; date of birth; gender and various address fields.

The data for the GeoSearch module are split into 1,000 cuts based on the first three digits of the zip code (zip3) for record. The GeoSearch program works on one zip3 cut at a time, with shell scripts submitting multiple streams of cuts to the system. This allows for parallel processing and restart capability.

The GeoSearch module also incorporates the adjacency of neighboring areas with different zip3 values (Miller, Bouch, Layne, 2012). This can eliminate the bias of limiting

the blocking strategy to exact match on zip3 and obviates missing links based on zip3 blocking.

After the initial set of links is created, a post-processing program is run to determine which of the links are retained. A series of checks are performed. First the date of death information from Numident is checked. Next a check is made for more than one SSN assigned to a source record. If so, the best link is selected. If no best SSN is determined, all SSNs assigned are dropped and the next module begins. A similar post-processing program is run at the end of all search modules.

3.4.3 NameSearch Module

The NameSearch module searches the reference files for records failing the Verification and GeoSearch Modules. Only name and date of birth data are used in this search process. NameSearch consists of multiple passes against the Numident Name Reference file, which contains all possible combinations of alternate names and alternate dates of birth for each SSN in the Census Numident file, and includes data for ITINs.

The typical NameSearch blocking strategy starts with a strict first pass, blocking records by exact date of birth and parts of names. Successive passes block on parts of the name and date of birth fields to allow for some name and date of birth variation. The typical match variables are first, middle and last names, generational suffix, date of birth, and gender.

After the initial set of links is created, a post-processing SAS program is run to determine which of the links are retained, similar to the program used after the GeoSearch.

3.4.4 DOBSearch Module

The DOBSearch module searches the reference files for the records that fail the NameSearch, using name and date of birth data. The module matches against a re-split version of the Numident Name Reference file, splitting the data based on month and day of birth.

There are typically four blocking passes in the DOBSearch module. The first pass blocks records by first name in the incoming file to last name in the DOB Reference file and last name in incoming file to first name in the DOB Reference file. This strategy accounts for switching of first and last name in the incoming file.

3.4.5 HHCompSearch Module

The HHComp Search module searches the reference files for records that fail the DOBSearch. To be eligible for this module, at least one person in the household of an unmatched person must have received a PIK.

This module creates an eligible universe by selecting all not-found persons from the input data where at least one person in their household received a PIK. A reference file is created at run-time.

For persons with a PIK in the eligible household, all of the geokeys from the PVS GeoBase are extracted for each of these PIKs. The geokeys are unduplicated and all persons are selected from the PVS GeoBase with these geokeys. Next, the program removes all household members with a PIK, leaving the unPIKed persons in the household. This becomes the reference file to search against. There are typically two passes in this module. Records are blocked by MAFID, name, date of birth, and gender.

3.5 Master File Creation

Another feature of the PVS is the creation of a master file for the analyst. The master file creation is the last step in the PVS process. This program collects the results of each search module and creates a variable describing the final disposition of the PIK assignment – the module it was assigned in or the reason for the inability to assign a PIK (e.g., respondent's SSN is in a reference file, but the name is not a match). The program adds the PIK, MAFID and Census geographical codes to the incoming file and removes SSN or ITIN.

IV. PVS RESULTS

This section describes the results for some of the files CARRA has processed through the PVS system. The primary success measure for the PVS is the number of people matched to a reference file and assigned a PIK.

Table 2 presents detailed PVS results for three types of files: 2010 commercial data, a 2011 Federal representative file,[7] and the 2010 Census of Population and Housing (decennial census). In each of the three files, some records lack the necessary information to continue through the PVS. Ninety nine percent, 100 percent, and 97 percent of records for commercial, the Federal, and Census data have the necessary information to proceed through PVS.

The Decennial Census file doesn't contain SSNs, while the some of the commercial files and most of the Federal files do. The files with SSNs are eligible to go through the Verification module. In the commercial and the Federal file one record represents a person (person-level) and do not reflect household rosters. Therefore these files are not put through the Household Composition module. The Decennial Census persons are enumerated within households, and these records were submitted to the Household Composition module, provided they failed previous modules.

The SSNs in the Federal file are verified at a much higher rate than the commercial files. This seems reasonable, as participation in Federal programs requires accurate input. As show in Table 2, 99.9 percent of the Federal file records where verified leaving a scant .1 percent to proceed through the other probabilistic modules. In the commercial files, 76 percent of records were verified in the Verification module.

Because records for the 2010 Decennial Census lack SSNs, this file is sent to the probabilistic GeoSearch module and PIKS are assigned to 86 percent of the records. Of the 60,351 records sent to GeoSearch for the Federal file, 56 percent have a PIK assigned. The rate for the commercial files is 29 percent. The GeoSearch module assigned more PIKs to the census data than the commercial data because census addresses are generally cleaner and more complete.

[7] The commercial file results are the average of two commercial files and the Federal file is left unnamed.

Table 2. Sample PVS Results

PVS/MAFMatch Results	Average Of 2 Commercial Files	2011 Federal File	2010 Decennial Census
# Records delivered	339,295,722	53,181,072	312,471,327
# Addresses to MAFMatch	339,295,722	53,181,072	
Matched to MAF	283,605,634	47,721,226	
% with MAFID	83.59%	89.73%	
# Person Records	339,295,722	53,181,072	312,471,327
NO SEARCH: Refuse/OptOut (VERFLG=R,O)	0	0	0
NO SEARCH: Blank Name (VERFLG=X)	22,491	0	10,367,975
Persons Available for PVS	339,273,232	53,181,072	302,103,352
With SSN (any 9 digit number)	257,979,068	53,118,553	0
No SSN - to search	81,294,164	62,519	302,103,352
To VERIFICATION (with SSN)	257,979,068	53,118,553	0
Verified	195,571,749	53,058,202	0
NOT Verified - to search	62,407,320	60,351	0
Percent VERIFIED	75.81%	99.89%	0
To GEO Search	143,701,483	122,870	302,103,352
Found in GEO Search (VERFLG=S)	41,847,479	68,525	259,873,717
NOT found in GEO Search	101,854,004	54,345	42,229,635
Percent Found in GEO Search	29.12%	55.77%	86.02%
To NAME Search	101,854,004	54,345	42,229,635
Found in NAME Search (VERFLG=T)	5,328,014	13,158	19,960,452
NOT Found in NAME Search	96,525,990	41,187	22,269,183
Percent Found in NAME Search	5.23%	24.21%	47.27%
To DOB Search	96,525,990	41,187	22,269,183
Found in DOB Search (VERFLG=D)	73,802	587	304,690
NOT Found in DOB Search	96,452,188	40,600	21,964,493
Percent Found in DOB Search	0.08%	1.43%	1.37%
To HHComp Search	0	0	21,964,493
Found in HHComp Search (VERFLG=U)	0	0	1,975,295
NOT Found in HHComp Search	0	0	19,989,198
Percent Found in HHComp Search	0	0	8.99%
Total TO PVS (Avail for PVS)	339,273,232	53,181,072	302,103,352
Total validated	242,821,044	53,140,472	282,114,154
Total not validated (of avail)	96,452,188	40,600	19,989,198
% Validated (OF AVAILABLE)	71.57%	99.92%	93.38%
% Validated (OF ALL)	71.57%	99.92%	90.28%

The NameSearch and DOBSearch yield another 6 percent PIKs for the commercial data, 26 percent for Federal data, and 49 percent for the Census data. Only the Census file is sent through the household composition module, assigning PIKS for an additional 9 percent of the records .

In the last two lines of Table 2, we present the overall number of records that get PIKS. One line shows the percentage of the records available to be searched and the last line describes percentages for the entire dataset, including records not sent to any of the modules. The Federal file has the highest percent validated, the Census has the second highest. While commercial files have the lowest percent validated, the rate is above 70 percent. The reference files are largely SSN based, resulting in higher validation rates in Federal files.

The PVS provided the linking capability for the 2010 Census Match Study (O'Hara and Marshall, 2011). Without the ability to link persons across files via the PIK, the research for the Census Match Study could not be undertaken. Assigning MAFIDs to addresses in administrative data through the PVS process also proved invaluable to the study.

V. CONTINUOUS ENHANCEMENTS

The PVS is programmed in the SAS language, making it accessible to programmers and analysts alike. The few C language subroutines used are transparent to analysts because they are called within SAS programs. With the SAS expertise within the Census Bureau, code alteration and support requirements are met with ease.

The modular nature of the PVS is advantageous for setting parameter estimates. It is also beneficial if problems are encountered and the program needs to be restarted. Further the modular nature of the PVS allows the analyst to select modules depending on their source data. The PVS also supports parallel processing via scripting, which allows more than one process to run at a time (up to 80 on CARRA's existing Linux machine), greatly speeding up the process.

Each search module utilizes the same Multi-Match engine, eliminating the need to write specialized code for different linkage types (GeoSearch, NameSearch, etc.). The engine is written in SAS and called from each search module. Each of the modules passes a parameter file to the Multi-Match engine.

Finally, PVS offers analysts the ability to easily examine links in each module. Links are presented with the source record, the reference record, the matching fields, the pass in which the records were matched, and the score for the match. This provides the analyst with invaluable information for checking the results of the matching and to discern patterns in their data.

A criticism (NORC 2011b) of the production use of PVS (versus research usage) is that it sets a single threshold controlling the probability of false linkages, but no control for the probability that a link is in fact a mismatch. Setting one threshold means the analyst can only change the probability of false matches ***or*** the failure to link matched records, but not both. Testing and research have indicated that a single threshold performs extremely well.

CARRA is continually updating and improving the PVS. Research is under way to estimate overall false match and non-match rates in the PVS. Concomitant with this effort, CARRA is developing a method to estimate the probability of a link, in order to provide analysts with a measure of the certainty of a link. Other current and planned research projects include: implementing the EM algorithm to automatically generate *m* and *u* probabilities, and developing a model to determine false match rate and matching bias.

VI. SUMMARY

The PVS has been used for over a decade at the U.S. Census Bureau. To date, over 120 survey, administrative record, and commercial files have been processed in a production environment through the PVS. It is robust, proven, and flexible software that provides an end to end process.

The PVS is well maintained software, providing orderly steps for probabilistically assigning unique person and address identifers (PIKs and MAFIDs) to data. The PIK or MAFID facilitate linkage across files. The PVS is written in SAS, with some calls to C routines, and can be well understood by users. It makes use of generalized matching software, Multi-Match, and provides many tools for the analyst to check parameters and matching results.

Core expertise and institutional knowledge resides in CARRA for using and updating the software, setting parameters, and interpreting PVS results. There are opportunities to increase functionality of the current PVS. These challenges and issues are actively being researched by CARRA.

REFERENCES

Bilmes, J.A., (1998). “A Gentle Tutorial of the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models.” Internation Computer Science Institute, U.C. Berkeley http://ssli.ee.washington.edu/~bilmes/mypubs/bilmes1997-em.pdf

Colosi, R., (2004). “Evaluation 7:An Examination of Address Matching Software.” United States Bureau of the Census.

Dunn, Halbert L. (1946). "Record Linkage" (PDF). American Journal of Public Health 36 (12): pp. 1412–1416. doi:10.2105/AJPH.36.12.1412.

Farber, J., Leggieri, C. (2002). “Building and Validating the Prototypical Statistical Administrative Records System,” United States Census Bureau.

Fellegi, I. P., and Sunter, A. B. (1969). "A Theory for Record Linkage," Journal of the American Statistical Association, 64, 1183 1210.

Herzog, T. N., Scheuren, F., and Winkler, W.E., (2007). Data Quality and Record Linkage Techniques, New York, N. Y.: Springer.

Kelly, R. (1984). “Blocking Considerations for Record Linkage Under Conditions of Uncertainty,” United States Census Bureau, Statistical Research Division.

Lahiri, P. A., and Larsen, M. D. (2005). “Regression Analysis with Linked Data,” Journal of the American Statistical Association, 100, 222-230.

Larsen, M. D. (2005). “Hierarchical Bayesian Record Linkage Theory,” Iowa State University, Statistics Department Technical Report.

Larsen, M.D. (2010). Record Linkage Modeling in Federal Statistical Databases. FCSM Research Conference, Washington, DC.

McGaughey, Anne (1994). “The 1995 Bureau of the Census Computer Name Standardizer.” Statistical Research Division.

Michelson, M. and Knoblock, C. A. (2006). “Learning Blocking Schemes for Record Linkage,” Proceedings of AAAI-2006.

Miller,C., Bouch, M., Layne M. (2012). “Spatial Adjacency Search Module for the PVS”.U.S. Census Bureau. Center for Administrative Records Research and Application (CARRA) Working Paper.

Newcombe, H. B., Kennedy, J. M., Axford, S. J., and James, A. P. (1959). "Automatic Linkage of Vital Records," *Science*, 130, 954-959.

NORC at the University of Chicago (2011a). "Assessment of the U.S. Census Bureau’s Person Idenfication Validation System.”

NORC at the University of Chicago (2011b). "Person Validation and Entity Resolution Conference Report.”

O’Hara, A., Marshall, L. (2011). “2010 Census Evaluations, Experiments and Assessments Plan: 2010 Match Study,” 2010 Census Planning Memoranda Series, United States Census Bureau.

Prevost, R., Leggieri, C. (1999). Expansion of Administrative Rececords Uses at the Census Bureau: A Long-Range Research Plan. *FCSM Research Conference*, Washington, DC.

Sadsinle, M., Hall, R., Fienberg, S. (2011). "Multiple Record Linkage: Generalizing the Felligi-Sunter Theory to More Than Two Datafiles," *JSM Proceedings.*

Wagner, D. (2012). "Documentation for the Person Identification Validation System," CARRA Internal Documentation, U.S. Census Bureau.

Wagner, D. (2012). "Documentation for the Multi-Match Record Linkage Software," CARRA Internal Documentation, U.S. Census Bureau.

Winglee, M.,Valliant, R., Scheuren F, (2005). "A Case Study in Record Linkage," Survey Methodolgy, Vol. 31., No 1., pp. 3-11..

Winkler, W. E. (1988). "Using the EM Algorithm for Weight Computation in the Fellegi-Sunter Model of Record Linkage," Proceedings of the Section on Survey Research Methods, American Statistical Association, 667-671.

Winkler, W. E. (1990b). "String Comparator Metrics and Enhanced Decision Rules in the Fellegi-Sunter Model of Record Linkage," Proceedings of the Section on Survey Research Methods, American Statistical Association, 354-359.

Winkler, W. E. (1994). "Advanced Methods for Record Linkage," Proceedings of the Section on Survey Research Methods, American Statistical Association, 467-472.

Winkler, W. E. (1995). "Matching and Record Linkage," in B. G. Cox, D. A. Binder, B. N. Chinnappa, A. Christianson, Colledge, M. A., and P. S. Kott (eds.) Business Survey Methods, New York: J. Wiley, 355-384 (also available at http://www.fcsm.gov/working-papers/wwinkler.pdf).

Winkler, W. E. (2003a), "Methods for Evaluating and Creating Data Quality," Proceedings of the ICDT Workshop on Cooperative Information Systems, Sienna, Italy, January 2003, longer version in Information Systems (2004), 29 (7), 531-550.

Winkler, W. E. (2006a). "Overview of Record Linkage and Current Research Directions," U.S. Bureau of the Census, Statistical Research Division Report, Statistical Report Series.

Winkler, W. E. (2007). "Automatically Estimating Record Linkage False Match Rates," U.S. Bureau of the Census, Statistical Research Division Report, Researcg Report Series.

Yancey, W.E. (2007). "BigMatch: A Program for Extracting Probable Matches from Large Files," Statistical Division Research Report, http://www.census.gov/srd/papers/pdf/RRC2007-01.pdf.

Yancey, W.E. (2002), "Improving EM Parameter Estimates for Record Linkage Parameters," Proceedings of the Section on Survey Research Methods, American Statistical Association, CD-ROM (also report RRS 2004/01at http://www.census.gov/srd/www/byyear.html).

United States Bureau of the Census, Geography Division.(1995). "Address Standardizer Documentation."

United States Census Bureau. (2010). "2010 Census TIGER/Line Shapefiles."