Cityscape

A Journal of Policy Development and Research

National Survey of Mortgage Originations

Volume 21, Number 2 • 2019

U.S. Department of Housing and Urban Development Office of Policy Development and Research

PD&R

Cityscape

A Journal of Policy Development and Research

National Survey of Mortgage Originations
Volume 21, Number 2 • 2019

U.S. Department of Housing and Urban Development
Office of Policy Development and Research

The goal of *Cityscape* is to bring high-quality original research on housing and community development issues to scholars, government officials, and practitioners. *Cityscape* is open to all relevant disciplines, including architecture, consumer research, demography, economics, engineering, ethnography, finance, geography, law, planning, political science, public policy, regional science, sociology, statistics, and urban studies.

Cityscape is published three times a year by the Office of Policy Development and Research (PD&R) of the U.S. Department of Housing and Urban Development (HUD). Subscriptions are available at no charge and single copies at a nominal fee. The journal is also available on line at huduser.gov/periodicals/cityscape.html.

PD&R welcomes submissions to the Refereed Papers section of the journal. Our referee process is double blind and timely, and our referees are highly qualified. The managing editor will also respond to authors who submit outlines of proposed papers regarding the suitability of those proposals for inclusion in *Cityscape*. Send manuscripts or outlines to cityscape@hud.gov.

Visit PD&R's website, huduser.gov, to find this report and others sponsored by PD&R. Other services of HUD User, PD&R's Research and Information Service, include listservs, special interest and bimonthly publications (best practices, significant studies from other sources), access to public use databases, and a hotline (1–800–245–2691) for help with accessing the information you need.

Contents

Symposium

National Survey of Mortgage Originations
Guest Editors: Robert B. Avery and Ron Borzekowski

Guest Editors' Introduction

National Survey of Mortgage Originations

Robert B. Avery
Federal Housing Finance Agency

Ron Borzekowski
Consumer Financial Protection Bureau

Disclaimer
The views expressed in this article are those of the authors and are not necessarily those of the Consumer Financial Protection Bureau, the Federal Housing Finance Agency or any other agency of the U.S. Government.

In 2012, the Federal Housing Finance Agency (FHFA) and the Consumer Financial Protection Bureau (CFPB) began work on the National Mortgage Database (NMDB®) and a new quarterly mail survey called the National Survey of Mortgage Originations (NSMO) associated with it. In November 2018, the two agencies released a public use file covering the first 4 years (2013–2016) of mortgage originations in NSMO.[1] The public use file contains survey responses and key administrative data obtained from other sources. The availability of matched administrative and survey data offers researchers and policymakers a unique view of the U.S. mortgage market, arguably one of the most critical markets in the U.S. economy.

NSMO is unique for two reasons. First, it fills an information void. There is no other recurring nationally representative survey of borrowers who have just obtained a mortgage. The U.S. Department of Housing and Urban Development's (HUD) American Housing Survey and the Federal Reserve's Survey of Consumer Finances and Survey of Household Economics contain some information on mortgages, but the number of recent mortgage borrowers they capture is small because of the different focus of those surveys. The National Association of Realtors® annual survey of homebuyers and sellers covers home purchasers but not refinancers and is not generally available to the public.

[1] The public use NSMO data and documentation are available at https://www.fhfa.gov/nsmodata.

Second, NSMO provides access to high-quality administrative data, as NSMO's sampling frame is derived from the NMDB. The NMDB contains data for a 5-percent random sample of closed-end, first-lien residential mortgages active at any time in the credit files of one of the three national credit bureaus from January 1998 to the present. The database is updated each quarter with a 5-percent sample of newly reported mortgages—typically, those just opened. The quarterly NSMO sample is a randomly drawn subset of 6,000 borrowers associated with those newly reported mortgages.[2] Based on the quarterly cycle, the typical respondent fills out the questionnaire between 6 and 9 months after origination.

The NMDB data include origination and performance information for each NMDB sample mortgage—and credit information about the borrowers associated with those mortgages—from origination to closure. Although extensive, the credit bureaus' files do not contain information on a number of key mortgage features, such as the loan's purpose (home purchase or refinance), whether it had an adjustable or fixed rate, its securitization status, or whether it was backed by an owner-occupied property, vacation home, or investor property. That information is available from a variety of sources, including high-quality matches to administrative file records maintained by Fannie Mae, Freddie Mac, the Federal Housing Administration, the U.S. Department of Veterans Affairs, the Rural Housing Service, and the Federal Home Loan Banks, which collectively account for more than 70 percent of mortgages in the United States. Further information was obtained from other public and proprietary data sources, including deed record filings, Home Mortgage Disclosure Act filings, and commercially available servicing databases.

The NMDB database combines information from each of those sources into a single record for each mortgage. Data merging was conducted behind a protected firewall at the credit bureau, using information about the borrowers, such as their names, Social Security numbers, addresses, and dates of birth, to facilitate high-quality matches. Variable values for the small percentage of loans that could not be matched to any of those sources were imputed using a variety of methods. The resulting matched database, as delivered to FHFA and CFPB, is de-identified—it does not identify specific borrowers.

The availability of high-quality administrative data for each sample loan means that NSMO does not have to ask the respondents to provide factual information about the mortgage. Thus, the survey instrument focuses on obtaining information about the borrowers' experience, perceptions, and expectations—which are not readily available anywhere else. The survey asks sampled borrowers 94 questions (including more than 300 subquestions) about their experience shopping for and closing on a mortgage; their knowledge and perceptions of the mortgage market and the mortgage process; their use of housing counselors; and their future expectations about house price appreciation and critical household and financial events. The survey also asks about demographic details and household composition information that is not available in the NMDB.

[2] At the time of this writing, that figure represents a sampling rate of about 1 in 300 new mortgages in the nation. About one-half of the survey responses have been completed online and the other half by mail. The overall response rate of NSMO is about 33 percent, and the usable response rate is about 28 percent. The NSMO survey instrument was improved substantially in 2015, was refined moderately in 2016, and continues to receive minimal adjustments moving forward. In 2015, the NSMO was sent to additional borrowers associated with a special sample of 1,000 mortgages originated in 2014 in completely rural counties. That sample was used for one of the articles in this volume.

The public use NMSO data contain variables drawn from both the survey responses and the NMDB administrative data file[3] and consist of 24,847 loans originated from 2013 through 2016, with additional years to be added in the future. When appropriately weighted, the data should be representative of new mortgages originated over the 4-year period.[4]

The four articles in this volume use the public use NSMO data to address a set of related and important questions about the knowledge and behavior of borrowers in the U.S. housing market over the past few years.[5] To make informed choices about buying a home or choosing and closing on a mortgage, consumers need information on a multitude of factors, including their own financial situation, the state of the housing market in their area, the details of available loan products, an understanding of the various service providers, and the process involved in the mortgage transaction. Some of that information is provided by private sources, although search costs or agency costs could be a factor in borrowers' accessing that information; however, in some cases, policymakers have mandated that consumers, or at least certain consumers, be provided particular information.

The first two articles focus on questions related to information provided primarily by private sources. In the first article, "Mortgage Experiences of Rural Borrowers in the United States: Insights from the National Survey of Mortgage Originations," Tim Critchfield, Jaya Dey, Nuno Mota, and Saty Patrabansh use a special supplement of NSMO that oversampled borrowers in completely rural areas—those without a significant population center—to compare the experience of borrowers in those areas with borrowers in urban and larger rural markets. Rural mortgage markets have been a particular concern for policymakers and researchers.[6] One conjecture is that because rural markets are relatively thin, they suffer from a lack of scale in information, and thus, mortgages are more expensive to originate. For example, lack of scale may make it more difficult for lenders and the mortgage secondary market to properly assess home values and thus make markets less competitive. In this article, the authors indirectly test those conjectures.

Collectively, the 644 counties covered by the special sample account for only 1 percent of U.S. residential mortgage originations, but those are arguably the areas most likely to reveal informational problems. Indeed, Critchfield and his colleagues show that borrowers in the special sample were less likely than other borrowers to be satisfied with the mortgage closing process and the required disclosures. They also tend to be less confident or knowledgeable about some details of mortgages. Together, those findings suggest that perhaps supply conditions lead to a poorer

[3] Because mortgage filings are typically a matter of public record, releasing the detailed administrative data contained in the NMDB as part of NSMO with any degree of geographic detail would not have been possible without creating significant privacy concerns. Consequently, almost all geographic information is suppressed in the NSMO public use file. Unfortunately, those concerns did not allow for the data of the special supplemental sample of completely rural borrowers used for one of the articles in this volume to be included in the public release. County-level house price indices and economic factors used for another article also could not be included in the public use file for the same reason.

[4] The administrative data in the NMDB are used to calculate nonresponse adjustment weights for NSMO. Missing values for almost all variables in the public use file have been imputed using an iterative regression—or logistic model-based approach—that also relies heavily on administrative data. The incidence of missing values is less than 10 percent for almost all variables.

[5] Use of information not in the public use NSMO is explicitly noted in two of the articles.

[6] The congressional bill that established the FHFA mandated that it pay particular attention to rural markets in its regulation of Fannie Mae and Freddie Mac.

information environment in which these rural consumers are making their decisions. The authors also show that, collectively, the borrowers in the special sample have higher contract interest rates than otherwise comparable borrowers in other areas—another potential sign of higher costs or poorer decisionmaking.

In the second article, "Perceptions and Expectations of Mortgage Borrowers: New Evidence from the National Survey of Mortgage Originations," Chad Redmer examines how cognizant different types of borrowers are of changes in recent house prices in their communities and the extent to which borrowers' expectations of the future may help forecasters predict future price changes. He finds, first, that borrowers overall have a good understanding of recent house price changes, and second, that their future expectations seem to incorporate some knowledge beyond what is available in public macro indicators. The first finding is important because borrowers who have just obtained a mortgage (the sample frame for the NSMO) are the group for whom house price knowledge may be most important. Poor decisions in either house purchases or refinances can have long-term implications for consumers' well-being. Redmer further finds that first-time homebuyers, the group with the most potential vulnerability, actually are among the more informed borrowers.

The second finding—that borrower expectations have additional predictive power over macro indicators—is more difficult to interpret. It could mean that consumers have additional information about their local markets that is not incorporated into traditional economic measures. Alternatively, it could stem from a self-fulfilling process whereby borrowers with bullish views tend to drive up prices in their markets. More research must be done to understand this finding better.

The last two articles focus more specifically on market interventions mandated by statute and regulations aimed at improving the information available to borrowers. The interventions considered include (1) homebuying handbooks that government agencies mandate originators give to borrowers and (2) homeownership education and counseling which is required for certain homebuyers.[7]

More specifically, federal regulations require that purchase-mortgage borrowers get a "special information booklet" to assist in their decisions. CFPB, which is now in charge of implementing the regulations, changed the booklet in October 2015. In "National Survey of Mortgage Originations Survey Data on *Your Home Loan Toolkit*," Brian Bucks, Tim Critchfield, and Susan Singer treat this change as a "natural experiment" and use the time series component of NSMO to examine the effect of the booklet change. They find that the percentage of homebuyers who remember receiving information from their lender increased with the introduction of the Toolkit. That increase was larger for borrowers with lower credit scores, household income, education, or experiences with prior mortgages. The percentage of respondents asking followup questions also increased after the new Toolkit introduction so that the total number of borrowers "digging deeper" also seems to have increased with this change.

[7] One can also see mandated standardized disclosures, such as the Loan Estimate and the Closing Disclosure, as similarly aimed at improving the information available to borrowers.

The fourth and final article presents a comparison of the outcomes, satisfaction, and knowledge metrics of borrowers who underwent housing counseling and otherwise similar borrowers who did not. Using propensity score methods, Robert B. Argento, Lariece M. Brown, Sergei Koulayev, Grace Li, Marina Myhre, Forrest Pafenberg, and Saty Patrabansh, in "First-Time Homebuyer Counseling and the Mortgage Selection Experience in the United States: Evidence from the National Survey of Mortgage Originations," compare borrowers who did and did not undergo counseling. They find that first-time homebuyers who reported receiving homeownership education and counseling also reported better mortgage knowledge, more confidence in their ability to explain the mortgage process, and a higher level of satisfaction with the mortgage they received.

The four articles in this symposium each use the survey data from NSMO to begin to address important questions about consumer knowledge and understanding in the U.S. housing and mortgage markets. Although each of these articles is a first step in its respective area, they all find important variations by geography or borrower characteristics, and all offer new areas for additional research. The articles chosen for this volume were authored by members of the team that helped develop the NMDB and NSMO, but it is our hope that these articles will stimulate researchers in the greater academic and the policy analysis community to use the NSMO data for their own investigations.

Acknowledgments

We thank the authors of the articles in this volume and numerous reviewers for comments and suggestions.

Mortgage Experiences of Rural Borrowers in the United States: Insights from the National Survey of Mortgage Originations

Tim Critchfield
Consumer Financial Protection Bureau

Jaya Dey
Freddie Mac

Nuno Mota
Fannie Mae

Saty Patrabansh
Federal Housing Finance Agency

Disclaimer
The views expressed in this article are those of the authors and are not necessarily those of the Consumer Financial Protection Bureau, Freddie Mac, Fannie Mae, or the Federal Housing Finance Agency.

Abstract

To date, research on rural mortgage markets in the United States has been limited by a lack of data on the specific mortgage experiences of borrowers living in rural areas. To fill this data gap, the National Survey of Mortgage Originations (NSMO) conducted a survey that oversampled people who took out mortgages in completely rural counties in 2014. This article shows results from this survey, contrasting the characteristics, experiences, and loan terms of mortgage borrowers in completely rural counties to those of borrowers in metropolitan and other non-metropolitan areas. Completely rural counties are those with no urban cluster or an urban population less than 2,500. We find that borrowers in completely rural counties paid slightly higher interest rates on average and were less satisfied that their mortgage was the one with the best terms to fit their needs than borrowers in other areas. These results persist even after controlling for income, credit quality, and other borrower characteristics. Completely rural borrowers were less likely than other borrowers to be satisfied with the mortgage closing process, the timeliness of disclosures, and the disclosure documents themselves. Finally, compared with borrowers in more urban areas, borrowers in completely rural areas tend to be less confident or less knowledgeable about some details of mortgages, and they are more likely to initiate contact with their lender.

Introduction

There is a widespread belief that lenders and credit markets in rural areas of the United States differ from those in urban areas. The literature on "relationship lending" argues that lending in rural areas differs fundamentally from lending in other areas because rural lenders have greater personal knowledge about their borrowers and local economic conditions.[1] The literature on community banking shows that lenders in rural areas tend to have fewer assets and generally have smaller geographic markets than larger financial institutions (Critchfield et al., 2004). The smaller scale of lending in rural areas could potentially constrain the supply of mortgages, make mortgages more costly to originate, and could adversely affect mortgages taken out by borrowers.[2]

To some degree, federal housing and mortgage policies reflect the distinctive features and challenges of mortgage lending in rural areas. For example, the Housing and Economic Recovery Act of 2008 (HERA) assigned Fannie Mae and Freddie Mac a *Duty to Serve* (DTS) three underserved markets by increasing the liquidity of mortgage investments and improving the distribution of mortgage investment capital to those markets.[3] One of the underserved markets specifically mentioned is the rural market, and another is manufactured housing—which is much more prevalent in rural areas (CFPB, 2014). Similarly, the Consumer Financial Protection Bureau (CFPB) provided exceptions for loans made by small creditors that operate predominantly in rural or underserved areas in several of its mortgage rules.

This article explores the underlying premise of these government policies: that mortgage borrowers in rural areas are potentially "less well-served" than those in other areas. Such an article has been difficult to write in the past because of a lack of data.[4] There are few available data sets containing detailed information on the characteristics of borrowers and their mortgages that are both representative and contain enough borrowers in rural areas to make meaningful comparisons. For example, data reported under the Home Mortgage Disclosure Act (HMDA) provide data on millions of mortgages and thus include many rural loans, but HMDA exempts small lenders and lenders with branches exclusively outside metropolitan areas, making the coverage of rural loans incomplete and potentially unrepresentative.

To fill the need for data on mortgages and mortgage borrowers in rural areas, the Federal Housing Finance Agency (FHFA) and CFPB conducted a special supplemental survey of 1,000 mortgages as part of the National Survey of Mortgage Originations (NSMO) which is based on the National Mortgage Database (NMDB®).[5] The supplemental sample that is the cornerstone of this article covers mortgages originated at any time in 2014 in counties defined as "completely rural." This article uses data for 267 borrowers from the special sample of completely rural mortgages and 6,273 respondents to the regular NSMO survey who took out a mortgage in 2014.[6]

---

[1] See, for example, https://www.stlouisfed.org/bank-supervision/2013-community-banking-conference/videos/small-business-lending-and-social-capital-are-rural-relationships-different or https://www.stlouisfed.org/Publications/Regional-Economist/April-2002/Community-Ties-Does-Relationship-Lending-Protect-Small-Banks-When-the-Local-Economy-Stumbles.

[2] See HAC (2012) for a description how mortgage finance in rural communities has evolved over the last decade.

[3] See 12 U.S.C. 4565(a)(1).

NSMO data allow us to compare the expectations, perceptions, knowledge, experience, and satisfaction of respondents in completely rural areas with those respondents in other rural areas, as well as urban areas. The remaining sections of the article summarize the relevant literature; define the three geographical groups used in the analysis; and describe the geographic differences in property, loan, and borrower characteristics, borrowers' experience and knowledge, and borrower's lender and mortgage choice.

Literature

This article bridges two largely distinct literatures, namely, (1) studies of rural housing and mortgage markets, and (2) the literature on borrower perceptions and experiences within the mortgage market.

The first set of studies have described the trends in employment, incomes, and housing market characteristics across the urban-rural divide (HAC, 2012; Mota, 2016; USDA, 2016) along with the effect and prevalence of government-sponsored enterprises in rural markets (Ambrose and Buttimer, 2005; MacDonald, 2001; Vandell, 1997). These studies compared rural and urban housing and mortgage markets at an aggregate level but did not consider differences in the characteristics of borrowers across rural and urban areas.

Because these studies did not consider the differences in characteristics between rural and urban borrowers, they generally find differences in borrower perceptions and experience by income, race and ethnicity, and prior experience with mortgages, rather than across area distinctions. Bucks and Pence (2008) showed that consumers were aware of the general terms of their mortgage (for example, adjustable-rate versus fixed-rate mortgages, number of months of required payment, and required monthly payment amount); they were less aware of details about potential interest rate changes in their adjustable-rate mortgages. Huang et al. (2017) reported that there were common misperceptions among borrowers about mortgage qualification criteria; specifically, borrowers tended to overestimate minimum credit score and down payment requirements. Cai and Shahdad (2015) found that one-third of homebuyers in their sample did not obtain quotes from multiple lenders when shopping for a mortgage. In addition, Cai and Shahdad concluded that borrowers with higher incomes, younger borrowers, and minority borrowers were more likely to obtain multiple lender quotes compared with other borrowers; while lower-income and first-time homebuyers were particularly likely to use family, friends, and co-workers when selecting a lender.

[4] This lack of data has limited the academic and policy research on rural mortgage markets, and both Fannie Mae and Freddie Mac cite lack of data and research on rural mortgage lending as one of the critical issues in rural mortgage markets. See Freddie Mac's "Freddie Mac Duty to Serve Underserved Markets Plan" from December 2017 and Fannie Mae's "Introduction of the Duty to Serve Underserved Markets Plan for the Manufactured Housing, Affordable Housing Preservation, and Rural Housing Markets" from December 2017.

[5] NSMO and NMDB are described in the guest editor's introduction.

[6] While the NSMO data for the regular sample is publicly available, the data for the supplemental NSMO sample is not publicly available for privacy considerations because membership in the supplemental sample conveys information about the location of the borrowers, which is suppressed in the public use NSMO. Out of the 1,000 surveys mailed in the supplemental survey, only 267 were useable records.

Definition of Urban and Rural Areas

We classify U.S. counties into three groups based on the U.S. Department of Agriculture's (USDA) county-level Rural Urban Continuum Code (RUCC) classification from 2013.[7] The first group, which we refer to as "metro" counties, consists of 1,167 counties in Metropolitan Statistical Areas (MSA) (RUCC codes 1, 2, or 3).[8] The second group includes 1,332 counties that are not in metropolitan areas, but have at least one urban cluster of 2,500 or more people (RUCC codes 4, 5, 6, or 7)—many of these counties, which we refer to as "non-metro" counties, are in Micropolitan Statistical Areas.[9] The last group, which we refer to as "completely rural" counties, comprises 644 counties that are designated as "completely rural or less than 2,500 urban population" (RUCC codes 8 or 9). The supplemental NSMO sample used in this article was drawn from mortgages in completely rural counties.

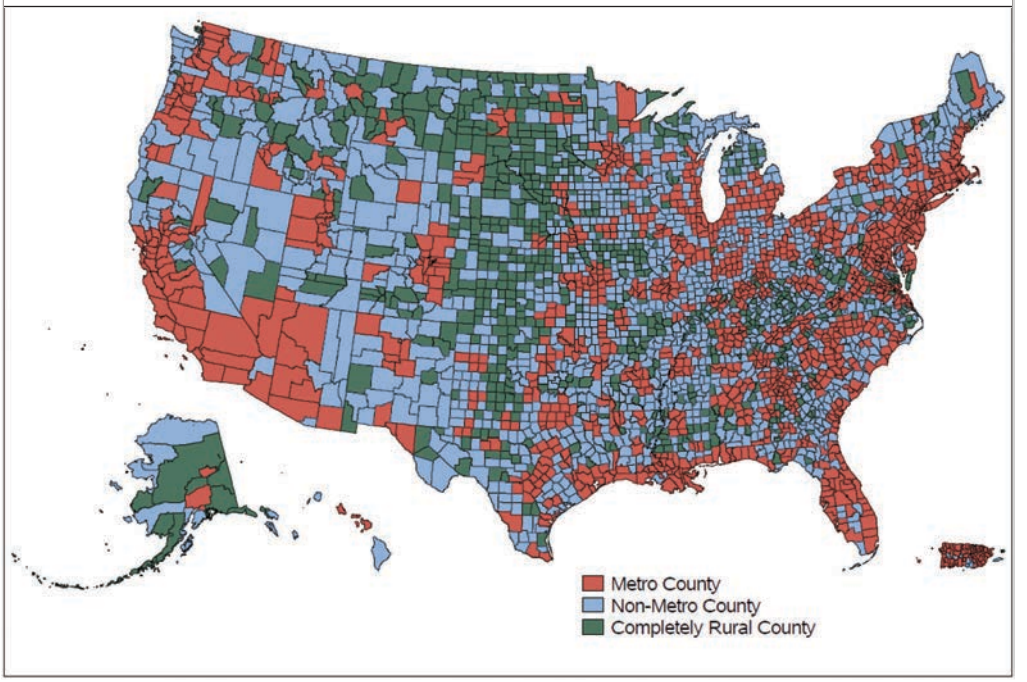
Using this classification, 37 percent are metro counties, 42 percent are non-metro counties, and 21 percent are completely rural counties. Exhibit 1 shows that most counties along the coasts are metro counties.[10] Non-metro counties are spread throughout the country, but the coasts have the lowest share ranging from 33 percent of counties in the Middle Atlantic Census division to 35 percent in the Pacific division. Completely rural counties are primarily located in parts of the Midwestern, Mountain, and Southern states.

Exhibit 2 shows the share of the 2014 population and housing units in each of the three county types according to the Census Bureau's American Community Survey (ACS) 2010–2014 5-Year Estimates.

[7] While there are many definitions of "rural" areas in the literature and government programs, including DTS, this article arises out of the supplemental NSMO sample and is restricted to how the sample was drawn. See USDA documentation at https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/documentation/.
[8] MSAs have at least one urbanized area of 50,000 or more population, plus adjacent territory that has a high degree of social and economic integration with the core as measured by commuting patterns.
[9] Micropolitan Statistical Areas have at least one urban cluster of at least 10,000 but less than 50,000 population, plus adjacent territory that has a high degree of social and economic integration with the core as measured by commuting patterns. For the USDA, "non-metro" would include RUCC codes 8 and 9, but we have broken them out for the purpose of this study.
[10] The Census Bureau divides the country into four regions and nine divisions.

Exhibit 1

Geographic Classification of Counties and County Equivalents



Source: USDA RUCC data, 2013

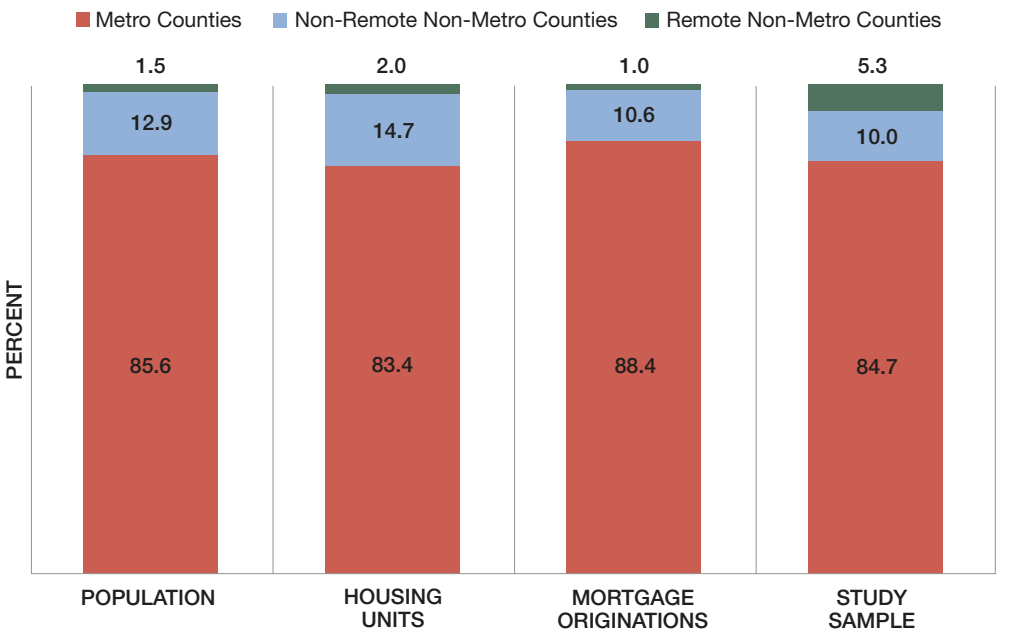
Most people and housing units in the United States—86 percent of people and 85 percent of housing units—were in metro counties in 2014. The remaining population and housing units were primarily located in non-metro counties. Completely rural counties accounted for only 2 percent (4.6 million people) of the U.S. population and 2 percent (2.6 million housing units) of housing units in 2014.

The final bars in exhibit 2 show the share of mortgage originations by geographic location. The share of mortgage originations in metro counties is 88.4 percent, a few percentage points higher than the share of people and housing units in these areas. Furthermore, just 1 percent of originations, or 55,000 loans, were in completely rural counties.[11] This highlights the importance of drawing a supplemental sample of mortgages in completely rural counties in order to obtain more accurate estimates. In particular, about 5 percent of our sample (345 out of 6,540) is for mortgages in completely rural areas, a sample size that is five times larger than a simple random sample of the country would be expected to yield for these areas.[12]

[11] County-type shares of NMDB 2014 originations in the NMDB data are similar to those in HMDA 2014 data, where metro accounts for 89.4 percent of originations, non-metro 9.7 percent, and completely rural 1.0 percent of originations.
[12] Analysis in this article, including the regressions, is based on analytic weights that account for both the sampling weight and the non-response adjustment. For each survey response, the sample weight adjusted for non-response was computed by multiplying the sampling weight and the non-response adjustment. Then, the analytic weight was computed separately for each of the following three groups: (1) mortgages in completely rural counties included in the special supplementary sample (267 mortgages), (2) mortgages in completely rural counties included in the regular sample (78 mortgages), and (3) mortgages in non-metro and metro counties included in the regular sample (6,273 mortgages). The analytic weight for a survey response was computed by multiplying the non-response-adjusted sample weight of that survey response by the total sample size of the group and dividing it by the sum of the non-response-adjusted sample weight of that group. The total number of completely rural mortgages in the analysis sample is 345 (267+78).

Exhibit 2

Population, Housing Unit, Mortgage Originations, and Sample Size by County Type, 2014



Note: Study sample includes a special oversample of mortgages originations in completely rural counties.
Sources: Census Bureau ACS 2010–2014 5-year estimates data for population and housing units by county; NMDB® data for 2014 first-lien mortgage originations; NSMO 2014 data for sample sizes by county

Geographic Differences in Property, Loan, and Borrower Characteristics

The primary data we use are survey responses from the regular quarterly NSMO samples and a special supplemental sample of borrowers in completely rural areas.[13] NSMO data cover a representative sample of first-lien residential mortgages taken out since 2013.

Exhibit 3 shows property and mortgage characteristics for our sample. Because purchasers and refinancers may have different expectations, knowledge, or experience, we present results for all originations, purchase mortgages, and refinance mortgages, separately.

Compared with those in metro areas, properties associated with mortgage originations in completely rural and non-metro areas were less likely to be single-family attached dwellings, and were more likely to be manufactured homes.[14] Manufactured housing is often titled as *chattel* (personal property) even though about three-fifths of manufactured-housing residents own the

[13] The supplemental sample uses the same questionnaire as the regular sample for 2014 mortgages with a few exceptions noted later.
[14] Single-family attached dwellings include townhouses, row houses, villas, apartments, and multi-unit dwellings. In the NSMO survey, mobile and manufactured homes are identified as manufactured housing because all such homes built after 1976 are defined by the HUD as manufactured housing.

Exhibit 3

Property, Mortgage, and Borrower Characteristics by Mortgage and Market Type (1 of 2)

Characteristics	All Mortgages			Purchase Mortgages			Refinance Mortgages		
	Metro (%) (n=5,541)	Non-Metro (%) (n=654)	Completely Rural (%) (n=345)	Metro (%) (n=3,044)	Non-Metro (%) (n=348)	Completely Rural (%) (n=161)	Metro (%) (n=2,497)	Non-Metro (%) (n=306)	Completely Rural (%) (n=184)
Purchase Mortgage Share	55	52	53	NA	NA	NA	NA	NA	NA
Property Type									
Single-Family Detached	83	86	86	82	88	89	84	85	83
Attached	16	7	2	17	8	2	15	5	2
Mobile or Manufactured	1	6	9	1	2	5	1	10	13
Land	0	1	3	0	2	4	0	0	2
Loan Amount									
Less than $50,000	3	10	14	2	8	10	4	12	17
$50,000 to $149,999	35	60	55	34	60	60	38	59	50
$150,000 to $299,999	39	25	27	41	27	29	37	24	25
$300,000 or More	22	5	4	23	5	1	22	5	8
Mortgage Term to Maturity									
30 Years or More	76	65	59	89	82	78	60	47	38
Median Terms (in Years)	30	30	30	30	30	30	30	20	15
Loan-to-Value (LTV) Ratio (Median)	80	80	80	88	90	95	73	74	73
Household Income									
Less than $35,000	6	12	10	6	12	8	7	11	13
$35,000 to $49,999	10	15	21	11	16	20	9	13	21
$50,000 to $74,999	18	23	28	20	21	30	16	24	25
$75,000 to $99,999	18	25	16	18	21	14	19	28	17
$100,000 to $174,999	30	20	20	29	21	21	31	18	20
$175,000 or More	17	7	5	17	8	7	17	5	3
Household Employment									
One or More Full-Time	87	83	81	90	86	86	84	79	75
None Full-Time	13	17	19	10	14	14	16	21	26
Household Type									
Couple	76	78	75	76	75	76	75	81	74
Single	24	22	25	24	25	24	25	19	26
Respondent Race/Ethnicity									
Non-Hispanic White	77	90	92	77	89	94	77	91	91
Hispanic and Non-White	23	10	8	23	11	7	23	9	9

Exhibit 3

Property, Mortgage, and Borrower Characteristics by Mortgage and Market Type (2 of 2)

Characteristics	All Mortgages			Purchase Mortgages			Refinance Mortgages		
	Metro (%) (n=5,541)	Non-Metro (%) (n=654)	Completely Rural (%) (n=345)	Metro (%) (n=3,044)	Non-Metro (%) (n=348)	Completely Rural (%) (n=161)	Metro (%) (n=2,497)	Non-Metro (%) (n=306)	Completely Rural (%) (n=184)
Respondent Age									
35 or Younger	28	29	30	40	39	45	13	17	13
36 to 45	23	24	23	23	24	25	23	24	20
46 to 55	23	19	17	17	16	10	29	23	24
56 to 65	16	17	19	12	14	12	22	21	25
66 or Older	10	11	12	7	8	8	13	15	17
Median	*45*	*45*	*43*	*39*	*39*	*37*	*51*	*49*	*52*
Respondent Education									
Some School	1	2	3	1	2	3	2	1	2
High School	10	21	19	8	13	14	11	29	26
Technical School	5	9	8	4	8	8	5	9	8
Partial College	20	23	21	19	24	21	22	22	22
College Degree	36	29	35	38	33	40	34	25	30
Postgraduate	28	17	14	29	20	14	26	13	13
Respondent Credit Score									
Lower Than 620	5	10	9	4	7	9	6	14	8
620 to 639	4	5	6	4	6	4	4	3	8
640 to 659	6	5	9	5	6	12	7	4	5
660 to 679	6	8	8	6	7	7	7	9	10
680 to 699	8	10	9	8	10	9	8	10	9
700 to 719	8	9	12	8	10	11	7	9	14
720 to 739	10	10	12	11	9	17	10	11	7
740 or Higher	52	43	35	54	46	31	50	40	39
Median	*744*	*725*	*716*	*747*	*729*	*717*	*739*	*720*	*714*

NA = data not available.

Notes: Attached properties include townhouses, row houses, villas, apartments, and multi-unit dwellings. VantageScore® 3.0 credit score ranges from 300 to 850.

Source: NSMO, 2014

land it is sited on (CFPB, 2014). Because chattel loans generally are not identifiable as mortgages in the NMDB credit files, the NSMO sample misses most chattel loans and thus substantially undercounts loans for manufactured housing. We found the percentage of initial purchases for manufactured housing much lower than for resale or any refinancing transactions for the same. Underrepresentation of manufactured housing may not be as severe for refinance mortgages or purchase mortgages by repeat buyers. Homeowners may be able to obtain mortgage refinancing on manufactured housing originally titled as chattel after a title change if the house is on a permanent foundation and especially if they also own the underlying land.[15]

Generally, mortgage loan amount was lowest in completely rural counties and highest in metro counties. In addition, loans with terms of less than 30 years were more common in completely rural areas. This partially reflects the greater share in completely rural areas of manufactured housing loans, which tend to have shorter terms. Nonetheless, these differences in loan terms persist even when manufactured housing are excluded from the samples. Purchase loans in completely rural counties had a higher median loan-to-value (LTV) ratio.[16]

This is partly explained by the slightly higher share of purchasers in completely rural counties that were first-time homebuyers, who tend to have loans with higher LTVs.[17]

NSMO is a representative sample of mortgages, but the survey is answered by a single respondent who may be one of multiple borrowers. Seventy-five percent of mortgage loans involve multiple borrowers. For simplicity, we refer to the respondents as "borrowers." In characterizing borrowers, we focus on age, household type, number of co-borrowers, race and ethnicity, education, employment, income, and credit scores.

Overall, borrowers in completely rural areas had lower incomes, were less likely to be employed full-time, and were more likely to identify as non-Hispanic White. Age differences across areas were small, though purchasers were slightly younger and refinancers were slightly older in completely rural counties than in metro and non-metro areas. Educational levels differed more noticeably with metro borrowers more likely than others to have a graduate degree. Although median credit scores were only slightly lower in completely rural counties than in more populous areas, the share of borrowers with a credit score of at least 740 was much lower in completely rural counties (35 percent) than in non-metro (43 percent) and metro (52 percent) counties.[18]

---

[15] According to the Manufactured Housing Resource Guide by the National Consumer Law Center, approximately three-quarters of the states have statutes that set forth a procedure to convert a manufactured home from personal to real property and document that conversion. See https://www.nclc.org/images/pdf/manufactured_housing/cfed-titling-homes.pdf.
[16] The LTV ratio is obtained by dividing the mortgage loan amount by the property value, and the difference between 100 percent and the LTV indicates the borrower's share of equity on the property. A typical mortgage will have an LTV of 80 percent, which indicates that the borrower has 20 percent equity in the property.
[17] First-time purchases are purchase mortgages taken out by borrowers who were younger than 55 years of age, who did not have any record of having a prior mortgage in the NMDB data, and who were buying a house they will primarily live in.
[18] VantageScore® 3.0 is the credit scoring model developed jointly by Equifax, Experian, and TransUnion. The NMDB contains VantageScore® 3.0 credit scores that range from 300 to 850. Generally, the higher a borrower's credit score is, the less risky the borrower is assumed to be. See https://your.vantagescore.com/resource/52/understanding-credit-scores.

The smaller share of completely rural borrowers with scores of 740 or greater suggests that loans in completely rural areas were likely to have higher interest rates because the best rates are typically offered to borrowers with scores above 740. To test this conjecture, we examine whether there were statistically significant differences in Freddie Mac's Primary Mortgage Market Survey (PMMS®) spread across the analysis groups (metro, non-metro, and completely rural).[19] We consider both the observed difference in PMMS spreads as well as an adjusted difference based on a regression model that accounts for borrowers, property, and loan attributes:

$$PMMS\_Spread_{i,c} = \alpha + \rho_{CR}\, CR_c + \rho_{NM}\, NM_c + \beta X_i + \varepsilon_{i,c}$$

The equation models the PMMS spread of the loan originated to borrower i in analysis group c ($PMMS\_Spread_{i,c}$) on indicator variables of whether analysis group c is completely rural (CR_c) or non-metro (NM_c), as well as a vector of borrower, property, and loan attributes (X_i). The coefficient estimates for ρ_{CR} and ρ_{NM} indicate the average difference in the PMMS spread for loans originated in completely rural and non-metro counties relative to metro counties.

Exhibit 4 displays the coefficient estimates and indicators of statistical significance for the sample of all mortgages, and then separately for purchase and refinance mortgages. This exhibit also provides the borrower, property, and loan attributes that are used as control variables for estimating the differences between counties.[20] Borrowers in completely rural counties indeed paid a slightly higher interest rate than borrowers in metro areas,[21] and the spread over the PMMS rate for borrowers in completely rural counties was 14 basis points higher than metro areas with a statistically significant difference. The results imply that completely rural borrowers paid 16 basis points more than non-metro borrowers. While 14 basis points appears small for the average of completely rural counties, a mortgage of $100,000 with a 4.00 percent rate that moved to 4.14 percent would cause the monthly payment to rise $8.10, or $97.20 per year, and cost the borrower $2,916.00 over the life of the loan. For purchase mortgages, the differences between completely rural and metro areas was 24 basis points and statistically significant, but the refinance mortgages difference was not significant. Notably, in this regression framework, the PMMS spread does not vary significantly with most borrower characteristics, whereas the purchase flag, loan amount, property type, and credit score are statistically significant predictors of the interest rate spread.

Geographic Differences in Borrowers' Experiences and Knowledge

In examining differences in mortgage borrowers' experience and knowledge by geography, we use the framework of the PMMS equation and exhibit 4 to examine borrowers' self-reported satisfaction, knowledge, and lender selection. More specifically, exhibit 5 shows the results of the NSMO asking borrowers how satisfied they were with several aspects of their mortgage and the

[19] Freddie Mac publishes the average PMMS rate by mortgage term on a weekly basis. The PMMS spread is calculated as the difference between the actual note rate of a mortgage and Freddie Mac's PMMS average prime offer note rate for that term at that time. This spread indicates how expensive a mortgage is compared to the average mortgage of similar term taken out in that week. The spread used in this article is unbounded while the spread in the NSMO public use file is bounded for privacy considerations.

[20] Some of the controls, such as race, ethnicity, age and gender, are characteristics that lenders do not or cannot use in loan pricing models. We include them to account indirectly for unmeasured characteristics that may be correlated with these controls.

[21] The R-squared for the models ranged from 0.06070 to 0.08412.

Exhibit 4

Spread Regression

		All Mortgages	Purchase Mortgages	Refinance Mortgages
Average for Metro (M) (Percentage Points)		0.22	0.19	0.27
Controlled Model				
Parameter	**Categories**	**Estimate**	**Estimate**	**Estimate**
	Intercept	0.22***(0.03)	0.15***(0.03)	0.23***(0.03)
	Non-Metro (NM)	– 0.03 (0.03)	– 0.04 (0.04)	– 0.02 (0.04)
	Completely Rural (CR)	0.14***(0.04)	0.23***(0.05)	0.05 (0.05)
	Purchase Mortgage	– 0.06***(0.02)	NA	NA
Property Type:	Mobile or Manufactured	0.10*(0.06)	0.16*(0.10)	0.06 (0.07)
	Attached	0.06**(0.02)	0.02 (0.03)	0.11***(0.03)
	Land	0.06 (0.12)	0.03 (0.03)	NA
Loan Amount:	Less Than $50,000	0.31***(0.04)	0.37***(0.06)	0.27***(0.06)
	$50,000 to $149,999	0.12***(0.02)	0.13***(0.03)	0.11***(0.03)
	$300,000 or More	– 0.15***(0.02)	– 0.11***(0.03)	– 0.19***(0.03)
Reported Household Income:	Less Than $35,000	0.01 (0.04)	0.008 (0.05)	0.02 (0.05)
	$35,000 to $49,999	– 0.03 (0.03)	– 0.05 (0.04)	0.001 (0.05)
	$50,000 to $74,999	– 0.01 (0.03)	– 0.04 (0.03)	0.02 (0.04)
	$100,000 to $174,999	– 0.01 (0.02)	– 0.004 (0.03)	– 0.03 (0.03)
	$175,000 or More	– 0.07**(0.03)	– 0.11***(0.04)	– 0.04 (0.04)
Household Type:	Single	– 0.001 (0.02)	– 0.007 (0.03)	0.006 (0.03)
Respondent Race/Ethnicity:	Hispanic and Non-White	– 0.03 (0.02)	– 0.03 (0.03)	– 0.02 (0.03)
Respondent Age:	35 or Younger	– 0.01 (0.02)	– 0.002 (0.03)	0.008 (0.04)
	46 to 55	– 0.01 (0.02)	0.05 (0.03)	– 0.08**(0.03)
	56 to 65	– 0.05*(0.03)	– 0.02 (0.03)	– 0.08**(0.03)
	66 or Older	– 0.04 (0.03)	0.03 (0.05)	– 0.10**(0.04)
Respondent Education:	Some School	0.03 (0.07)	0.07 (0.10)	0.008 (0.10)
	High School	0.06**(0.03)	0.06 (0.04)	0.06 (0.04)
	Technical School	0.05 (0.04)	0.05 (0.05)	0.06 (0.05)
	Partial College	0.05**(0.02)	0.07**(0.03)	0.04 (0.03)
	Postgraduate	– 0.02 (0.02)	– 0.01 (0.03)	– 0.02 (0.03)
Respondent Credit Score:	Lower Than 620	0.16***(0.03)	0.05 (0.05)	0.24***(0.05)
	620 to 639	0.06 (0.04)	– 0.05 (0.06)	0.19***(0.06)
	640 to 659	0.12***(0.03)	0.04 (0.05)	0.20***(0.05)
	660 to 679	0.08**(0.03)	0.08*(0.05)	0.08*(0.04)
	680 to 699	0.11***(0.03)	0.06 (0.04)	0.18***(0.04)
	700 to 719	0.01 (0.03)	– 0.02 (0.04)	0.03 (0.04)
	720 to 739	0.04 (0.03)	0.04 (0.04)	0.04 (0.04)
Number of Observations		6,540	3,553	2,987
R-Squared		0.07	0.06	0.08

NA = data not available.

** denotes significance at 10-percent level. ** denotes significance at 5-percent level. *** denotes significance at 1-percent level.*

Notes: The controlled difference model has the following covariates for the respondent: age, race/ethnicity, credit score, and education; and household: household type, annual income, mortgage loan amount, and property type. The controlled difference model intercept reflects the percentages for respondents who were 36 to 45 years old, non-Hispanic White, and in a coupled household; who had a college degree or higher, credit score of 740 or higher, annual household income from $75,000 to $99,999, and mortgage loan amount from $150,000 to $299,999; and who lived in a single-family detached house.

Survey question: "What is the interest rate on this mortgage?"

Source: NSMO, 2014

mortgage process.[22] Most borrowers were "very satisfied" with their mortgage and the mortgage process, but borrowers in completely rural counties were less likely be "very satisfied," as most of the completely rural–metro differences in exhibit 5 are negative and statistically significant.

Exhibit 5

Satisfaction of Borrowers with Mortgage and Mortgage Process

	Percent	Controlled Difference Model		
		County Type Difference (Percentage Point)		
Share "Very Satisfied"	M	NM – M	CR – M	CR – NM
All Mortgages				
Best Terms to Fit Needs	78	2	– 8***	– 10***
Lowest Interest Rates Qualified	70	3	– 3	– 6*
Lowest Closing Costs	58	1	– 6**	– 6*
Lender	76	1	0	– 1
Settlement Agent	70	1	– 4	– 5
Application Process	62	0	– 3	– 3
Loan Closing Process	66	0	– 8***	– 9***
Disclosure Documents	65	1	– 8***	– 8***
Timeliness of Documents	64	1	– 9***	– 11***
Purchase Mortgages				
Best Terms to Fit Needs	78	2	– 7**	– 9**
Lowest Interest Rates Qualified	70	– 1	– 8**	– 7*
Lowest Closing Costs	55	– 2	– 5	– 3
Lender	76	– 3	2	5
Settlement Agent	69	2	– 4	– 6
Application Process	61	– 3	– 5	– 2
Loan Closing Process	64	– 1	– 9**	– 8*
Disclosure Documents	63	– 1	– 8**	– 7
Timeliness of Documents	63	1	– 10***	– 12***
Refinance Mortgages				
Best Terms to Fit Needs	78	3	– 10***	– 13***
Lowest Interest Rates Qualified	70	7**	2	– 5
Lowest Closing Costs	61	3	– 7*	– 10**
Lender	76	5**	– 2	– 8*
Settlement Agent	71	0	– 3	– 4
Application Process	63	3	1	– 4
Loan Closing Process	69	1	– 7*	– 9*
Disclosure Documents	66	3	– 7*	– 10**
Timeliness of Documents	66	1	– 8**	– 10**

CR = completely rural counties. M = metro counties. NM = non-metro counties.
** denotes significance at 10-percent level. ** denotes significance at 5-percent level. *** denotes significance at 1-percent level.*
Notes: The controlled difference model has the following covariates for the respondent: age, race/ethnicity, credit score, and education; and household: household type, annual income, mortgage loan amount, and property type. The controlled difference model intercept reflects the percentages for respondents who were 36 to 45 years old, non-Hispanic White, and in a coupled household; who had a college degree or higher, credit score of 740 or higher, annual household income from $75,000 to $99,999, and mortgage loan amount from $150,000 to $299,999; and who lived in a single-family detached house.
Survey questions: "Overall, how satisfied are you that the mortgage you got was one with the following? Overall, how satisfied are you with the following?"
Source: NSMO, 2014

[22] The survey asked borrowers if satisfaction with the mortgage they got was: (1) the best terms to fit their needs, (2) the lowest interest rate for which they qualified, and (3) the lowest closing costs. The survey asked the borrowers about satisfaction with: (1) their lender or broker, (2) their settlement agent, (3) the application process, (4) the loan closing process, (5) the information in the disclosure documents, and (6) timeliness of mortgage disclosure documents.

The share of borrowers in completely rural counties who were "very satisfied" that they received a mortgage with the best terms to fit their needs was 8 and 10 percentage points lower than the shares in metro counties and non-metro counties, respectively. Completely rural borrowers were also 6 percentage points less likely than those in metro areas to report that they were very satisfied they received the lowest closing costs. Completely rural borrowers were 8 to 9 percentage points less likely to be very satisfied with the closing process, disclosure documents, and timeliness of mortgage documents. This pattern largely holds for refinance mortgages, and for both completely rural/metro and completely rural/non–metro comparisons. For purchase mortgages, these differences are less likely to be significant.

The NSMO data offer several measures of borrowers' expectations at the start of the mortgage process and familiarity with aspects of mortgage lending. These include borrowers' indications of how concerned they were about qualifying for a mortgage and how firm an idea they had about the type of mortgage they wanted. The difference in concern about qualifying for a mortgage between metro, non-metro, and completely rural borrowers was virtually nil. Exhibit 6, however, shows that completely rural borrowers were less likely to have a firm idea about the type of mortgage they wanted than borrowers in the other two areas.

The survey measured borrowers' understanding of the mortgage process by asking about their familiarity with their own credit history; available interest rates and mortgage products; and requirements—such as income and down-payment requirements—to obtain a mortgage.

Exhibit 6

Borrower Concern about Qualifying and Idea about Mortgage Wanted

	Percent	Controlled Difference Model		
		County Type Difference (Percentage Point)		
	M	NM – M	CR – M	CR – NM
All Mortgages				
Not at All Concerned	52	– 1	– 1	– 1
Have Firm Idea	59	0	– 6**	– 6*
Purchase Mortgages				
Not at All Concerned	47	3	1	– 3
Have Firm Idea	54	– 1	– 4	– 3
Refinance Mortgages				
Not at All Concerned	57	– 5*	– 3	2
Have Firm Idea	65	1	– 8**	– 9*

CR = completely rural counties. M = metro counties. NM = non-metro counties.
** denotes significance at 10-percent level. ** denotes significance at 5-percent level.*
Notes: The controlled difference model has the following covariates for the respondent: age, race/ethnicity, credit score, and education; and household: household type, annual income, mortgage loan amount, and property type. The controlled difference model intercept reflects the percentages for respondents who were 36 to 45 years old, non-Hispanic White, and in a coupled household; who had a college degree or higher, credit score of 740 or higher, annual household income from $75,000 to $99,999, and mortgage loan amount from $150,000 to $299,999; and who lived in a single-family detached house.
Survey questions: "When you began the process of getting this mortgage, how concerned were you about qualifying for a mortgage? How firm an idea did you have about the mortgage you wanted?"
Source: NSMO, 2014

Generally, borrowers in completely rural counties were less familiar with aspects of mortgage lending than borrowers in non-metro or metro counties, as shown in exhibit 7. Borrowers were typically very familiar with their credit history with little differences among the areas. Completely rural counties had a smaller fraction of borrowers who were very familiar with interest rates available, the types of mortgages available, the mortgage process, down-payment requirements, and income requirements. Completely rural borrowers and refinancers among them were also less likely to be very familiar with the money required for closing. These results were consistent with the notion that borrowers in completely rural counties have less information or fewer lenders to choose from than borrowers in metro areas.

Exhibit 7

Borrower Familiarity with Aspects of Mortgage Lending

	Percent	Controlled Difference Model		
		County Type Difference (Percentage Point)		
Share "Very Familiar"	M	NM – M	CR – M	CR – NM
All Mortgages				
Credit History or Score	76	2	– 1	– 3
Interest Rates Available	58	1	– 8***	– 9***
Mortgage Types Available	48	– 2	– 9***	– 7**
Mortgage Process	55	0	– 10***	– 11***
Down Payment to Qualify	59	1	– 9***	– 9***
Income Needed to Qualify	57	0	– 8***	– 8**
Money Needed for Closing	51	– 1	– 9***	– 8**
Purchase Mortgages				
Credit History or Score	75	– 1	0	1
Interest Rates Available	54	0	– 8**	– 9*
Mortgage Types Available	45	– 4	– 10***	– 6
Mortgage Process	49	1	– 9**	– 10**
Down Payment to Qualify	60	– 1	– 8**	– 7*
Income Needed to Qualify	54	2	– 6	– 8*
Money Needed for Closing	48	– 4	– 5	– 2
Refinance Mortgages				
Credit History or Score	77	4	– 3	– 7*
Interest Rates Available	64	1	– 8**	– 10**
Mortgage Types Available	51	0	– 8**	– 8*
Mortgage Process	64	– 3	– 13***	– 10**
Down Payment to Qualify	59	0	– 10**	– 10**
Income Needed to Qualify	60	– 3	– 10**	– 7
Money Needed for Closing	55	0	– 13**	– 13**

CR = completely rural counties. M = metro counties. NM = non-metro counties.
** denotes significance at 10-percent level. ** denotes significance at 5-percent level. *** denotes significance at 1-percent level.*
Notes: The controlled difference model has the following covariates for the respondent: age, race/ethnicity, credit score, and education; and household: household type, annual income, mortgage loan amount, and property type. The controlled difference model intercept reflects the percentages for respondents who were 36 to 45 years old, non-Hispanic White, and in a coupled household; who had a college degree or higher, credit score of 740 or higher, annual household income from $75,000 to $99,999, and mortgage loan amount from $150,000 to $299,999; and who lived in a single-family detached house.
Survey question: "When you began the process of getting this mortgage, how familiar were you with each of the following?"
Source: NSMO, 2014

The NSMO also probed borrowers on their knowledge about mortgage concepts by asking them how well they could explain the concepts to someone else. Based on borrowers' responses, the mortgage concepts from the least to the most challenging are:

(1) difference between fixed and adjustable rates,
(2) consequences of not making required payments,
(3) difference between interest rate and annual percentage rate (APR),
(4) amortization of a loan, and
(5) difference between prime and subprime loans.

We classify the first two as simple concepts and the last three as complex concepts.

Overall, borrowers were more knowledgeable about simple concepts than they were about complex concepts as shown in exhibit 8. For the simple concept of adjustable versus fixed-rate mortgages, borrowers in completely rural counties were less likely to say they could explain the

Exhibit 8

Borrowers' Abilities to Explain Aspects of Mortgages

Share Able to Explain "Very Well"	Percent	Controlled Difference Model		
		County Type Difference (Percentage Point)		
	M	NM – M	CR – M	CR – NM
All Mortgages				
Fixed Versus Adjustable Rate	70	– 1	– 8***	– 7**
Consequence of Not Paying	67	1	– 3	– 3
Amortization of Loan	39	0	– 6**	– 5*
Interest Rate Versus APR	28	0	– 4	– 4
Prime Versus Subprime	22	– 1	– 2	– 1
Purchase Mortgages				
Fixed Versus Adjustable Rate	67	1	– 1	– 2
Consequence of Not Paying	66	2	1	0
Amortization of Loan	36	– 3	– 5	– 3
Interest Rate Versus APR	26	– 3	– 4	– 1
Prime Versus Subprime	20	– 2	0	2
Refinance Mortgages				
Fixed Versus Adjustable Rate	72	– 4	– 17***	– 13***
Consequence of Not Paying	68	– 1	– 9**	– 7
Amortization of Loan	41	2	– 5	– 7
Interest Rate Versus APR	31	4	– 3	– 7*
Prime Versus Subprime	24	2	– 4	– 5

APR = annual percentage rate. CR = completely rural counties. M = metro counties. NM = non-metro counties.
** denotes significance at 10-percent level. ** denotes significance at 5-percent level. *** denotes significance at 1-percent level.*
Notes: The controlled difference model has the following covariates for the respondent: age, race/ethnicity, credit score, and education; and household: household type, annual income, mortgage loan amount, and property type. The controlled difference model intercept reflects the percentages for respondents who were 36 to 45 years old, non-Hispanic White, and in a coupled household; who had a college degree or higher, credit score of 740 or higher, annual household income from $75,000 to $99,999, and mortgage loan amount from $150,000 to $299,999; and who lived in a single-family detached house.
Survey question: "How well could you explain to someone the following?"
Source: NSMO, 2014

differences very well. After adjusting for differences in characteristics of borrowers and loans, the share of borrowers in completely rural counties who could explain the differences very well was 8 percentage points lower than those in metro areas, and 7 percentage points lower than in non-metro counties. Refinancing transactions drove the overall differences, as the share of refinancers in completely rural counties able to explain the difference between adjustable and fixed rates very well was 17 percentage points lower than those in metro areas and 13 percentage points lower than those in non-metro counties. Fewer borrowers reported familiarity with complex mortgage concepts and there were no significant differences among the geographic areas.

The NSMO asked several questions about the shopping and the mortgage application process, namely, whether the borrower:

(1) picked the lender or broker before the loan;
(2) applied directly to a lender (as opposed to through a broker or a builder);
(3) initiated contact with the lender or broker (as opposed to those who were contacted by the lender or broker first, or those were who were put in touch with the lender or broker by a third party);
(4) seriously considered multiple lenders and brokers; and
(5) applied to multiple lenders or brokers.

We interpret the first three items as reflecting how proactive borrowers were with lender selection. The last two items reflect how much borrowers shopped across multiple lenders.

Exhibit 9 shows that most borrowers were proactive with lender selection—they picked the lender before the loan, initiated contact, and applied directly with a lender. With respect to shopping behavior, more than half of borrowers seriously considered applying to multiple lenders, but only one-fourth or fewer applied to more than one lender.

Outside of metro areas, borrowers were more proactive in initiating contact and applying directly to a lender. Completely rural borrowers seeking a loan for home purchase were more proactive than their home-purchase counterparts in metro areas, as they were 13 and 16 percentage points more likely to have applied directly or to have initiated contact, respectively. These borrowers may have fewer options for lenders in their area as solicitation by lenders may be less common in completely rural areas or lenders in those areas may have less competition.

These differences, however, do not translate into differences in the likelihood that a completely rural borrower considered or applied to multiple lenders. Borrowers who refinanced in non-metro counties were more proactive than those in metro areas, but they did not show a different level of engagement with multiple lenders than other areas.

The NSMO asked borrowers how likely they were to: sell their property, move but keep the property, refinance, or pay off the mortgage to have a mortgage-free property in the next couple of years. Borrowers' responses demonstrated their expectations about their property and mortgage. In our analysis, we combined borrowers who were very likely to take an action together with those who were somewhat likely to take that action and referred to these borrowers as generally likely to take that action.

Borrowers were generally very unlikely to take any of these actions, as seen in exhibit 10. Borrowers who refinanced were the most likely to report any of these anticipated actions, with 40

Exhibit 9

Borrowers' Mortgage Shopping and Application Steps

	Percent	Controlled Difference Model		
		County Type Difference (Percentage Point)		
	M	NM – M	CR – M	CR – NM
All Mortgages				
Picked Lender Before Loan	71	0	4	4
Applied Directly to a Lender	63	9***	10***	1
Borrower Initiated Contact	66	8***	12***	4
Considered Multiple Lenders	52	– 1	0	1
Applied to Multiple Lenders	23	– 2	0	3
Purchase Mortgages				
Picked Lender Before Loan	73	0	3	3
Applied Directly to a Lender	58	9***	13***	4
Borrower Initiated Contact	64	10***	16***	6
Considered Multiple Lenders	55	– 1	3	4
Applied to Multiple Lenders	26	– 5*	1	6
Refinance Mortgages				
Picked Lender Before Loan	68	1	5	4
Applied Directly to a Lender	69	10***	9**	– 1
Borrower Initiated Contact	69	5*	7**	3
Considered Multiple Lenders	49	– 2	– 4	– 2
Applied to Multiple Lenders	19	0	– 1	– 1

CR = completely rural counties. M = metro counties. NM = non-metro counties.
** denotes significance at 10-percent level. ** denotes significance at 5-percent level.*
Notes: The controlled difference model has the following covariates for the respondent: age, race/ethnicity, credit score, and education; and household: household type, annual income, mortgage loan amount, and property type. The controlled difference model intercept reflects the percentages for respondents who were 36 to 45 years old, non-Hispanic White, and in a coupled household; who had a college degree or higher, credit score of 740 or higher, annual household income from $75,000 to $99,999, and mortgage loan amount from $150,000 to $299,999; and who lived in a single-family detached house. Survey questions: "Which of the following best describes your shopping process? How did you apply for the mortgage? Who initiated first contact between you and the lender/mortgage broker you used for the mortgage you took out? How many different lenders/mortgage brokers did you seriously consider before choosing where to apply for this mortgage? How many different lenders/mortgage brokers did you end up applying to?"
Source: NSMO, 2014

Exhibit 10

Borrowers' Mortgage and Property Expectations (1 of 2)

Share Stating "Likely" to	Percent	Controlled Difference Model		
		County Type Difference (Percentage Point)		
	M	NM – M	CR – M	CR – NM
All Mortgages				
Sell Property	34	– 3*	– 2	2
Move but Keep Property	23	– 2	– 1	1
Refinance Mortgage	31	– 3*	– 2	1
Have Mortgage-Free Property	22	2	2	0
Purchase Mortgages				
Sell Property	30	0	5	5
Move but Keep Property	23	– 2	4	6
Refinance Mortgage	34	– 1	– 5	– 4
Have Mortgage-Free Property	21	0	2	2

Exhibit 10

Borrowers' Mortgage and Property Expectations (2 of 2)

Share Stating "Likely" to	Percent	Controlled Difference Model		
		County Type Difference (Percentage Point)		
	M	NM – M	CR – M	CR – NM
Refinance Mortgages				
Sell Property	40	– 7**	– 9**	– 2
Move but Keep Property	23	– 3	– 6*	– 3
Refinance Mortgage	28	– 6**	0	6
Have Mortgage-Free Property	22	4	2	– 1

CR = completely rural counties. M = metro counties. NM = non-metro counties.
** denotes significance at 10-percent level. ** denotes significance at 5-percent level. *** denotes significance at 1-percent level.*
Notes: The controlled difference model has the following covariates for the respondent: age, race/ethnicity, credit score, and education; and household: household type, annual income, mortgage loan amount, and property type. The controlled difference model intercept reflects the percentages for respondents who were 36 to 45 years old, non-Hispanic White, and in a coupled household; who had a college degree or higher, credit score of 740 or higher, annual household income from $75,000 to $99,999, and mortgage loan amount from $150,000 to $299,999; and who lived in a single-family detached house.
Survey question: "How likely is it that in the next couple of years you will do the following?"
Source: NSMO, 2014

percent reporting they might sell the property in the next few years. Refinancers in completely rural counties were 9 percentage points less likely to sell their property in the next few years. Refinance borrowers in non-metro counties were 7 percentage points less likely than borrowers in metro areas to anticipate selling their homes, and were also less likely to report that they expect to refinance again in the next few years compared to both metro and completely rural refinancers.

Geographic Differences in Borrowers' Lender and Mortgage Choice

The detail of the question on how borrowers chose their lenders is an advantage of the NSMO survey. NSMO asked specifically about the importance of seven factors:

(1) reputation of the lender or broker,
(2) having an established banking relationship with the lender or broker,
(3) having a local office or branch of the lender or broker nearby,
(4) recommendation from a friend, relative or co-worker,
(5) recommendation from a real estate agent or builder, and
(6) whether the lender or broker was a friend.

Unfortunately, the survey questions changed from a three-point scale of "very," "somewhat," and "not at all" to a two-point scale of "important" and "not important" in the seventh wave of the survey when the supplemental sample was administered along with mostly 2015 mortgages from the regular sample. Consequently, we examined lender and mortgage choice with a different comparison group than that used in the earlier part of the article. Supplemental sample borrowers (267) were combined with regular sample borrowers who took out their mortgage in 2015 (6,199) to create an analysis file of borrowers who answered the question in the two-point scale. This led to a sample size of 6,466 mortgages with 5,508 (85 percent) in metro, 618 (10 percent) in non-metro, and 340 in completely rural counties shown in exhibits 11 and 12.[23]

[23] These county shares are comparable to the ones for the 2014 sample, reported in exhibit 2.

Borrowers most often identified lender reputation as an important factor for lender selection, followed by having an established banking relationship and a local office or branch. Having an established banking relationship was more important to borrowers in non-metro and completely rural counties than in metro areas, shown in exhibit 11. In contrast, metro borrowers felt the agent or builder recommendation was more important. Differences were even more stark in the importance of an established banking relationship and agent or builder recommendation for purchase mortgages. Furthermore, non-metro borrowers also stated that having a local office or branch was important more frequently than metro borrowers, particularly for refinancers.

Exhibit 11

Importance of Factors in Choosing a Lender

Share Stating Factor "Important"	Percent	Controlled Difference Model		
		County Type Difference (Percentage Point)		
	M	NM – M	CR – M	CR – NM
All Mortgages				
Reputation	71	– 4*	0	3
Established Bank Relationship	55	9***	9***	0
Local Office or Branch	49	8***	5	– 3
Friend/Relative Recommended	40	– 2	1	3
Agent/Builder Recommended	36	– 6***	– 12***	– 6**
Was a Friend or Relative	14	0	– 3	– 3
Purchase Mortgages				
Reputation	72	– 4	– 4	0
Established Bank Relationship	50	12***	19***	7
Local Office or Branch	55	5*	0	– 5
Friend/Relative Recommended	48	– 1	– 5	– 3
Agent/Builder Recommended	55	– 11***	– 21***	– 10**
Was a Friend or Relative	15	– 1	– 4	– 3
Refinance Mortgages				
Reputation	71	– 3	2	5
Established Bank Relationship	60	7**	0	– 7
Local Office or Branch	44	10***	7*	– 4
Friend/Relative Recommended	32	– 3	6*	9**
Agent/Builder Recommended	17	– 2	– 5*	– 3
Was a Friend or Relative	14	1	– 1	– 3

CR = completely rural counties. M = metro counties. NM = non-metro counties.
** denotes significance at 10-percent level. ** denotes significance at 5-percent level. *** denotes significance at 1-percent level.*
Notes: Given inclusion of 2015 NSMO observations, sample sizes for exhibits 8 and 9 differ from those in remaining tables. Sample sizes for this exhibit are: 5,508 for M, 618 for NM, and 340 for CR. The controlled difference model has the following covariates for the respondent: age, race/ethnicity, credit score, and education; and household: household type, annual income, mortgage loan amount, and property type. The controlled difference model intercept reflects the percentages for respondents who were 36 to 45 years old, non-Hispanic White, and in a coupled household; who had a college degree or higher, credit score of 740 or higher, annual household income from $75,000 to $99,999, and mortgage loan amount from $150,000 to $299,999; and who lived in a single-family detached house.
Survey question: "How important were each of the following in choosing the lender/mortgage broker you used for the mortgage you took out?"
Source: NSMO, 2014-2015

NSMO also asked about the importance of seven factors in deciding on a mortgage:

(1) lower interest rate,
(2) fixed-interest rate for the life of the loan,
(3) lower annual percentage rate (APR),
(4) lower closing fees,
(5) lower monthly payment,
(6) a term of 30 years, and
(7) no mortgage insurance.

Exhibit 12 shows the factors for selecting a mortgage are listed in the order of importance with near universal agreement that getting a lower interest rate was very important. For all mortgages, one difference between borrowers in non-metro areas relative to metro borrowers was their ranking of importance of having a lower monthly payment and not having mortgage insurance. The shares reporting these features as being important for mortgage selection in non-metro counties was between 4 and 6 percentage points lower than in metro counties. Completely rural purchasers were 7 percentage points less likely to indicate that having lower monthly payment was important than metro purchasers. Completely rural purchasers were also 8 to 10 percentage points less likely to indicate that a mortgage term of 30 years was important than non-metro and metro purchasers, respectively.

Exhibit 12

Importance of Factors in Choosing a Mortgage (1 of 2)

	Percent	Controlled Difference Model		
		County Type Difference (Percentage Point)		
Share "Very Familiar"	M	NM – M	CR – M	CR – NM
All Mortgages				
Lower Interest Rate	98	0	0	0
Fixed-Interest Rate	89	1	– 1	– 3
Lower APR	88	– 1	2	3
Lower Closing Fees	84	0	0	1
Lower Monthly Payment	80	– 4**	– 3	1
30-Year Term	61	– 2	– 3	0
No Mortgage Insurance	58	– 6***	– 2	4
Purchase Mortgages				
Lower Interest Rate	97	– 1	0	1
Fixed-Interest Rate	89	0	– 1	– 1
Lower APR	87	0	1	1
Lower Closing Fees	82	1	– 2	– 3
Lower Monthly Payment	81	– 3	– 7**	– 4
30-Year Term	69	– 2	– 10**	– 8*
No Mortgage Insurance	54	– 6*	– 2	5

Exhibit 12

Importance of Factors in Choosing a Mortgage (2 of 2)

	Percent	Controlled Difference Model		
		County Type Difference (Percentage Point)		
Share "Very Familiar"	M	NM – M	CR – M	CR – NM
Refinance Mortgages				
Lower Interest Rate	98	1	0	– 1
Fixed-Interest Rate	90	2	– 1	– 3
Lower APR	88	– 2	3	6*
Lower Closing Fees	85	– 1	3	4
Lower Monthly Payment	80	– 4	1	5
30-Year Term	52	– 3	3	6
No Mortgage Insurance	62	– 6*	– 1	5

APR = annual percentage rate. CR = completely rural counties. M = metro counties. NM = non-metro counties.
** denotes significance at 10-percent level. ** denotes significance at 5-percent level. *** denotes significance at 1-percent level.*
Notes: Given inclusion of 2015 NSMO observations, sample sizes for exhibits 8 and 9 differ from those in remaining tables. Sample sizes for this exhibit are: 5,508 for M, 618 for NM, and 340 for CR. The controlled difference model has the following covariates for the respondent: age, race/ethnicity, credit score, and education; and household: household type, annual income, mortgage loan amount, and property type. The controlled difference model intercept reflects the percentages for respondents who were 36 to 45 years old, non-Hispanic White, and in a coupled household; who had a college degree or higher, credit score of 740 or higher, annual household income from $75,000 to $99,999, and mortgage loan amount from $150,000 to $299,999; and who lived in a single-family detached house.
Survey question: "How important were each of the following in determining the mortgage you took out?"
Sources: NSMO, 2014-2015

Conclusions

Rural credit markets are commonly viewed as differing from those in more populous areas for several reasons, including a greater share of smaller and locally focused lenders. This article offers, to our knowledge, the first comprehensive analysis that contrasts mortgage borrowers' expectations, knowledge, and outcomes in completely rural areas to those of borrowers in metro and non-metro areas. To do so, we take advantage of survey data from the NSMO, including a special sample of mortgage borrowers in completely rural counties.

The comparisons we provide do not paint a simple picture of how borrowers and mortgage markets differ by geographic location, as completely rural borrowers differ from other borrowers on some dimensions but not others. Nonetheless, the results provide some suggestive evidence that completely rural borrowers may have been initially less familiar with the mortgage market conditions, products, and requirements. Completely rural borrowers were significantly less likely than borrowers in metro areas to say they were very familiar with elements of the mortgage process—such as the down-payment requirements—when they began it. There were few significant differences by geographic location, however, in borrowers' self-reported level of concern in qualifying for a mortgage, or in their ability to explain specific mortgage features to someone else at the time of the survey. It is not clear how to reconcile these findings, but one possibility is that completely rural borrowers were less familiar at the outset because they expected to or relied more heavily on the lender to qualify for a loan and to become familiar with the details of getting a mortgage.

Our results may also point to a relatively greater importance of borrower-lender relationships outside of metro areas. For example, borrowers in completely rural and non-metro areas were more likely than those in metro areas to rate having an established relationship as important in choosing their lender. Additionally, completely rural borrowers were similarly satisfied with the lender, settlement agent, and the application process as were borrowers in more populous areas, even though they paid somewhat higher interest rates and were less likely to be very satisfied with the mortgage they obtained, or with the closing process.

We conclude this article with a note of caution. Results presented in this article are for three groups of counties as a whole. Because counties are very heterogeneous in each of the three groups, the results do not necessarily apply to any individual county.

Acknowledgements

The authors thank Robert B. Avery, Ron Borzekowski, Kenneth P. Brevoort, Lariece Brown, Brian Bucks, Daniel E. Coates, Thomas Daula, Michael Eriksen, Samuel Frumkin, Ian H. Keith, Michael LaCour-Little, Forrest W. Pafenberg, Shiv Rawal, David Sanchez, Susan Singer, and Rebecca Sullivan for providing comments and Julia Nguyen for assisting with data.

Authors

Tim Critchfield is an interdisciplinary statistician in the Office of Research at the Consumer Financial Protection Bureau.

Jaya Dey is a senior economist with the Affordable Lending Analytics and Research group at Freddie Mac.

Nuno Mota is an economist with the Economic and Strategic Research group at Fannie Mae.

Saty Patrabansh is the manager of the National Mortgage Database Program at the Federal Housing Finance Agency.

References

Ambrose, Brent, and Richard Buttimer. 2005. "GSE Impact on Rural Mortgage Markets," *Regional Science and Urban Economics* 35 (4): 417–443.

Avery, Robert B., Kenneth P. Brevoort, Tim Critchfield, Ian H. Keith, Ismail E. Mohamed, Forrest W. Pafenberg, Saty Patrabansh, Jay D. Schultz, and Claudia E. Wood. 2017. *National Mortgage Database Technical Report 1.2*. Federal Housing Finance Agency and the Consumer Financial Protection Bureau. https://www.fhfa.gov/PolicyProgramsResearch/Programs/Documents/NMDB-Technical-Report_1.2_10302017.pdf.

Avery, Robert B., Mary F. Bilinski, Audrey Clement, Tim Critchfield, Ian H. Keith, Ismail E. Mohamed, Saty Patrabansh, Chad Redmer, Jay D. Schultz, and Rebecca Sullivan. 2018. *National Survey of Mortgage Originations Technical Documentation: Technical Report 2*. Federal Housing Finance Agency and the Consumer Financial Protection Bureau, National Mortgage Database Program. https://www.fhfa.gov/DataTools/Downloads/Documents/NSMO-Public-Use-Files/NSMO-Technical-Documentation-20181108.pdf.

Bucks, Brian, and Karen Pence. 2008. "Do Borrowers Know Their Mortgage Terms?" *Journal of Urban Economics* 64 (2): 218–233.

Cai, Qiang, and Sarah Shahdad. 2015. *What is the Mortgage Shopping Experience of Today's Homebuyer? Lessons from Recent Fannie Mae Acquisitions*, Fannie Mae: Perspectives. http://fanniemae.com/portal/research-insights/perspectives/041315-cai-shahdad.html.

Consumer Financial Protection Bureau (CFPB). 2014. *Manufactured-Housing Consumer Finance in the United States*. Washington, DC: Consumer Financial Protection Bureau. http://files.consumerfinance.gov/f/201409_cfpb_report_manufactured-housing.pdf.

Critchfield, Tim, Tyler Davis, Lee Davison, Heather Gratton, George Hanc, and Katherine Samolyk. 2004. "The Future of Banking in America – Community Banks: Their Recent Past, Current Performance, and Future Prospects" [Vol. 16, No. 3], FDIC Banking Review Series. https://www.fdic.gov/bank/analytical/banking/2005jan/article1.pdf.

Huang, Li-Ning, Michael LaCour-Little, Mark Palim, and Ying Pan. 2017. "How Much do U.S. Households Know About Qualifying for a Mortgage?" Working Paper, May 31 2017. Washington, DC: Fannie Mae. http://fanniemae.com/resources/file/research/datanotes/pdf/consumer-knowledge-working-paper.pdf.

Housing Assistance Council (HAC). 2012. *Taking Stock: Rural People, Poverty, and Housing in the 21st Century*. Washington, DC: Housing Assistance Council. http://www.ruralhome.org/storage/documents/ts2010/ts_full_report.pdf.

MacDonald, Heather I. 2001. "Fannie Mae and Freddie Mac in Nonmetropolitan Housing Markets: Does Space Matter?" *Cityscape* 5 (3): 219–264.

Mota, Nuno. 2016. Rural Mortgage Lending Over the Last Decade. Working paper 10.25.2016. Washington, DC: Fannie Mae. http://fanniemae.com/resources/file/research/datanotes/pdf/working-paper-102716.pdf.

U.S. Department of Agriculture (USDA). 2016. "Rural America at a Glance, 2016 Edition." *Economic Information Bulletin* 162. Washington, DC: U.S Department of Agriculture, Economic Research Division. https://www.ers.usda.gov/webdocs/publications/80894/eib-162.pdf.

Vandell, Kerry D. 1997. "Improving Secondary Markets in Rural America." In *Proceedings – Rural and Agricultural Conferences: Financing Rural America*, April: 85–120. Kansas City, MO: Federal Reserve Bank of Kansas City.

Perceptions and Expectations of Mortgage Borrowers: New Evidence from the National Survey of Mortgage Originations

Chad Redmer
United States Naval Academy

The views expressed in this paper are those of the author and should not be attributed to the United States Naval Academy or Department of Defense. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Abstract

Research on house price expectations has been limited by a lack of data which have meaningful cross-sectional variation in home purchase and refinance experiences across respondents. The National Survey of Mortgage Originations (NSMO) fills this gap and allows for the exploration of whether members of different demographic groups exhibit different levels of knowledge of house price changes in their local areas and whether they have different levels of capacity to anticipate future price changes as a result. This study finds that first-time homebuyers, a demographic group considered potentially vulnerable during the origination process, are consistently more aware of recent and impending house price trends than repeat purchasers. This study also shows that borrowers with lower incomes have less awareness about price trends than borrowers with higher incomes, and less-educated borrowers have less awareness about price trends than borrowers with higher levels of education. Additionally, this article provides evidence that expectations of future house price changes in consumers' local markets have informational value for estimating future house prices beyond the value provided by local economic data. This study provides a basis for understanding whether consumers who are less knowledgeable about house prices in their area may make poorer decisions regarding their mortgage or home purchase and may require more support or protection as a result.

Introduction

Economists understand that consumer expectations of house prices can play a role in the real economy. Several studies have shown that expectations have predictive information in the aggregate. What is less understood is whether differences in cross-sectional house price expectations can help predict cross-sectional variations in outcomes and whether there is variation in the relationship across different demographic groups. That is, do members of different demographic groups exhibit more or less knowledge of house price changes in their local areas, and do they have more or less capacity to forecast future price changes as a result? This article addresses these two questions.

These questions are important not just for macroeconomic purposes, but for consumer well-being and policy formulation as well. Consumers who are less knowledgeable about house prices in their area may make poorer decisions regarding their mortgage or house purchase. Because housing is a significant portion of overall wealth for most households, the accuracy of perceptions and expectations regarding house price trends could have lasting effects on the lifecycle consumption decisions of households.

This article addresses the questions posed earlier using previously unavailable survey data from the newly released NSMO, which is based on the National Mortgage Database (NMDB®).[1] The survey database is large enough to provide meaningful cross-sectional variation in home purchase and refinance experiences across respondents. The survey offers a rich set of questions about house prices, including what respondents think has happened to prices in their local area (backcasting) and what respondents expect to happen (forecasting). Because time has passed since the responses reflected in the publicly released NSMO data were gathered, it is possible to compare both respondents' backcasts and forecasts with actual house price changes in their areas.

Another advantage of NSMO is that it surveys only borrowers who have just obtained a mortgage. These borrowers are likely the most motivated and informed consumers among the overall population of homeowners and their answers should be the most accurate since they experienced the mortgage process recently. This makes NSMO better for comparisons of the accuracy of backcasts and forecasts across groups than surveys of the whole population. House price expectation data from comprehensive national surveys suffer from the fact that many respondents may have little need to be informed since they are not active participants in the housing market. The emphasis on borrowers active in the mortgage market is also important because these are the consumers for whom consumer protection would have been most relevant. The identification of subgroups who turn out to be systematically less informed may be particularly important in designing consumer-focused policies.

The remainder of this article is organized as follows. The next section provides background information on the literature and data used for the analysis. The two sections after that provide results for backcasting and forecasting, respectively. The section following those examines differences across groups in how they form expectations. The final section concludes.

[1] NSMO and NMDB are described in the guest editors' introduction.

Background

Despite the importance of understanding borrowers' perceptions and expectations concerning local house prices, little economic literature compares them with the price changes that actually occurred in the relevant local areas. Case et al. (2015) evaluated survey responses regarding house price trend perceptions and expectations collected from recent homebuyers in four specific housing markets in the same season annually over the 2003-through-2014 period. The responses are compared to changes in the Case-Shiller Home Price Index over the same period for those four markets. Niu and van Soest (2014) used national survey responses from the RAND American Life Panel, conducted quarterly from 2009 through 2013, to compare state level responses concerning future price expectations with the Federal Housing Finance Agency (FHFA) House Price Index (HPI) state-level indices. The latter study does not focus on respondents most likely to be interested in house prices, and neither study uses a sample that is representative.

This article extends the existing literature using responses to NSMO, which was specifically designed to provide detailed information on borrower perceptions and expectations. In addition to questions concerning mortgage shopping behavior and mortgage closing experience, NSMO includes questions concerning borrowers' perceptions of recent house price trends in their neighborhood and their expectations of future house price trends. The survey also contains detailed information on borrower demographics and several questions which solicit the borrower's appetite for financial risk, concerns about loan qualification, and knowledge about mortgages.

Using NSMO responses from borrowers who originated mortgages from 2013 through 2016, this article compares respondents' perceptions of what happened to house prices in their local area with what actually occurred prior to the survey, and compares what they forecasted would happen with what actually did occur after the survey. The challenge in this exercise is to use a definition of area (neighborhood) comparable to that used by respondents and which also has available accurate summary information on house price changes. For this article, the county is used as the representation of "neighborhood," and actual house price changes are derived from newly available FHFA county-level house price indices. This is less than ideal, as it defines an area which is surely larger than that generally meant by respondents when they refer to a neighborhood. Indices, however, computed at more granular geographic levels (often based on very few observations) would include too much unexplained variation for a meaningful analysis of differences. In addition, the FHFA county indices are available at the quarterly level, which is generally not true for indices with more granular geography. This is important because of the relatively short window ("couple of years") used in the NSMO questionnaire for items regarding past and future house price change perceptions and expectations.

Another advantage in using a county-level definition of local area is that it allows for use of other aggregate measures of economic activity including population change, unemployment rates, per capita income changes, mortgage delinquency, and mortgage volumes. The metrics measuring mortgage delinquency and volume are new and are constructed from the NMDB. These aggregates serve as estimates of the "local economic fundamentals" driving county-level house price changes. This allows for comparisons of respondent perceptions and forecasts with actual changes at both the gross level and when controlling for these fundamentals.

The statistics for the specific county-level variables used in this analysis are presented in exhibit 1. NSMO data used in this analysis were solicited in 15 different waves of a quarterly survey. Thus, there is considerable variation in the timing of when respondents provided their answers (generally 6 to 9 months after origination). The county-level HPI estimates were available quarterly and thus could be aligned to when the survey was answered. Data regarding the county-level seriously delinquent rate and average mortgage loan amount were also available quarterly, both in terms of levels and changes.[2] Percent change in population, percent change in per capita income, and the level and change in the unemployment rate, however, were only available annually. Perforce, this implies alignment issues in their use, which is partially accounted for by using quarterly dummy variables for time of response in the analysis. Another consequence of the availability of data is a slight reduction in the number of observations that could be used for analysis. Population and unemployment data were not available for 72 observations in the full NSMO sample and thus those respondents were not used. An additional 2,928 observations were dropped for the forecasting section of the analysis since they were interviewed in the most recent waves of the survey considered for this study, and therefore, the 2 years of future price changes data were not available at the time of this analysis.

Exhibit 1

Summary Statistics for County HPI and Economic Data

Local Economic Fundamental Variable	N	25th	50th	75th
Previous 2-Year County HPI Growth	24,775	4.46	9.34	17.29
Future 2-Year County HPI Growth*	21,847	6.73	10.97	16.10
Previous 2-Year County Population Growth	24,775	0.46	1.66	3.08
Previous 2-Year County Income Growth	24,775	4.55	6.38	8.61
Previous 2-Year County Mortgage Balance Growth	24,775	0.48	2.39	4.87
2-Year Lagged County Unemployment Rate	24,775	5.17	6.43	7.87
Change in 2-Year Lagged County Unemployment Rate	24,775	– 1.17	– 0.76	– 0.41
2-Year Lagged County Serious Delinquency Rate	24,775	2.11	3.18	4.66
Change in 2-Year Lagged County Serious Delinquency Rate	24,775	– 1.32	– 0.70	– 0.36

HPI = House Price Index.
*Note: *2-year forward HPI comparison data is not yet available for surveys completed in 2017Q1 or 2017Q2.*

[2] A mortgage is considered to be seriously delinquent if it has been delinquent for 90 days or more or has entered the bankruptcy or foreclosure process.

Borrower Backcasts of House Price Changes

Borrowers were asked, "In the last couple of years, how have house prices changed in the neighborhood where this property is located?"—a question referred to as a "backcast" in this study. The distribution of their responses is presented in exhibit 2. The distribution of actual percent change in county HPI over the 2 years prior to the quarter in which the borrower completed the survey is presented in exhibit 3. The exhibits demonstrate that both borrower backcasts and county-level HPIs skewed heavily toward price increases, though in both instances there is some evidence of price decreases.

The alignment between borrower backcasting of local house price trends and actual house price trends is examined in two ways: first, with a simple correlation of the percent change in respondents' county HPI with their backcast of local house price trends for corresponding prior 2-year periods, and second, with a multivariate analysis examining the conditional correlation of actual changes with respondents' backcasts, controlling for local economic fundamentals measured over the same 2-year period. In both the simple and multivariate backcast tests, backcast information is represented by two dummy variables for the responses "Significant Increase" and "Significant Decrease," with "Little/No Change" treated as the base response.

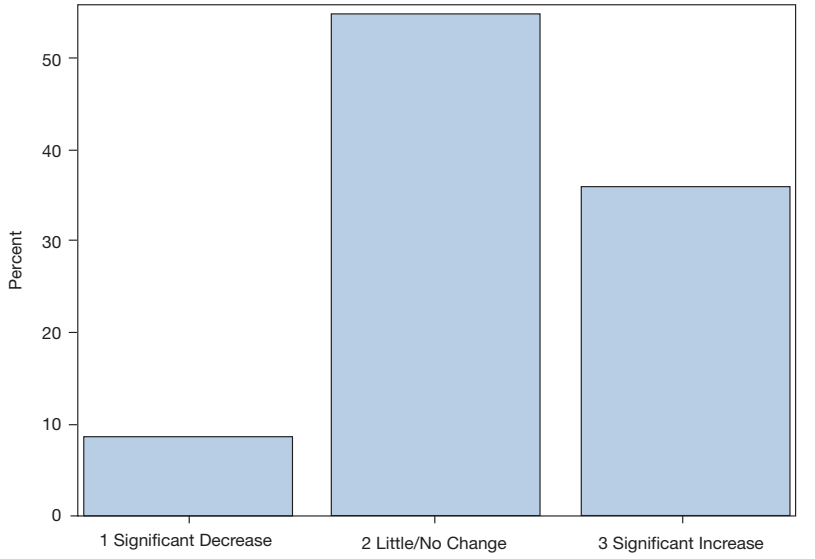
The two analyses test different hypotheses. The simple analysis tests how aware borrowers are of general changes in house prices. The multivariate analysis tests a more subtle question: do the borrowers possess information on actual local house price changes over and above that which could be explained by the other economic variables? Additionally, both analyses test how this information varies across groups.

Exhibit 4 presents estimates from the simple and multivariate backcast tests. The highly significant coefficients on the two dummy variables in the simple test have appropriate signs but are unbalanced in their effects. The coefficient on "Significant Increase" is nearly seven times larger than the coefficient on "Significant Decrease" in absolute value, illustrating that borrowers classify relatively smaller declines in prices as significant.

In the multivariate regression, all control variables are significant, providing evidence that the selected fundamentals are good indicators of actual house price growth. Importantly, the coefficients on the response dummy variables are still significant, albeit with somewhat smaller magnitude for "Significant Increase." This indicates that borrowers have additional knowledge, beyond that gained from local economic conditions, that shapes their backcast of recent local house price trends. One possible explanation is that a borrower may be aware of a more recent appraisal or sale price of one of their neighbor's homes that provides them with an information advantage over just the macroeconomic variables.

Exhibit 2

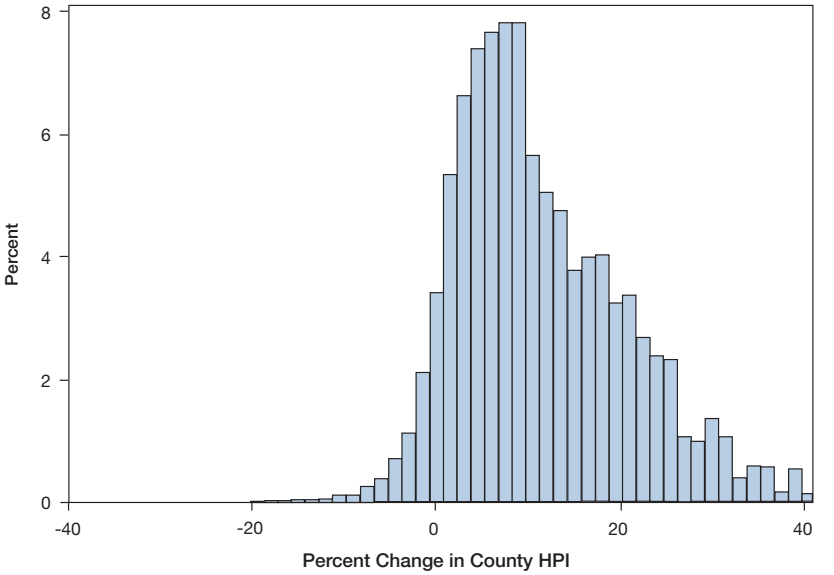
Perception of Recent Trends in Neighborhood House Prices



Note: Distribution of responses to the question, "In the last couple of years, how have house prices changed in the neighborhood where this property is located?".
Source: NSMO

Exhibit 3

Distribution of Actual Recent Trends in County House Prices



HPI = House Price Index.
Note: Distribution of actual change in county house prices in the 2 years prior to the quarter of survey completion.
Source: FHFA HPI

Exhibit 4

Parameter Estimates for Simple and Multivariate Backcast Regressions

	Backcast Alignment with Actual House Price Trends	
Parameter	**Simple**	**Multivariate**
N	24,775	24,775
R^2	0.0908	0.4838
Intercept	9.34	– 1.27*
	(0.08)	(0.53)
"Significant Increase"	5.86***	3.16***
	(0.12)	(0.10)
"Significant Decrease"	– 0.84***	– 0.74***
	(0.21)	(0.16)
Population Growth	-	1.82***
		(0.02)
Income Growth	-	0.33***
		(0.01)
Unemployment Level	-	0.77***
		(0.03)
Change in Unemployment	-	– 3.56***
		(0.11)
Average Mortgage Balance	-	0.10***
		(0.01)
Delinquency Rate	-	– 0.74***
		(0.03)
Change in Delinquency Rate	-	– 3.38***
		(0.06)
Survey Quarter Fixed Effects	**No**	**Yes**

** = p<0.05; ** = p<0.01; *** = p<0.001.*
Note: Standard errors are given in parentheses.

As indicated earlier, the simple and multivariate analyses are repeated for different subsets of respondents. An evaluation of the relative "accuracy" of the collective backcasts of each group is based on the relative R^2 statistics of the estimated equations, with a higher R^2 indicating more accuracy. The following demographic and behavioral groups are examined:

Age[3]:	Less than 30, 30 to 54, 55 or more
Purchaser status[4]:	First-time homebuyer, repeat purchaser
Refinance status:	Purchaser, refinancer
Education[5]:	Four-year college degree or higher, less than four-year college degree
Household income:	Less than $50,000, $50,000 to $149,999, $150,000 or more
Number of borrowers:	Single borrower, borrower with spouse/partner
Property use:	Primary residence, second or investment home
Credit score[6]:	Below 720, 720 to 799, 800 or more
Financial riskiness:	Above average, average, no risk
Qualification concern:	Very, somewhat, not at all

Results for both the simple and multivariate analyses for the subgroups are presented in exhibit 5. Like the overall analysis, both the simple and multivariate subset analyses produce similar results. With a few exceptions, those subgroups which were more accurate with the simple analysis were also more accurate when control factors were added. Refinancers were collectively more accurate than purchasers, perhaps because they lived in the area during the measurement period. Within the purchaser group, first-time homebuyers appeared to be more accurate, perhaps because homeownership is a bigger decision for them and they put more energy into the shopping process as a result.

Other groups that were collectively more accurate than comparison groups are borrowers aged 30 to 54 (relative to borrowers younger than 30 or aged 55 or older), college graduates (relative to borrowers with less formal education), higher-income households (relative to lower-income households), owner-occupants (relative to investors/second-home borrowers), borrowers who identified that they were willing to take above average financial risk (relative to those with less appetite for risk), and borrowers who indicated that they were either somewhat or not concerned about mortgage qualification (relative to those who were very concerned). This analysis also showed that neither credit score nor the number of borrowers on a mortgage was consistent in ranking accuracy when comparing the simple and multivariate analyses.

[3] Age is defined as the older of the respondent and spouse age if the spouse is a borrower on the loan.
[4] The purchaser demographic excludes refinancers, while all other demographics capture all respondents.
[5] Household education level is defined as the highest level indicated between the respondent and spouse. The four-year college degree group includes respondents who indicated they were a "college graduate" or had completed "postgraduate studies."
[6] Credit score uses VantageScore® 3.0, which ranges from 300 to 850, with higher scores implying less risk. The score used is the higher of the respondent and spouse at origination if the spouse is a borrower on the loan.

Exhibit 5

Relative Accuracy of House Price Trend Backcast Within Demographic Subsets

Demographic Category	Subset	N	R^2 Simple	R^2 Multivariate	Mean HPI Change (%) Within Response Category		
					Significant Increase	Little/No Change	Significant Decrease
Age	<30	1,870	.0981	.4659	13.8	8.1	5.6
	30–54	13,499	.1027	.4945	15.4	9.5	7.9
	55+	9,406	.0696	.4801	15.3	9.7	9.7
Purchase Status	First–Time	3,290	.0830	.4978	14.5	8.9	8.4
	Repeat	8,231	.0711	.4569	14.9	9.8	8.6
Refinance Status	Purchase	11,521	.0756	.4689	14.8	9.5	8.6
	Refinance	13,254	.1029	.5031	15.5	9.2	8.5
Education	4–Year Degree	17,562	.1028	.4900	15.4	9.4	8.3
	<4–Year Degree	7,213	.0616	.4765	14.5	9.2	8.9
Household Income	$150K+	3,941	.0866	.5438	16.0	10.7	8.7
	$50K–$149,999	15,428	.0924	.4859	15.3	9.4	8.6
	<$50K	5,406	.0664	.4667	13.8	8.4	8.2
Number of Borrowers	2	18,910	.0938	.4835	15.2	9.3	8.4
	1	5,865	.0816	.4877	15.1	9.5	8.9
Property Use	Owner/Occupants	22,751	.0925	.4860	15.1	9.2	8.6
	Investment/Rental	2,024	.0825	.4719	16.0	10.5	7.9
Credit Score	800+	5,923	.0806	.5016	15.2	9.6	9.1
	720–799	11,057	.1006	.4982	15.6	9.4	8.6
	<720	7,795	.0833	.4706	14.7	9.1	8.1
Risk Appetite	Above Average	5,752	.1029	.5112	15.3	9.6	7.6
	Average	12,580	.0939	.4815	15.4	9.3	8.7
	No Risk	6,443	.0710	.4721	14.7	9.1	8.9
Qualifying Concern	Very	3,485	.0722	.4749	15.0	9.7	8.5
	Somewhat	7,101	.0961	.4888	15.5	9.3	8.5
	Not at All	14,189	.0938	.4883	15.1	9.3	8.5
Benchmark	**All**	**24,775**	**.0908**	**.4838**	**15.2**	**9.3**	**8.5**

HPI = House Price Index.

Borrower Forecasts of House Prices

Borrowers were asked, "What do you think will happen to the prices of homes in this neighborhood over the next couple of years?"—a question referred to as a "forecast" in this study. The distribution of their responses is presented in exhibit 6. The distribution of actual percent change in county HPI over the 2 years after the quarter in which the borrower completed the survey is presented in exhibit 7. Again, the figures demonstrate that both borrower forecasts and county-level HPIs skewed heavily toward price increases, though in both instances, there is some evidence of price decreases.

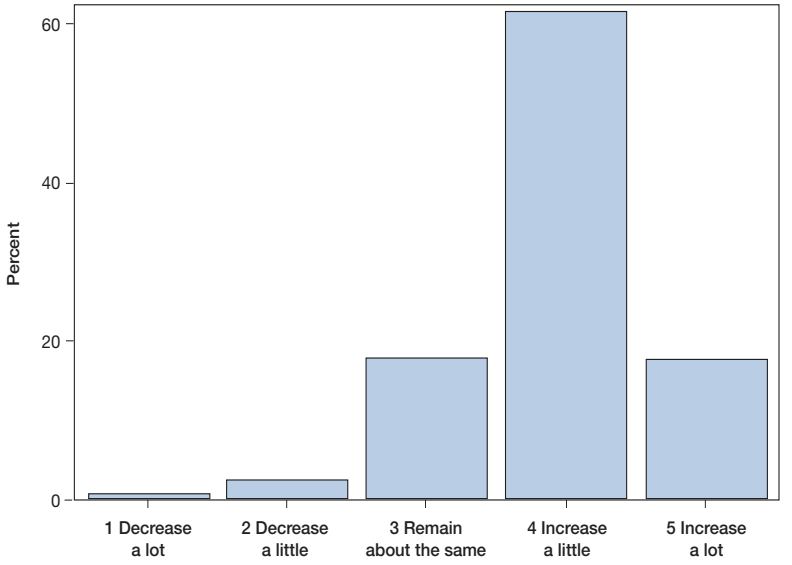
The alignment between borrower forecasting of local house price trends and actual house price trends is examined in the same two ways as studied for backcasting (simple and multivariate). A five-point scale was used, however, in soliciting respondents' forecasts versus a three-point scale for backcasts. Thus, forecast information is represented by four dummy variables for "Increase a lot," "Remain about the same," "Decrease a little, and "Decrease a lot," with "Increase a little" treated as the base group since it was the modal response.

Exhibit 8 presents estimates from the simple and multivariate forecast tests. The highly significant coefficients on the four dummy variables in the simple test have appropriate signs but are unbalanced in their effects. The effects of coefficients on "Increase a lot" and "Remain about the same" have similar magnitude in absolute value. The coefficients on "Decrease a little" and "Decrease a lot," while greater in magnitude than on "Remain about the same," do not follow a linear pattern. This may just be a residual of the fact that house prices generally rose over the sample period. It may also mean that borrowers are more sensitive to the amount of a decline in prices than they are to a comparable increase.

In the multivariate regression, all local economic fundamental control variables except for the index of average loan balance and lagged 2-year delinquency rate are significant, providing evidence that the other selected fundamentals are good indicators of actual house price growth. The failure of the mortgage market variables to show predictive power suggests that these variables lag house price changes rather than lead them. The coefficients on the response dummy variables "Increase a lot" and "Remain about the same" are still significant, however the response dummy variables "Decrease a little" and "Decrease a lot" are no longer significant. This could be an indication that borrowers possess information on future house prices above what could be estimated when prices are trending upward.

Exhibit 6

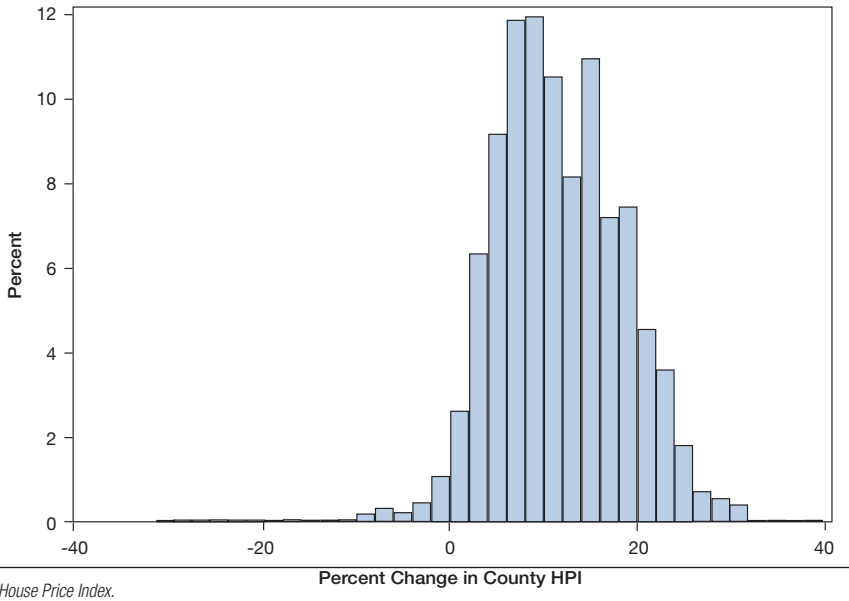
Expectation of Future Trends in Neighborhood House Prices



Note: Distribution of responses to the question, "What do you think will happen to the prices of homes in this neighborhood over the next couple of years?".
Source: NSMO

Exhibit 7

Distribution of Actual Future Trends in County House Prices



HPI = House Price Index.
Note: Distribution of actual change in county house prices in the 2 years after the quarter of survey completion.
Source: FHFA HPI

Exhibit 8

Parameter Estimates for Simple and Multivariate Forecast Regressions

	Forecast Alignment with Actual House Price Trends	
Parameter	**Simple**	**Multivariate**
N	21,847	21,847
R^2	0.0407	0.5141
Intercept	11.87***	6.57***
	(0.06)	(0.17)
"Increase a lot"	1.96***	0.29***
	(0.12)	(0.09)
"Remain about the same"	– 2.27***	– 0.25**
	(0.12)	(0.09)
"Decrease a little"	– 2.49***	– 0.29
	(0.29)	(0.21)
"Decrease a lot"	– 3.10***	– 0.12
	(0.55)	(0.39)
Past HPI Growth	-	0.21***
		(0.00)
Population Growth	-	1.16***
		(0.02)
Income Growth	-	0.36***
		(0.01)
Unemployment Level	-	– 0.07**
		(0.02)
Change in Unemployment	-	– 0.81***
		(0.08)
Average Mortgage Balance	-	– 0.01
		(0.01)
Delinquency Rate	-	0.14***
		(0.02)
Change in Delinquency Rate	-	– 0.02
		(0.05)
Survey Quarter Fixed Effects	**No**	**Yes**

HPI = House Price Index.
** = p<0.05; ** = p<0.01; *** = p<0.001.*
Note: Standard errors are given in parentheses.

Conversely, this could also be an indication that borrowers may not possess information of future prices over what could be estimated when there is a downturn in house price trends.

The simple and multivariate analyses of forecasting is repeated for the same subsets of respondents examined in the backcasting section. Results for both the simple and multivariate analyses for the subgroups are presented in exhibit 9. There is less consistency between the simple and multivariate analyses than in the backcasting section. In multiple cases, there were reversals among pairs or a complete reverse ordering among triples.

Groups that were collectively more accurate than comparison groups at forecasting are young borrowers (relative to borrowers 30 and older), first-time homebuyers (relative to repeat purchasers), single-borrower households (relative to two-borrower households), owner-occupants (relative to investors/second-home borrowers), and borrowers who indicated that they were very concerned about mortgage qualification (relative to those that were somewhat or not at all concerned).

Several groups reversed ordering between the simple and multivariate analyses. Based on the simple analyses, the forecasts of the following groups were better aligned with actual price trends: refinancers (relative to purchasers), less formally educated (relative to college graduates), lower-income households (relative to higher-income households), and borrowers who take no financial risks (relative to borrowers who accept financial risk). In the multivariate analyses, these categories had a complete reversal in ordering. In this case, purchasers, higher-educated borrowers, higher-income borrowers, and borrowers willing to take above average financial risks may provide some information on future house prices above what could be estimated from fundamentals. Again, the credit score comparison groups showed no consistency in patterns between the simple and multivariate analyses.

Finally, there was very little consistency in the demographics that were relatively better at both backcasting and forecasting. The only group that was consistent in simple and multivariate test results and relatively better at both backcasting and forecasting was the first-time homebuyer group (relative to repeat purchasers).

How Expectations of Future House Prices are Formed

In the previous section, borrower responses regarding their expectations for future prices changes in their neighborhood were compared with actual future changes to determine if the respondent data have forecasting value. This section examines how these expectations are formed. This question is important in improving the understanding of how consumers process house price information, particularly when the answers are found separately for different subsets.

Here, the responses to the question "What do you think will happen to the prices of homes in this neighborhood over the next couple of years?" are treated as the dependent variable. There is no reason to believe that the responses follow a cardinal ordering so they are treated as ordinal, and the process is modeled as an ordered logistic regression.[7] Three sets of explanatory variables are

[7] See Cameron and Trivedi (2005) for a detailed review of ordered logistic regression methodology.

Exhibit 9

Relative Accuracy of House Price Trend Backcast Within Demographic Subsets

Demographic Category	Subset	N	R^2 Simple	R^2 Multivariate	Mean HPI Change (%) Within Response Category				
					Increase "a lot"	Increase "a little"	Remain Same	Decrease "a little"	Decrease "a lot"
Age	<30	1,685	.0444	.5369	13.0	9.9	6.9	6.1	5.2
	30–54	11,898	.0405	.5280	14.9	11.9	8.0	7.6	7.0
	55+	8,264	.0407	.4860	15.1	11.9	7.5	8.4	6.0
Purchase Status	First–Time	2,943	.0481	.5532	13.5	10.8	7.7	6.2	8.1
	Repeat	7,321	.0296	.5036	14.3	11.6	8.0	8.3	8.2
Refinance Status	Purchase	10,264	.0366	.5192	14.0	11.4	7.9	7.3	8.2
	Refinance	11,583	.0443	.5116	15.4	11.9	7.6	7.7	6.0
Education	4–Year Degree	15,457	.0357	.5220	14.7	11.8	8.3	7.2	6.6
	<4-Year Degree	6,390	.0511	.4999	14.8	11.3	6.7	8.2	6.5
Household Income	$150K+	3,419	.0295	.5252	16.2	13.0	10.3	7.5	10.1
	$50K–$149,999	13,586	.0398	.5183	14.8	11.7	7.9	7.9	8.0
	<$50K	4,842	.0417	.4995	13.1	10.5	6.5	7.1	4.4
Number of Borrowers	2	16,649	.0395	.5084	14.7	11.6	7.8	7.8	7.8
	1	5,198	.0447	.5363	14.6	11.9	7.5	7.1	5.1
Property Use	Owner/Occupants	20,013	.0444	.5215	14.7	11.6	7.6	7.5	6.5
	Investment/Rental	1,834	.0150	.4729	15.0	12.3	9.3	9.0	7.2
Credit Score	800+	5,210	.0395	.5197	15.1	11.7	8.0	6.3	9.6
	720–799	9,737	.0354	.5170	15.1	11.9	8.3	7.7	5.7
	<720	6,900	.0484	.5133	14.2	11.3	7.0	7.9	6.7
Risk Appetite	Above Average	5,062	.0302	.5290	14.9	11.8	8.3	7.3	7.1
	Average	11,104	.0410	.5217	14.7	11.7	7.8	7.4	6.5
	No Risk	5,681	.0464	.4887	14.4	11.4	7.4	7.9	6.5
Qualifying Concern	Very	3,116	.0588	.5389	15.0	11.8	7.7	8.4	5.9
	Somewhat	6,363	.0348	.5061	14.6	11.8	7.9	7.7	7.1
	Not at All	12,368	.0396	.5151	14.7	11.6	7.6	7.1	6.8
Benchmark	**All**	**21,847**	**.0407**	**.5141**	**14.6**	**11.6**	**7.9**	**7.9**	**6.7**

HPI = House Price Index.

included in the analysis: (1) dummy variables representing the respondent's backcasts; (2) the set of county-level economic fundamentals variables; and (3) variables representing the demographic and other subgroups while controlling for backcasting and local economic fundamentals. Exhibit 10 presents results for the three different models. Column 1 uses only those variables in set (1) above; column 2 uses only the variables in (2); column 3 uses all three sets of variables. In each case, the columns present the log odds coefficients of the ordered logistic. Here, positive values should be interpreted as increasing the odds of a more optimistic response category and negative values as decreasing the same.

Exhibit 10

Three Models of House Price Expectation Formation (1 of 2)

Parameter	Models of House Price Expectation Formation Model 1 Past Perceptions *(N=24,775)*	Model 2 Local Fundamentals *(N=24,775)*	Model 3 Demographic Effects *(N=24,775)*
	All ordered logistic regression coefficients are reported in log odds.		
"Significant Increase"	1.35***	-	1.14***
	(0.03)		(0.04)
"Significant Decrease"	– 0.69***	-	– 0.62***
	(0.05)		(0.05)
Past HPI Growth	-	0.03***	0.02***
		(0.00)	(0.00)
Population Growth	-	0.11***	0.09***
		(0.01)	(0.01)
Income Growth	-	0.03***	0.02***
		(0.00)	(0.00)
Unemployment Level	-	– 0.02	0.00
		(0.01)	(0.01)
Change in Unemployment	-	– 0.01	0.01
		(0.04)	(0.04)
Average Mortgage Balance	-	0.02***	0.00
		(0.00)	(0.00)
Delinquency Rate	-	0.01	0.02*
		(0.01)	(0.01)
Change in Delinquency Rate	-	– 0.01	– 0.02
		(0.02)	(0.02)

Exhibit 10

Three Models of House Price Expectation Formation (2 of 2)

Parameter	Models of House Price Expectation Formation Model 1 Past Perceptions *(N=24,775)*	Model 2 Local Fundamentals *(N=24,775)*	Model 3 Demographic Effects *(N=24,775)*
Age <30	-	-	0.04 (0.06)
Age ≥55	-	-	0.19*** (0.03)
Spouse/Partner	-	-	0.01 (0.04)
Income ≤$50K	-	-	– 0.26*** (0.04)
Income ≥$150K	-	-	0.01 (0.04)
College Graduate	-	-	0.06 (0.04)
Credit Score <720	-	-	0.12*** (0.04)
Credit Score ≥800	-	-	– 0.05 (0.04)
First-Time Homebuyer	-	-	0.01 (0.05)
Refinance	-	-	– 0.25*** (0.03)
Owner–Occupants	-	-	0.03 (0.05)
Very Concerned	-	-	– 0.02 (0.05)
No Concerns	-	-	– 0.07* (0.03)
Above Average Financial Risk	-	-	0.23*** (0.04)
No Financial Risk	-	-	– 0.18*** (0.04)
Survey Quarter Fixed Effects	**No**	**Yes**	**Yes**

HPI = House Price Index.

** = p<0.05; ** = p<0.01; *** = p<0.001.*

Notes: Model 1 captures recent perceptions of borrowers through responses to the question: "In the last couple of years, how have house prices changed in the neighborhood where this property is located?" Model 2 includes local economic fundamentals as independent variables. Model 3 includes the perceptions from Model 1 and the fundamentals from Model 2 and adds demographic group dummy variables to capture differences relative to the to the base group which has the following characteristics: age 30 to 54, one-borrower household, income from $50,000 to $149,999, less than 4 years of college education, credit score from 720 to 799, repeat homebuyer, investor/second home owner, somewhat concerned about mortgage qualification, and takes average financial risks. All ordered logistic regression coefficients are reported in log odds. Exponentiating the log odds coefficients returns the proportional odds ratio. Standard errors are given in parentheses.

It is not surprising that when regressing backcasts of recent house price changes on forecasts of future changes (column 1), the coefficients on the dummy variables are highly significant and properly signed. Holding all else constant, if the respondent selected "Significant Increase" as a backcast on recent prices, the odds of that respondent selecting a more optimistic categorical response as their forecast are nearly four times higher than selecting a less optimistic response. Conversely, if the respondent selected "Significant Decrease" as a backcast, the odds of selecting a more optimistic categorical response as their forecast are less than half as high as the odds of selecting a lower categorical response.

The results presented in column 2 indicate that more positive economic fundamentals including actual recent house price changes, population growth, income growth, and average mortgage balances slightly increase the odds of a respondent selecting a more optimistic categorical response as a forecast. Somewhat unexpectedly, unemployment indicators did not seem to be related to consumer forecasts.

The results presented in column 3 put these pieces together and measure the marginal, rather than gross, impact of both backcasts and economic fundamentals variables as well as the impact of variables representing demographic and other group factors. Somewhat surprisingly, the coefficients on the backcasting and economic fundamental variables are relatively similar to their gross effects shown in columns 1 and 2. This suggests that they have separate impacts on respondent forecasts. Coefficients on the group variables show that older borrowers, borrowers with low credit scores, and borrowers that take above average financial risk are significantly more likely to select a more optimistic response (conditioned on their backcast and fundamentals). Other groups were less optimistic. Borrowers with low income, refinancers, and borrowers that take no financial risk are more likely to select a more pessimistic response.

Conclusion

Assessing awareness of house price trends is an important part of consumer protection in the mortgage market. This article extends the literature by identifying variation across recent mortgage borrowers—those with the greatest incentive to be informed—in their awareness of changes in the house prices in their local markets. On a positive note, this study finds that first-time homebuyers, a demographic considered especially vulnerable during the origination process, are consistently more aware of recent and impending house price trends than repeat borrowers. On the other hand, this study finds that lower-income and less-educated borrowers have relatively less awareness about price trends.

This article also provides evidence that expectations of future house price changes in consumers' local markets have informational value for forecasting beyond the value provided by local economic data. Why this is true, however, is not clear and needs to be further researched. This article finds that different demographic groups display varying degrees of optimism in forming their forecasts even when controlling for their backcasts and economic fundamentals. Borrowers willing to take above average risk stand out as the most optimistic forecasters, while lower income borrowers tend to be the least optimistic.

It is important to note that this study defined a relatively aggregated area—the county—as the "neighborhood" of the borrower. It is likely that there is neighborhood variation within counties that is not reflected in this definition, however, data at more local levels were unavailable for detailed exploration. Further, the results presented are in aggregate and may not reflect a particular locale.

Finally, this article should be treated as a "first stab" at the problem. NSMO contains a wealth of information on borrower behavior, expectations, and attitudes. While the article identifies the fact that consumers vary in the accuracy of their backcasts and forecasts, it does not address how this variation affects their behavior and, importantly, their outcomes in the mortgage market. These issues remain topics for future research.

Author

Chad Redmer is a Ph.D. student in the Economics Department at George Washington University and a Commander in the U.S. Navy assigned to teach in the Economics Department at the U.S. Naval Academy. Please contact the author at caredmer@gwmail.gwu.edu with any questions or comments.

Acknowledgments

The author thanks Robert B. Avery, Timothy Critchfield, Ian Keith, Will Larson, Saty Patrabansh, Jay Schultz, and Rebecca Sullivan for providing comments, Anthony Yezer for providing continuous guidance and encouragement throughout my GWU experience, and Darryl Getter for a lifetime of advising.

References

Cameron, A. Colin, and Pravin K. Trivedi. 2005. Microeconometrics: Methods and Applications. New York: Cambridge University Press.

Case, Karl E., Robert J. Shiller, and Anne Thompson. 2015. What Have They Been Thinking? Home Buyer Behavior in Hot and Cold Markets. NBER Working Paper No. 18400. Cambridge, MA: National Bureau of Economic Research.

Niu, Geng, and Arthur van Soest. 2014. House Price Expectations. IZA Discussion Paper No. 8536. Bonn, NRW, Germany: Institute for the Study of Labor (IZA).

First-Time Homebuyer Counseling and the Mortgage Selection Experience in the United States: Evidence from the National Survey of Mortgage Originations

Robert B. Argento
Lariece M. Brown
Freddie Mac

Sergei Koulayev
Consumer Financial Protection Bureau

Grace Li
Fannie Mae

Marina Myhre
U.S. Department of Housing and Urban Development

Forrest Pafenberg
Saty Patrabansh
Federal Housing Finance Agency

Disclaimer
The views expressed in this article are those of the authors and are not necessarily those of Freddie Mac, the Consumer Financial Protection Bureau, Fannie Mae, the U.S. Department of Housing and Urban Development, or the Federal Housing Finance Agency.

Abstract

The existing literature on homebuyer education and counseling (HEC) often focuses on the evaluation of specific programs, generally using mortgage loan performance as the metric of success. This article contributes to the literature in two ways. First, it provides evidence on the benefits of HEC to mortgage borrowers in aspects other than mortgage performance. Second, the article evaluates HEC in general, not just one specific program. It does so by drawing from a nationally representative sample of all first-time homebuyers in the United States who took out a mortgage between 2013 and 2016. The study data comes from the National Survey of Mortgage Originations (NSMO), a new survey co-sponsored by the Federal Housing Finance Agency (FHFA) and the Consumer Financial Protection Bureau (CFPB). We find that 17 percent of a nationally representative sample of first-time homebuyers reported receiving some form of HEC. In an analysis of early loan performance, we find that while borrowers reporting HEC had higher delinquency rates, these differences decrease and are not statistically significant when controlling for observable differences between those reporting HEC and the group without HEC. Using propensity score matching, we find that first-time homebuyers who reported receiving HEC also reported better mortgage knowledge, more confidence in their ability to explain the mortgage process, and higher level of satisfaction with the mortgage they received.

Introduction

Homebuyer education and counseling (HEC) is often viewed as a strategy for achieving sustainable homeownership, particularly among low- to moderate-income households, and, as such, an important aspect of providing access to sustainable mortgage credit. Advocates of HEC promote these programs to better prepare first-time homebuyers for successful homeownership by helping them make good home purchase and mortgage decisions and by improving their financial management skills (DeMarco et al., 2016). Lenders may require HEC for potential homebuyers with more marginal credit characteristics, such as higher loan-to-value (LTV) ratios, to mitigate risk. Therefore, HEC has become a standard part of many government, nonprofit, and industry programs geared toward low- to moderate-income first-time homebuyers on the premise that both homebuyers and society will benefit.

However, there has not been definitive evidence or a consensus on the benefits and cost-effectiveness of HEC. The existing literature on housing education and counseling often focuses on the evaluation of specific programs, generally using mortgage loan performance as the metric of success. The literature shows mixed results on HEC's effectiveness. In addition, the question of whether borrowers benefit in other ways from HEC, even when mortgage performance in the short term may not be improved, is largely unaddressed. This article contributes to the literature in two ways. First, it provides evidence about potential benefits of HEC to nonperformance aspects, such as mortgage knowledge and satisfaction with the mortgage terms and the mortgage process. Second, it uses a new nationally representative sample of first-time homebuyers rather than focusing on a specific program.

In this article, we examine the incidence and effectiveness of HEC among the general population of first-time homebuyers who took out a mortgage for home purchases between 2013 and 2016. We use responses from the NSMO, a new survey of a nationally representative sample of new mortgages in the United States. NSMO is a component of the National Mortgage Database® (NMDB).[1]

[1] NSMO and NMDB are described in the guest editors' introduction.

We find that 17 percent of first-time homebuyers reported receiving some form of HEC. In an analysis of early loan performance, we find that while borrowers reporting HEC had higher delinquency rates, these differences decrease and are not statistically significant when controlling for observable differences between the groups. Using propensity score matching to account for selection on observables, we find that first-time homebuyers who reported receiving HEC also reported better mortgage knowledge, more confidence in their ability to explain the mortgage process, and higher level of satisfaction with their mortgage.

Literature

The literature studying the effectiveness of prepurchase HEC is large and diverse; see Myhre and Watson (2017), Mayer and Temkin (2016), and Collins and O'Rourke (2011, 2010) for reviews. Nevertheless, the existing research is limited in several respects. First, nearly every study that we are aware of has focused on measuring counseling effectiveness among participants of a particular lending program.[2] An important exception is "The First-Time Homebuyer Education and Counseling Demonstration" study currently being conducted by the U.S. Department of Housing and Urban Development (HUD). This is a large-scale randomized experiment in which participants are recruited from a pool of mortgage applicants from three large national lenders (DeMarco et al., 2016).

Different lending programs have different counseling programs. There is a large variety of prepurchase counseling programs, which differ by (a) mode of delivery (individual counseling, classroom instruction, telephone instruction, home study); (b) program duration and content; (c) who delivers the instruction (nonprofit, government, lender); and (d) whether it is voluntary or mandatory. Typically, a study focuses on a specific counseling program (Agarwal et al., 2014a, 2014b, 2010; Brown, 2016) or by study design (Smith, Hochberg, and Greene, 2014; Carswell, 2009). Few studies span across counseling programs; see the article by Quercia and Spader (2008) for an analysis of the relative effectiveness of various types of counseling; also, see Avila, Nguyen, and Zorn (2013), whose study utilized a large sample of loans not restricted to a specific counseling program.

The selection specifics in the lending program and the specifics of the counseling program can make it difficult to generalize results of a given study. Therefore, we believe there is a need for evidence about the effectiveness of HEC that spans across lending programs and modes of instruction.

The existing studies of HEC have exclusively focused on the goal of home purchase or on the debt repayment behavior (mortgage delinquency and default, credit score, repayment of nonmortgage debt). Little evidence exists on the impact of HEC on borrower's mortgage knowledge and behaviors related to mortgage choice, such as mortgage shopping. The HUD experiment is the first attempt in this direction, and early results point to improvements in financial literacy and a "greater appreciation for communication with lenders" (DeMarco et al., 2016).

However, beyond the HEC context, there is a large body of literature studying the effectiveness of various financial education programs in improving financial literacy and related behaviors. For instance, Collins (2013) examined the impact of a mandatory financial education program

[2] See Avila, Nguyen, and Zorn (2013) for a study of Freddie Mac's Affordable Gold and Home Possible programs and Mayer and Temkin (2016) for NeighborWorks America's programs.

on very low-income households, finding improvements in self-reported behaviors. Broadly, low levels of financial literacy have been linked to suboptimal financial behaviors (see review by Lusardi and Mitchell, 2014). In the mortgage context, Alexandrov and Koulayev (2017) suggested that a lack of mortgage shopping may have prevented borrowers from realizing significant price savings. Moulton, Collins, Loibl, and Samak (2013) showed that many borrowers underestimate their total or monthly nonmortgage debt and are overconfident in their ability to pay down their debt, relative to the actual repayment behavior; and Gerardi, Goette, and Meier (2010) linked low financial literacy to delinquencies on subprime mortgage loans. Thus, there is a need to study the effectiveness of HEC as a specific type of financial education in addition to its role in default risk management.

This article addresses both aspects of HEC's effectiveness. Our results apply to the general population of first-time homebuyers who took out a mortgage between 2013 and 2016 and are not specific to a particular lending or counseling program. We explore a variety of outcomes beyond debt repayment, such as self-reported mortgage knowledge, shopping, mortgage selection, satisfaction with the mortgage process, and the use of government-mandated mortgage disclosures.

Finally, we point out that results measuring the effect of HEC on loan performance vary greatly across studies. Part of the difference in results across studies may be related to the time period when the loans were originated. In a natural field experiment of a Tennessee prepurchase homebuyer education program in 2002 funded by a HUD housing counseling grant, Brown (2016) found that borrowers who received HEC had a 42-percent reduced chance of foreclosure compared with the control group but no statistically significant difference in default (defined as first incidence of becoming 90 days delinquent). On the other hand, in a randomized field experiment sponsored by the Federal Reserve Bank of Philadelphia, Smith and colleagues (2014) found that one-on-one prepurchase counseling conducted by a HUD-approved housing counseling agency in 2007 had positive long-term effects on credit score and debt levels of participants compared with a control group who only received a 2-hour prepurchase workshop—but no impact on timeliness of mortgage payments, as most borrowers stayed current on their mortgage. Avila, Nguyen, and Zorn (2013) found that the effect of HEC on loan performance was largest among loans originated between 2006 and 2008. Similarly, Mayer and Temkin (2016) found a 30-percent reduction in 90-plus days' delinquency for 2007-vintage loans originated by NeighborWorks. Li and colleagues (2016) analyzed the same program but for post-crisis originations in the period from 2010 to 2012 and found a 14-percent reduction.

Analysis Sample and Outcomes

For this article, we use data from the NSMO survey conducted for a sample of borrowers with mortgages originated in 2013 through 2016. There were 24,847 respondents to this survey for these origination years. We focus specifically on first-time homebuyers to identify those with less experience in the mortgage process relative to repeat borrowers. Given the detailed administrative and credit data available in NMDB, we can identify homebuyers as those borrowers who have not had an active mortgage. We also restrict our analysis to homebuyers younger than 55 years to increase the likelihood that the analysis is focused on first-time buyers instead of those who paid off an earlier mortgage.[4]

[4] For NSMO, we match the survey respondent and the spouse (when one exists) to the borrowers in the credit file. At least one of them must be the borrower on the mortgage. If both are borrowers, we require that both have no prior mortgage in the credit file for the loan type to be designated as a first-time homebuyer loan. If only one is the borrower, we require that that person has no prior mortgage in the credit file for the loan type to be designated as first-time homebuyer loan.

NSMO identifies the recipients of HEC by asking, "Did you take a course about homebuying or talk to a housing counselor?" Followup questions ask what delivery mode was used for the homebuying course or housing counseling (in-person, one-on-one; in-person, group; over the phone; or online), how long the course or counseling lasted, and whether it was helpful. Exhibit 1 summarizes responses to these questions. Among the 3,305 first-time homebuyers that we have identified, 562 reported receiving some form of HEC, which amounts to a rate of 17 percent.[5] This is itself a novel finding; there is currently no reliable estimate of the percentage of recent mortgage borrowers who receive some form of HEC. Among those first-time homebuyers who report receiving any kind of HEC, 43 percent report receiving it in a group setting, while 18 percent report a one-on-one counseling interaction, as shown in exhibit 1.[6]

Exhibit 1

Types of Counseling

	Weighted Frequency (%)	Survey Counts
Did you take a course about homebuying or talk to a housing counselor?		
No	83.18	2,743
Yes	16.82	562
How was the homebuying course or counseling provided?		
In Person, One-on-One	18.49	102
In Person, in a Group	43.02	244
Over the Phone	12.39	68
Online	49.89	281
How many hours was the homebuying course or counseling?		
Less Than 3 Hours	43.07	247
3–6 Hours	28.91	166
7–12 Hours	22.75	123
More Than 12 Hours	5.26	26
Overall, how helpful was the homebuying course or counseling?		
Very	55.20	302
Somewhat	37.66	218
Not at All	7.15	42

Notes: Responses do not sum up to 100 percent because categories are not mutually exclusive. Frequencies are adjusted for population weights. Survey counts are unweighted actual responses.
Source: NSMO, 2013–2016

The scope of this analysis is a comparison of the 562 first-time homebuyers reporting HEC with the remaining 2,743 first-time homebuyers who did not report receiving HEC. Substantively, the NSMO does not distinguish between homebuyer education (group classroom instruction about the homebuying process) and homebuyer counseling (individual, one-on-one sessions with a housing counselor that are tailored to a client's financial situation and stage in the homebuying process). For instance, 18 percent of respondents indicated a one-on-one session, which probably corresponds to counseling rather than homebuyer education. It is unclear whether the remaining 82 percent received education, counseling, or both. For instance, those who report group instruction may have also received individual consultation. NSMO does not identify whether

[5] Our definition of a first-time mortgage borrower is the following: (a) no mortgage in the previous credit history; (b) both the borrower and the coborrower, if present, are 55 years of age or younger. This approach results in approximately 36 percent of first-time mortgage borrowers among all borrowers for home purchase in 2013–2016.
[6] Sample weights are used for frequencies, summary means, and the linear probability models. For each survey response, the sample weight is adjusted for non-response by multiplying the sampling weight and the non-response adjustment.

the provider of homebuyer education or counseling meets HUD or National Industry Standards (NIS). In fact, the reported counseling could involve a for-profit organization or an online tool that does not meet HUD or NIS standards. For example, the provider could be the lender or mortgage insurance company. Therefore, we chose not to compare groups who self-reported different types of HEC in NSMO to one another.

The NSMO survey focuses on borrower experiences when obtaining a mortgage. We identify a broad set of questions related to mortgage knowledge and mortgage-related behaviors. For each question, we compare the responses of first-time homebuyers who reported receiving HEC with those who did not report receiving HEC while controlling for the set of relevant borrower covariates. We believe the observed differences in responses are informative of how HEC may have affected the underlying mortgage knowledge and mortgage-related behaviors. For this reason, we label these responses as "outcomes."

Each outcome is an indicator variable equal to one if the respondent provides a certain answer to the relevant question, and zero otherwise. The question, "How well could you explain to someone the process of taking out a mortgage?" generates an indicator variable equal to one if the respondent replied "very well" and zero otherwise. The mean of the outcome variable represents the share of respondents who provided a particular answer; for the question above, the weighted sample mean is 41.32 percent, indicating that 41.32 percent of respondents said "very well," as shown in exhibit 2. Among HEC recipients, the share of respondents who said "very well," 48.42 percent is much higher than the 39.89 percent of non-HEC respondents who said they could explain the process of taking out a mortgage "very well". For the complete list of questions and corresponding outcomes, see exhibit 2.[7] Next, we provide a broad overview of the categories of survey questions we analyze and how first-time homebuyers responded to them.

Exhibit 2

Summary Statistics of Responses for Select Questions (1 of 3)

	Weighted Means (%)		
	Sample	**No HEC**	**HEC**
When you began the process of getting your mortgage, how familiar were you (and any cosigners) with each of the following? (=1 if very familiar)			
The mortgage interest rates available at that time	33.58	32.81	37.39
The different types of mortgages available	23.66	22.84	27.75
The process of taking out a mortgage	15.59	14.68	20.10
The downpayment needed to qualify for a mortgage	35.99	36.26	34.65
The income needed to qualify for a mortgage	32.20	31.45	35.91
Your credit history or credit score	64.43	63.60	68.57
The money needed at closing	25.20	25.03	26.06
How well could you explain to someone the… (=1 if very well)			
Process of taking out a mortgage	41.32	39.89	48.42
Difference between a fixed- and an adjustable-rate mortgage	55.59	55.01	58.49

[7] The current survey instrument is available at https://www.fhfa.gov/nsmo.

Exhibit 2

Summary Statistics of Responses for Select Questions (2 of 3)

	Weighted Means (%)		
	Sample	**No HEC**	**HEC**
How well could you explain to someone the… (=1 if very well)			
Difference between a prime and a subprime loan	12.44	12.50	12.13
Difference between a mortgage's interest rate and its APR	19.17	19.05	19.76
Amortization of a loan	23.98	24.41	21.86
Consequences of not making required mortgage payments	55.73	54.30	62.81
How many different lenders/brokers did you seriously consider before choosing where to apply for this mortgage? (=1 if more than one lender/broker)			
Seriously considered more than one lender/broker	54.64	54.26	56.52
How many different lenders/brokers did you end up applying to? (=1 if more than one lender/broker)			
Applied to more than one lender/broker	31.71	30.97	35.37
Did you seek input about your closing documents from any of the following people? (=1 if yes, =0 otherwise)			
Lender/mortgage broker	72.92	72.91	73.00
Settlement/closing agent	22.70	22.51	23.65
Real estate agent	60.78	61.13	59.04
Personal attorney	15.83	15.83	15.87
Title insurance agent	17.41	16.69	21.00
Trusted friend or relative who is not a cosigner on the mortgage	44.44	45.01	41.60
Housing counselor	2.45	0.77	10.73
Any of the sources	88.17	87.97	89.17
How important were each of the following in choosing the lender/broker you used for the mortgage you took out? (=1 if important or very important)			
Having an established banking relationship	43.98	44.64	40.71
Having a local office or branch nearby	52.04	51.68	53.82
Used previously to get a mortgage	7.69	7.19	10.15
Lender/broker is a personal friend or relative	12.64	12.27	14.42
Recommendation from a friend/relative/co-worker	50.07	50.51	47.90
Recommendation from a real estate agent/home builder	51.44	51.07	53.28
Reputation of the lender/broker	59.12	58.73	61.04
Spoke my primary language, which is not English	9.64	8.89	13.31
Any of non-price factors are important	91.79	91.79	91.84

Exhibit 2

Summary Statistics of Responses for Select Questions (3 of 3)

	Weighted Means (%)		
	Sample	**No HEC**	**HEC**
Overall, how satisfied are you that the mortgage you got was the one with the… (=1 if very satisfied)			
Best terms to fit your needs	73.47	73.36	73.59
Lowest interest rate for which you could qualify	68.47	68.61	57.28
Lowest closing costs	56.53	62.05	62.40
Any option = very	84.26	87.47	63.62
Overall, how satisfied are you with the… (=1 if very satisfied)			
Lender or mortgage broker you used	72.94	72.80	73.59
Application process	59.57	60.03	57.28
Loan closing process	61.71	61.58	62.40
Information in the mortgage disclosure documents	61.24	60.75	63.62
Timeliness of mortgage disclosure documents	60.35	60.91	57.61
Settlement agent	65.90	66.07	65.06
Overall satisfied with mortgage process	87.19	86.90	88.64
Did you face any unpleasant surprises at your loan closing? (=1 for "yes")			
Did you face any unpleasant surprises at your loan closing?	16.63	15.50	22.25
Administrative data			
Mortgage interest rate spread at origination greater than 100 BPS	12.93	12.22	16.43
LTV ratio at origination greater than 95 percent	44.24	42.24	54.11
Mortgage debt to income ratio greater than 45 percent	12.57	12.29	13.94
Second lien	2.14	1.43	5.62
Ever 60 days delinquent	3.65	3.12	6.31
Ever 90 days delinquent	2.51	2.19	4.11

BPS = basis points. HEC = homebuyer education and counseling. LTV = loan-to-value.
Note: Sample size is 3,305.
Source: NSMO, 2013–2016.

Knowledge of Mortgage Process

Respondents were asked how familiar they were with their credit score, the current level of interest rates, available mortgage types, the mortgage process in general, and the downpayment needed when applying for mortgages when they began the process of getting this mortgage. For each item, we compare the response of "very familiar" to the responses "somewhat" or "not at all." While 64 percent of first-time homebuyers responded they were very familiar with their credit history and credit score, only 16 percent reported high familiarity with the process of taking out a mortgage.

Another survey question asked how well the respondent could explain to others the process of taking out a mortgage, the difference between a fixed- and adjustable-rate mortgage, the difference between the mortgage interest rate and the annual percentage rate (APR), and the consequences of not making required mortgage payments. We compare the self-reported "very well" to "somewhat" or "not at all" responses. Most first-time homebuyers responded that they understood the difference between a fixed- and an adjustable-rate mortgage (56 percent) and the consequences of not making required payments (56 percent). Respondents were least comfortable explaining the differences between prime and subprime loans (12 percent) and the difference between the interest rate and the APR (19 percent).

Assigned Importance to Non-Price Factors

Respondents were asked if lender attributes—such as lender reputation, the lender's online presence, whether the lender has a branch nearby, and whether the lender can speak the borrower's language—were important or very important when choosing the lender. These attributes are normative, non-price criteria. We create a combined indicator equal to one if at least one attribute was deemed very important. We created variables with a value of one if the respondent answered "very important" for any of the non-price factors. A very high percentage, 92 percent, of first-time homebuyers indicated that at least one non-price attribute was "very important." This suggests that their mortgage choice might not have been guided by mortgage cost alone.

Satisfaction with Mortgage Process and Mortgage Terms

The survey contains a variety of questions that focus on a borrower's satisfaction with the mortgage experience, beginning with the application process and ending with mortgage terms. Generally, borrowers are satisfied with both the process and mortgage terms. For example, close to 68 percent of respondents report being satisfied with the interest rate on the mortgage.

Mortgage Shopping

Shopping behavior is measured two ways. First, the respondent is asked, "How many lenders/brokers did you seriously consider before taking out this mortgage?" and by this measure, 55 percent of first-time homebuyers reported that they seriously considered more than one lender or broker. Borrowers were also asked, "How many different lenders/brokers did you end up applying to?". By this second measure, 32 percent indicated that they applied to more than one lender or broker.

Seeking Input About Closing Documents

Eighty-eight percent of respondents reported they sought input about closing documents from at least one source, with the lender and real estate agent being the most popular sources (73 percent and 61 percent, respectively). Forty-four percent asked for input from a friend or relative.

Mortgage Terms

Using survey responses and the mortgage administrative data linked to the survey, we can observe the characteristics associated with the mortgages selected by the borrower. Available information includes underwriting information such as LTV at origination and debt-to-income ratio (DTI), as well as product features such as interest rate, term, and fixed-rate versus adjustable-rate. Between 2013 and 2016, the mortgage market was relatively homogeneous, with only a small fraction of loans having "special" mortgage features, such as a balloon payment, prepayment penalty, or interest-only payments.[8] For this reason, we do not include these mortgage features as part of our analysis.

For our analysis, we created indicator variables to represent a mortgage with an LTV at origination greater than 95 percent, a mortgage interest rate spread at origination greater than 100 basis points,[9] a DTI ratio greater than 45 percent, and a second lien present at origination. For the first-time homebuyers in this study, 44 percent have an origination LTV greater than 95 percent, 2 percent are securing second liens with the first mortgage, and 13 percent have the higher DTI and interest-rate spread at origination.[10]

Some mortgage terms, including LTV and DTI, may reflect decisions or circumstances of a household that may predate HEC, such that any association between these terms and reported HEC should be taken with caution. In the more detailed empirical analysis, discussed below, we consider the possibility that these mortgage terms affect the selection into reported HEC as an additional specification. This is particularly true if borrowers participated in HEC to qualify for a specific mortgage program. Again, we are not able to confirm the exact nature, timing, or entity associated with the counseling.

Empirical Strategy

For each outcome discussed above, we examine the empirical relationship between the incidence of that outcome among survey respondents and the self-reported HEC, controlling for relevant borrower characteristics. Because the borrowers in our analysis were not randomly assigned to receive HEC, we employ propensity score matching (PSM) to control for the observable borrower characteristics related to which borrowers reported receiving homebuyer education or counseling. PSM controls for observable borrower and loan characteristics.[11] This is an improvement relative to other papers that were limited to detailed information for only borrowers who received counseling. Such papers relied on loan characteristics (Agarwal et al., 2014a, 2014b, 2010) or credit attributes (Mayer and Temkin, 2016; Roll and Moulton, 2016) to address selection.[12] Regardless, PSM is not able to control for unobserved attributes that might affect both counseling choices and mortgage selection, such as job security, mobility, available assets, or funds for a downpayment. Further, borrowers may select into HEC based on pre-HEC levels of mortgage knowledge in a way that is not fully captured by the included covariates. Overall, the results of our analysis are best

[8] Note that only first-lien residential mortgages are in the NSMO sample, and second liens such as Home Equity Lines of Credit (HELOC) are not captured.
[9] A basis point is one-hundredth of a percent.
[10] We have 45 fewer observations associated with DTI because we omit DTI observations that exceed 60 percent. The relevant counts are reported in exhibit 2.
[11] For an explanation of the method, see, for example, Dehejia and Wahba (2002).
[12] See Mayer and Temkin (2016) for a review.

interpreted as correlations or suggestive evidence of the relationship between HEC and aspects of the mortgage selection process, not proof of causal relationships.

We begin by summarizing how borrowers who reported HEC differ from those who reported that they did not participate in HEC. We have identified several borrower characteristics that are predetermined and potentially related to self-reported education or counseling participation. Exhibit 3 presents the breakdown of first-time homebuyers by borrower characteristics. The first column presents the frequency of a borrower type, adjusted for population weight to be as representative as possible to the general population of first-time homebuyers. The second column presents the counseling rate among borrowers of a given type which can be compared with the overall counseling rate of 17 percent. From this exhibit, we make the following observations.

- Credit Score: Borrowers with a lower credit score are generally more likely to report HEC participation, with the highest participation rates reported for borrowers with a credit score of 620–639.[13] Among these borrowers, the reported HEC rate is 27 percent.
- Age: Younger borrowers (35 years or younger) are less likely to report HEC participation than older borrowers. The younger borrowers' reported HEC participation rate is 15 percent compared with 21 percent for those 36 to 45 years old.
- Race and Ethnicity: The HEC rate among Blacks is 41 percent, more than twice the average in the sample.
- Education: Borrowers with a college degree have a lower reported HEC rate, at 16 percent, compared with a 24-percent rate among high school graduates.
- Household Income: Household income shows a strong relationship with HEC rate. Households with less than $50,000 in combined yearly income exhibit HEC rates of approximately 26 percent, well above that of higher income groups.
- Marital Status: At 15 percent, married couples have the lowest reported HEC rate, compared with 19 percent for both not-married couples and singles.

Exhibit 3

Types of Borrowers and the Counseling Rate for Each Type (1 of 3)

Covariate	Weighted Percent	
	Frequency	Counseling Rate
Loan Amount		
Less Than $50,000	1.88	21.91
$50,000 to $99,999	16.51	20.38
$100,000 to $149,999	25.65	19.25
$150,000 to $199,999	19.91	15.93
$200,000 to $249,999	12.95	15.72

[13] The credit score in the data is VantageScore® 3.0.

Exhibit 3

Types of Borrowers and the Counseling Rate for Each Type (2 of 3)

Covariate	Weighted Percent	
	Frequency	Counseling Rate
$250,000 to $299,999	8.18	16.91
$300,000 to $349,999	4.45	14.05
$350,000 to $399,999	3.08	10.21
$400,000 or More	7.41	7.84
Credit Score		
Lower Than 620	6.51	19.61
620 to 639	5.36	26.83
640 to 659	8.71	20.98
660 to 679	8.03	21.28
680 to 699	9.18	19.94
700 to 719	10.62	16.17
720 to 739	12.55	14.29
740 or Higher	39.04	13.40
Age at Last Birthday		
35 or Younger	73.81	15.47
36 to 45	18.60	21.32
Older Than 45	7.59	18.93
Race		
White/Caucasian	81.68	15.10
Black/African-American	6.31	40.65
Asian	8.22	13.19
Other	3.79	22.20
Hispanic or Latino		
No	13.08	21.60
Yes	86.92	16.10
Highest Level of Education Achieved		
Some Schooling	1.50	33.27
High School Graduate	8.95	23.99
Technical School	5.30	19.76
Some College	17.06	19.96
College Graduate	41.05	15.75
Postgraduate Studies	26.14	12.46
Income Relied Upon in Underwriting		
Less Than $35,000	8.39	25.72
$35,000 to $49,999	17.61	26.28
$50,000 to $74,999	26.62	19.02
$75,000 to $99,999	20.26	13.01
$100,000 to $174,999	19.74	8.35
$175,000 or More	7.37	9.32

Exhibit 3

Types of Borrowers and the Counseling Rate for Each Type (3 of 3)

Covariate	Weighted Percent	
	Frequency	Counseling Rate
Household Type		
Married	52.64	15.10
Not Married but With Partner	16.95	18.77
Single	30.41	18.73
Gender		
Female	44.47	18.63
Male	55.53	15.37
Employment Type		
Employed Full Time	87.78	16.22
Self-Employed	4.90	16.67
Other	7.33	24.20
LTV Ratio at Origination		
95% or Below	55.76	13.84
Greater Than 95%	44.24	20.58
Mortgage DTI Ratio		
45% or Below	87.43	16.51
Greater Than 45%	12.57	18.61
Has Second Lien		
No	97.86	16.22
Yes	2.14	44.22

DTI = debt-to-income. LTV = loan-to-value.
Source: NSMO, 2013–2014

To implement PSM we first estimate a logistic regression model of the likelihood that a borrower reports participating in HEC. The specification of this regression mirrors the specification used in the linear probability model. Based on the estimates of the logistic regression, the propensity score is computed for each observation in the sample. Borrowers reporting HEC are then matched to non-HEC borrowers with similar values for the probability of reporting HEC.[14] Within each matched group, the survey responses of those reporting HEC are compared with the responses of borrowers who did not report HEC using a linear regression model.

[14] We used the k-nearest neighbors method to identify matched observations. The results reported in the table include the 20 nearest neighbors. Additional sensitivity testing was conducted for k=10 and k=15.

In exhibit 4, we present results of a linear probability model of HEC with the variables used in the logistic regression of PSM. Credit score, race, age, education, and income are statistically significant predictors of reported HEC. When other factors, such as income, are controlled for, loan amount, household type, and employment type are not statistically significant in the counseling regression model.

As mentioned earlier, we consider an analysis in which higher LTV, higher DTI, and the presence of a second lien are considered outcome variables and an alternate analysis in which these variables are selection variables used in the logistic model of reporting HEC. The linear probability model of the alternate model is included on the right side of exhibit 4. Having a second lien is a statistically significant predictor of counseling.

Exhibit 4

Weighted Linear Probability Model of HEC Choice (1 of 2)

Control	Without Administrative Data			With Administrative Data		
	Estimate (%)	Standard Error (%)	t-Statistic	Estimate (%)	Standard Error (%)	t-Statistic
Intercept	26.90	7.69	3.50	24.62	7.69	3.20
Loan Amount						
Less Than $50,000 *(Omitted)*	–	–	–	–	–	–
$50,000 to $99,999	0.62	4.89	0.13	0.60	4.91	0.12
$100,000 to $149,999	3.17	4.86	0.65	3.35	4.90	0.68
$150,000 to $199,999	2.84	4.97	0.57	2.86	5.02	0.57
$200,000 to $249,999	4.56	5.12	0.89	5.13	5.17	0.99
$250,000 to $299,999	8.50	5.35	1.59	9.48	5.40	1.75
$300,000 to $349,999	7.05	5.79	1.22	8.14	5.82	1.40
$350,000 to $399,999	5.17	6.18	0.84	5.70	6.20	0.92
$400,000 or More	3.04	5.75	0.53	3.39	5.80	0.58
Credit Score						
Lower Than 620 *(Omitted)*	–	–	–	–	–	–
620 to 639	9.12**	3.71	2.46	9.36**	3.68	2.54
640 to 659	4.01	3.30	1.21	4.03	3.29	1.22
660 to 679	5.22	3.37	1.55	4.67	3.36	1.39
680 to 699	4.55	3.27	1.39	4.37	3.27	1.34
700 to 719	1.02	3.20	0.32	1.20	3.18	0.38
720 to 739	1.55	3.14	0.49	1.49	3.14	0.47
740 or Higher	1.97	2.81	0.70	2.39	2.85	0.84
Age at Last Birthday						
35 or Younger *(Omitted)*	–	–	–	–	–	–
36 to 45	4.16**	1.71	2.43	4.31**	1.71	2.53
Over 45	1.05	2.47	0.43	1.26	2.45	0.51

Exhibit 4

Weighted Linear Probability Model of HEC Choice (2 of 2)

Control	Without Administrative Data			With Administrative Data		
	Estimate (%)	Standard Error (%)	t-Statistic	Estimate (%)	Standard Error (%)	t-Statistic
Race						
White/Caucasian *(Omitted)*	–	–	–	–	–	–
Black or African-American	23.64***	2.69	8.79	23.40***	2.69	8.71
Asian	1.85	2.47	0.75	2.41	2.47	0.97
Other	6.79**	3.34	2.03	6.73**	3.33	2.02
Hispanic or Latino						
Yes	–	–	–	–	–	–
No	2.68	1.95	1.37	2.71	1.96	1.38
Highest Level of Education Achieved						
Some Schooling	–	–	–	–	–	–
High School Graduate	– 6.70	5.66	– 1.18	– 7.78	5.62	– 1.38
Technical School	– 8.60	5.95	– 1.44	– 8.42	5.91	– 1.42
Some College	– 10.19*	5.50	– 1.85	– 10.24*	5.47	– 1.87
College Graduate	– 10.54*	5.45	– 1.93	– 10.35*	5.41	– 1.91
Postgraduate Studies	– 11.60**	5.56	– 2.09	– 11.65**	5.52	– 2.11
Income Relied Upon in Underwriting						
Under $35,000 *(Omitted)*	–	–	–	–	–	–
$35,000 to $49,999	0.80	2.71	0.30	1.83	2.74	0.67
$50,000 to $74,999	– 6.27**	2.70	– 2.32	– 5.34*	2.73	– 1.95
$75,000 to $99,999	– 12.20***	2.97	– 4.11	– 11.07***	3.00	– 3.69
$100,000 to $174,999	– 17.25***	3.19	– 5.41	– 16.63***	3.23	– 5.14
$175,000 or More	– 14.71***	4.17	– 3.53	– 15.65***	4.23	– 3.70
Household Type						
Married *(Omitted)*	–	–	–	–	–	–
Not Married but With Partner	2.96*	1.80	1.65	2.93*	1.79	1.64
Not Married, No Partner	– 1.26	1.57	– 0.80	– 1.06	1.57	– 0.68
Gender						
Female *(Omitted)*	–	–	–	–	–	–
Male	– 2.12	1.32	– 1.60	– 2.13	1.32	– 1.61
Employment Type						
Employed Full Time *(Omitted)*	–	–	–	–	–	–
Self-Employed	0.95	2.98	0.32	2.04	2.97	0.69
Other	3.54	2.52	1.40	3.64	2.54	1.43
LTV Ratio at Origination						
95% or Below *(Omitted)*	–	–	–	–	–	–
Greater than 95%	–	–	–	1.55	1.50	1.04
Mortgage DTI Ratio						
45% or Below *(Omitted)*	–	–	–	–	–	–
Greater Than 45%	–	–	–	– 2.67	2.00	– 1.34
Has Second Lien						
No *(Omitted)*	–	–	–	–	–	–
Yes	–	–	–	29.04***	4.38	6.63

– = not applicable. DTI = debt-to-income. LTV = loan-to-value.
** p<0.05, ** p<0.01, *** p<0.001*
Source: NSMO, 2013–2016.

Empirical Results

The NSMO survey responses are used to explore the relationship between reported participation in HEC and aspects of the mortgage shopping and selection process. The categories we investigate include knowledge of the mortgage process, importance associated with non-price lender attributes, satisfaction with the mortgage process, number of lenders considered, seeking input on closing documents, and the post-origination loan performance.

For each response or mortgage attribute, we estimate OLS and PSM models.[15] Exhibits 5 and 6 present the results from all regressions. Exhibit 5 includes the results when higher LTV, higher DTI, and the presence of a second lien are considered outcome variables, and exhibit 6 includes an alternate analysis where these variables are considered as potential selection variables for reported HEC. We report the estimate of the HEC coefficient, multiplied by 100, with standard errors. Because all outcomes are indicator variables, the interpretation of the coefficient is the percentage point change in the response of interest. To help assess the economic magnitude of the change, the first column of the exhibit presents the sample average of the response or attribute variable.

In every case in which a significant result is found, all models report similar magnitude. Using the numbers reported for the PSM model in exhibit 6, highlights of the results include the following:

1. Borrowers who reported receiving HEC also reported large positive and statistically significant differences in financial knowledge regarding available mortgage interest rates, the different types of mortgages available, and their credit history or credit score. Holding other characteristics constant, borrowers reporting HEC were 9-percentage points more likely to report that they could explain the process of taking out a mortgage (with the population average of 41.32 percent) and 7-percentage points more likely to report they could explain the consequences of not making required mortgage payments (with the population average of 55.73 percent).

2. Borrowers who reported receiving HEC were 9-percentage points more likely to consult a housing counselor about their closing documents than were average first-time homebuyers (with a population average of 2.45 percent). This is a non-trivial result because a counseling course does not necessarily include an individual consultation, as it may have been in a group setting or online.

3. Borrowers were generally satisfied with at least one aspect of the mortgage they received, at 84.26 percent. Yet, borrowers who reported receiving HEC were 4-percentage points more likely to report being satisfied with their mortgage.

4. A somewhat surprising result is that borrowers who reported HEC were 5-percentage points more likely to report facing "unpleasant surprises" at the loan closing. It is not clear if the first-time homebuyers who reported HEC had more challenging closing processes or were more inclined to carefully review their loan estimates and closing documents.

[14] The PSM regression includes weights based on the number of observations in each group of matched respondents.

Exhibit 5

Estimates of the Relationship Between Reported HEC and the Outcome of Interest (1 of 2)

	Mean (%)	OLS			PSM		
		Coefficient (%)		Standard Error (%)	Coefficient (%)		Standard Error (%)
When you began the process of getting your mortgage, how familiar were you (and any cosigners) with each of the following? (=1 if very familiar)							
The mortgage interest rates available at that time	33.58	7.46	***	2.48	4.68	**	2.37
The different types of mortgages available	23.66	5.36	**	2.31	4.77	**	2.18
The process of taking out a mortgage	15.59	3.76	*	2.02	3.19		1.98
The downpayment needed to qualify for a mortgage	35.99	– 0.21		2.46	– 0.35		2.40
The income needed to qualify for a mortgage	32.20	4.56	*	2.48	3.55		2.40
Your credit history or credit score	64.43	6.95	***	2.41	5.77	**	2.36
The money needed at closing	25.20	– 0.83		2.27	– 1.87		2.23
How well could you explain to someone the… (=1 if very well)							
Process of taking out a mortgage	41.32	9.93	***	2.58	9.03	***	2.49
Difference between a fixed and an adjustable-rate mortgage	55.59	7.71	***	2.48	5.79	**	2.49
Difference between a prime and subprime loan	12.44	1.94		1.63	2.06		1.57
Difference between a mortgage's interest rate and its APR	19.17	2.84		2.01	2.84		2.02
Amortization of a loan	23.98	2.65		2.06	3.60	*	2.05
Consequences of not making required mortgage payments	55.73	9.37	***	2.49	6.20	**	2.45
How many different lenders/brokers did you seriously consider before choosing where to apply for this mortgage? (=1 if more than one lender/broker)							
Seriously considered more than one lender/broker	54.64	4.94	*	2.58	3.00		2.49
How many different lenders/brokers did you end up applying to? (=1 if more than one lender/broker)							
Applied to more than one lender/broker	31.71	3.77		2.47	3.99	*	2.41
Did you seek input about your closing documents from any of the following people? (=1 if yes, =0 otherwise)							
Lender/mortgage broker	72.92	1.72		2.27	1.39		2.28
Settlement/closing agent	22.70	1.01		2.17	3.12		2.15
Real estate agent	60.78	– 1.83		2.56	– 0.11		2.44
Personal attorney	15.83	1.04		1.93	-0.68		1.89
Title insurance agent	17.41	3.88	*	2.11	2.83		2.03
Trusted friend or relative who is not a cosigner on the mortgage	44.44	– 1.13		2.56	0.25		2.45
Housing counselor	2.45	9.05	***	1.46	8.91	***	1.34
Any of the sources	88.17	2.52		1.65	1.90		1.61

Exhibit 5

Estimates of the Relationship Between Reported HEC and the Outcome of Interest (2 of 2)

	Mean (%)	OLS		PSM	
		Coefficient (%)	Standard Error (%)	Coefficient (%)	Standard Error (%)
How important were each of the following in choosing the lender/broker you used for the mortgage you took out? (=1 if important or very important)					
Having an established banking relationship	43.98	– 5.70	2.55	– 5.10	2.48
Having a local office or branch nearby	52.04	2.02	2.62	2.13	2.51
Used previously to get a mortgage	7.69	1.71	1.52	1.15	1.52
Lender/broker is a personal friend or relative	12.64	1.92	1.76	3.84 **	1.69
Recommendation from a friend/relative/ coworker	50.07	-0.94	2.61	– 1.08	2.49
Recommendation from a real estate agent/ home builder	51.44	2.05	2.57	3.38	2.50
Reputation of the lender/broker	59.12	3.24	2.55	4.32 *	2.46
Spoke my primary language, which is not English	9.64	1.77	1.66	1.61	1.67
Any of non-price factors are important	91.79	0.31	1.37	– 1.55	1.43
Overall, how satisfied are you that the mortgage you got was the one with the… (=1 if very satisfied)					
Best terms to fit your needs	73.47	0.82	2.33	1.40	2.23
Lowest interest rate for which you could qualify	68.47	0.04	2.41	0.10	2.37
Lowest closing costs	56.53	3.32	2.51	4.43 *	2.44
Any option = very	84.26	4.27 **	1.74	3.52 **	1.77
Overall, how satisfied are you with the… (=1 if very satisfied)					
Lender or mortgage broker you used	72.94	0.96	2.29	0.69	2.22
Application process	59.57	– 3.16	2.60	– 1.39	2.47
Loan closing process	61.71	0.50	2.56	0.66	2.42
Information in the mortgage disclosure documents	61.24	2.22	2.55	2.19	2.41
Timeliness of mortgage disclosure documents	60.35	– 3.12	2.61	– 1.62	2.47
Settlement agent	65.90	– 2.53	2.48	– 1.17	2.39
Overall satisfied with mortgage process	87.19	2.16	1.61	1.83	1.67
Did you face any unpleasant surprises at your loan closing? (=1 for "yes")					
Did you face any unpleasant surprises at your loan closing	16.63	5.83 ***	2.11	5.50 ***	2.05
Administrative data					
Mortgage interest rate spread at origination greater than 100 BPS	12.93	2.38	1.73	2.77	2.05
LTV at origination greater than 95%	44.24	1.44	2.34	0.51	2.50
Mortgage DTI ratio greater than 45%	12.57	– 2.09	1.67	– 1.30	1.86
Second lien	2.14	4.56 ***	1.18	4.36 ***	0.95
Ever 60 days delinquent	3.65	1.41	1.23	0.59	1.20
Ever 90 days delinquent	2.51	0.91	1.05	0.53	1.00

BPS = basis points. DTI = debt-to-income. LTV = loan-to-value. OLS = ordinary least squares. PSM = propensity score matching.
** p<0.05, ** p<0.01, *** p<0.001*
Note: Sample size is 3,305 for OLS and 3,134 for PSM.
Source: NSMO, 2013–2016.

Exhibit 6

Estimates of the Relationship Between Reported HEC and the Outcome of Interest with Selection on LTV, DTI, and Second Lien (1 of 2)

	Mean (%)	OLS			PSM		
		Coefficient (%)		Standard Error (%)	Coefficient (%)		Standard Error (%)
When you began the process of getting your mortgage, how familiar were you (and any cosigners) with each of the following? (=1 if very familiar)							
The mortgage interest rates available at that time	33.58	7.74	***	2.53	5.20	**	2.41
The different types of mortgages available	23.66	6.25	***	2.35	4.78	**	2.21
The process of taking out a mortgage	15.59	4.26	**	2.06	3.89	*	1.99
The downpayment needed to qualify for a mortgage	35.99	0.38		2.48	– 0.44		2.44
The income needed to qualify for a mortgage	32.20	5.12	**	2.52	3.36		2.45
Your credit history or credit score	64.43	7.14	***	2.43	5.68	**	2.41
The money needed at closing	25.20	– 0.25		2.31	– 0.83		2.26
How well could you explain to someone the… (=1 if very well)							
Process of taking out a mortgage	41.32	10.17	***	2.61	8.77	***	2.53
Difference between a fixed and an adjustable-rate mortgage	55.59	7.82	***	2.51	4.80	*	2.54
Difference between a prime and subprime loan	12.44	2.03		1.67	2.27		1.63
Difference between a mortgage's interest rate and its APR	19.17	3.26		2.06	3.05		2.05
Amortization of a loan	23.98	2.48		2.10	3.40		2.10
Consequences of not making required mortgage payments	55.73	9.26	***	2.52	6.88	***	2.51
How many different lenders/brokers did you seriously consider before choosing where to apply for this mortgage? (=1 if more than one lender/broker)							
Seriously considered more than one lender/broker	54.64	5.17	**	2.61	2.08		2.54
How many different lenders/brokers did you end up applying to? (=1 if more than one lender/broker)							
Applied to more than one lender/broker	31.71	4.62	*	2.49	3.47		2.47
Did you seek input about your closing documents from any of the following people? (=1 if yes, =0 otherwise)							
Lender/mortgage broker	72.92	2.47		2.26	1.72		2.32
Settlement/closing agent	22.70	0.50		2.19	3.31		2.17
Real estate agent	60.78	– 2.03		2.59	0.25		2.49
Personal attorney	15.83	1.46		1.94	– 0.55		1.97
Title insurance agent	17.41	3.60	*	2.13	2.80		2.08
Trusted friend or relative who is not a cosigner on the mortgage	44.44	– 1.50		2.58	0.51		2.50
Housing counselor	2.45	9.33	***	1.50	9.29	***	1.35
Any of the sources	88.17	2.83	*	1.60	2.25		1.64

Exhibit 6

Estimates of the Relationship Between Reported HEC and the Outcome of Interest with Selection on LTV, DTI, and Second Lien (2 of 2)

	Mean (%)	OLS			PSM		
		Coefficient (%)		Standard Error (%)	Coefficient (%)		Standard Error (%)
How important were each of the following in choosing the lender/broker you used for the mortgage you took out? (=1 if important or very important)							
Having an established banking relationship	43.98	– 4.64		2.57	– 1.99		2.53
Having a local office or branch nearby	52.04	2.18		2.65	2.14		2.56
Used previously to get a mortgage	7.69	1.71		1.55	1.00		1.57
Lender/broker is a personal friend or relative	12.64	2.09		1.78	3.25	*	1.77
Recommendation from a friend/relative/ coworker	50.07	– 1.24		2.64	– 2.59		2.55
Recommendation from a real estate agent/ home builder	51.44	1.12		2.60	2.11		2.56
Reputation of the lender/broker	59.12	3.22		2.58	4.39	*	2.52
Spoke my primary language, which is not English	9.64	1.76		1.67	1.97		1.67
Any of non-price factors are important	91.79	0.47		1.37	– 0.86		1.49
Overall, how satisfied are you that the mortgage you got was the one with the… (=1 if very satisfied)							
Best terms to fit your needs	73.47	0.81		2.36	1.91		2.28
Lowest interest rate for which you could qualify	68.47	0.51		2.44	0.26		2.41
Lowest closing costs	56.53	2.89		2.54	5.08	**	2.50
Any option = very	84.26	4.39	**	1.77	3.49	*	1.82
Overall, how satisfied are you with the… (=1 if very satisfied)							
Lender or mortgage broker you used	72.94	1.72		2.29	0.78		2.28
Application process	59.57	– 2.65		2.61	– 1.30		2.52
Loan closing process	61.71	0.60		2.59	1.05		2.47
Information in the mortgage disclosure documents	61.24	2.35		2.57	1.97		2.46
Timeliness of mortgage disclosure documents	60.35	– 2.92		2.63	– 2.02		2.52
Settlement agent	65.90	– 2.71		2.51	– 0.03		2.46
Overall satisfied with mortgage process	87.19	2.33		1.61	1.73		1.70
Did you face any unpleasant surprises at your loan closing? (=1 for "yes")							
Did you face any unpleasant surprises at your loan closing	16.63	5.24	**	2.15	4.90	**	2.09
Administrative data							
Mortgage interest rate spread at origination greater than 100 BPS	12.93	2.81		1.76	2.28		2.10
Ever 60 days delinquent	3.65	1.20		1.21	0.82		1.20
Ever 90 days delinquent	2.51	0.73		1.01	0.66		1.04

BPS = basis points. DTI = debt-to-income. HEC = homebuyer education and counseling. LTV = loan-to-value. OLS = ordinary least squares. PSM = propensity score matching.
** p<0.05, ** p<0.01, *** p<0.001*
Note: Sample size is 3,260 for OLS and 3,012 for PSM.
Source: NSMO, 2013–2016.

We also analyze whether HEC had any effect on early loan performance. We measure loan performance as ever being 60 or more days delinquent since origination. For loan originations to first-time homebuyers, this delinquency rate is 3.65 percent. We consider the full payment history available for any given origination, resulting in comparatively longer performance timelines for earlier originations. Without controls, the difference in the delinquency rate between those reporting HEC and those not reporting HEC is 3.19 percentage points, 6.31 percent compared with 3.12 percent, as shown in exhibit 2. In both the OLS and PSM modeling frameworks, the performance gap between the reported HEC and non-HEC group decreases when controlling for observable differences between the groups. The difference between HEC recipients and other first-time homebuyers is not statistically significant. Our finding of no effect is in line with Smith, Hochberg, and Greene (2014).

It may be tempting to conclude that counseling does not matter for first-time homebuyers. However, we should recognize that selection into counseling or reported HEC is not by random assignment. In fact, the reported HEC group may not have been eligible for a mortgage without completing some form of HEC. The reported HEC group was selected because they were determined to be a potentially higher risk than even observably similar non-HEC group members. This is consistent with the higher rates of counseling observed among the traditionally riskier borrower characteristics in exhibit 3.

Although encouraging, these results, particularly those with respect to mortgage knowledge, should be interpreted with caution for two reasons.

First, the outcomes on mortgage knowledge are self-reported, and it has been found that consumers, in some instances, may overestimate their actual financial literacy (Lusardi and Mitchell, 2011). For instance, counseling may make consumers more confident in their knowledge, as opposed to actually improving it. Further research is needed linking receipt of HEC to actual, rather than self-reported, knowledge. As noted earlier, early results by DeMarco et al. (2016) of the ongoing HUD experiment indicate a modest positive effect of HEC on respondents' performance in a four-question financial literacy test.

Second, consumers may select into HEC along dimensions related to mortgage knowledge. For instance, one plausible hypothesis is that borrowers who feel less confident in their mortgage knowledge would be more likely to use HEC, as they stand to gain more from this type of education. It is worth emphasizing that our results reject this hypothesis. Almost all the coefficients for the relationship between HEC and self-reported mortgage knowledge are positive, even if not all are statistically significant. As pointed out by Collins and O'Rourke (2011, 2010) at the time of their review, the literature was inconclusive as to which direction the selection would operate. Our results point in the direction of positive selection.

Conclusion

HEC is believed to improve the homebuying process for potential homebuyers. Although we did not administer a mortgage literacy assessment, we do find evidence consistent with improved familiarity and confidence with the mortgage process and related terminology. Moreover, first-time homebuyers who reported receiving HEC had a higher level of satisfaction with their mortgage. The results around loan performance, as measured by delinquencies, are less conclusive. Subsequent analyses could continue to monitor mortgages reported in the NMDB and NSMO data for patterns during different economic conditions.

A key limitation of this survey in identifying the effect of HEC on mortgage knowledge is that we are unable to determine whether the responses on the survey reflect the postpurchase (and thus post-HEC) state of knowledge or the prepurchase state of knowledge. Future research should employ methods to distinguish between the pre-HEC and post-HEC states of knowledge. By examining within-person changes in knowledge, one may be able to eliminate most if not all concerns related to selection into HEC that may be related to mortgage knowledge. An example of this approach is Carswell (2009): measures of financial distress were obtained before and after counseling, with an aim of identifying the effect of HEC on within-person changes in financial distress.

Acknowledgments

The authors thank Robert B. Avery, Xian Fang Bak, Ron Borzekowski, Daniel E. Coates, Thomas Daula, Samuel Frumkin, Ian H. Keith, Michael LaCour-Little, Gen Melford, David Sanchez, Susan Singer, Rebecca Sullivan, Cynthia Waldron, Stacey Walker, and Peter Zorn for providing comments and Julia Nguyen for assisting with data. Grace Li was a senior economist for Fannie Mae at the time this article was written.

Authors

Robert B. Argento is a quantitative analytics senior at Freddie Mac.

Lariece M. Brown is a quantitative analytics director at Freddie Mac.

Sergei Koulayev is an economist at the Consumer Financial Protection Bureau.

Grace Li is a statistical quantitative senior at SunTrust Investment Services.

Marina Myhre is a social science analyst at the U.S. Department of Housing and Urban Development.

Forrest Pafenberg is a former supervisory economist at the Federal Housing Finance Agency.

Saty Patrabansh is the manager of the National Mortgage Database Program at the Federal Housing Finance Agency.

References

Agarwal, Sumit, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, and Douglas D. Evanoff. 2014a. The Effectiveness of Mandatory Mortgage Counseling: Can One Dissuade Borrowers from Choosing Risky Mortgages? NBER Working Paper No. 19920. Cambridge, MA: National Bureau of Economic Research.

———. 2014b. "Predatory Lending and the Subprime Crisis." *Journal of Financial Economics* 113 (1): 29–52.

———. 2010. "Learning to Cope: Voluntary Financial Education Programs and the Housing Crisis," *American Economic Review: Papers and Proceedings* 100 (2): 495–500.

Alexandrov, Alexei, and Sergei Koulayev. 2017. No Shopping in the U.S. Mortgage Market: Direct and Strategic Effects of Providing Information. Working paper 2017-01. Washington, DC: Consumer Financial Protection Bureau Office of Research.

Avila, Gabriela, Hoa Nguyen, and Peter Zorn. 2013. The Benefits of Pre-Purchase Homeownership Counseling. Working paper (April). Tysons Corner, VA: Freddie Mac.

Brown, Scott R. 2016. "The Influence of Homebuyer Education on Default and Foreclosure Risk: A Natural Experiment," *Journal of Policy Analysis and Management* 35 (1): 145–172.

Carswell, Andrew T. 2009. "Does Housing Counseling Change Consumer Financial Behaviors? Evidence from Philadelphia," *Journal of Family and Economic Issues* 30 (4): 339–356.

Collins, J. Michael 2013. "The Impacts of Mandatory Financial Education: Evidence from a Randomized Field Study," *Journal of Economic Behavior & Organization* 95: 146–158.

Collins, J. Michael, and Collin M. O'Rourke. 2011. Homeownership Education and Counseling: Do We Know What Works? Research Institute for Housing America Research Paper No. 1102.

———. 2010. "Financial Education and Counseling—Still Holding Promise," *Journal of Consumer Affairs* 44 (3): 483–498.

DeMarco, Donna, Nichole Fiore, Debbie Gruenstein Bocian, Shawn Moulton, Laura Peck, and Abt Associates Inc. 2016. *The First-Time Homebuyer Education and Counseling Demonstration: Early Insights*. Washington, DC: U.S. Department of Housing and Urban Development Office of Policy Development and Research. https://www.huduser.gov/portal/sites/default/files/pdf/First-Time-Home-Buyers.pdf.

Gerardi, Kristopher, Lorenz Goette, and Stephan Meier. 2010. Financial Literacy and Subprime Mortgage Delinquency: Evidence from a Survey Matched to Administrative Data. Working paper 2010-10. Atlanta, GA: Federal Reserve Bank of Atlanta.

Li, Wei, Bing Bai, Laurie Goodman, and Jun Zhu. 2016. *NeighborWorks America's Homeownership Education and Counseling: Who Receives It and Is It Effective?* Research report. Washington, DC:

Urban Institute. http://www.urban.org/sites/default/files/publication/84476/2000950-NeighborWorks-America's-Homeownership-Education-and-Counseling-Who-Receives-It-and-Is-It-Effective.pdf.

Lusardi, Annamaria, and Olivia S. Mitchell. 2014. "The Economic Importance of Financial Literacy: Theory and Evidence," *Journal of Economic Literature*, 52 (1): 5–44.

———. 2011. "Financial Literacy and Retirement Planning in the United States," *Journal of Pension Economics & Finance* 10 (4): 509–525.

Mayer, Neil S., and Kenneth Temkin. 2016. "Prepurchase Counseling Effects on Mortgage Performance: Empirical Analysis of NeighborWorks® America's Experience," *Cityscape* 18 (2): 73–98.

Moulton, Stephanie, J. Michael Collins, Cäzilia Loibl, and Anya Samak. 2015. "Effects of Monitoring on Mortgage Delinquency: Evidence From a Randomized Field Study," *Journal of Policy Analysis and Management* 34 (1): 184–207.

Myhre, Marina L., and Nicole Elsasser Watson. 2017. *Housing Counseling Works*. Washington, DC: U.S. Department of Housing and Urban Development Office of Policy Development and Research. https://www.huduser.gov/portal/sites/default/files/pdf/Housing-Counseling-Works.pdf.

Quercia, Roberto G., and Jonathan Spader. 2008. "Does Homeownership Counseling Affect the Prepayment and Default Behavior of Affordable Mortgage Borrowers?" *Journal of Policy Analysis and Management* 27 (2): 304–325.

Roll, Stephen, and Stephanie Moulton. 2016. The Impact of Credit Counseling on Consumer Outcomes: Evidence from a National Demonstration Program. Washington, DC: Federal Deposit Insurance Corporation. http://gflec.org/wp-content/uploads/2016/04/Roll-Stephen-and-Moulton-Stephanie-The-Impact-of-Credit-Counseling-on-Consumer-Outcomes.pdf.

Smith, Marvin M., Daniel Hochberg, and William H. Greene. 2014. *The Effectiveness of Pre-Purchase Homeownership Counseling and Financial Management Skills*. Philadelphia, PA: Federal Reserve Bank of Philadelphia.

Spader, Jonathan, and Roberto G. Quercia. 2009. *Pre-purchase Homeownership Counseling and Mortgage Search*. Working paper. Chapel Hill, NC: UNC Center for Community Capital.

Turnham, Jennifer, and Anna Jefferson. 2012. *Pre-Purchase Counseling Outcome Study: Research Brief Housing Counseling Outcome Evaluation*. Research brief. Washington, DC: U.S. Department of Housing and Urban Development. https://www.huduser.gov/publications/pdf/pre_purchase_counseling.pdf.

National Survey of Mortgage Originations Survey Data on Your Home Loan Toolkit

Brian Bucks
Tim Critchfield
Susan Singer
Consumer Financial Protection Bureau

Disclaimer
The views expressed in this article are those of the authors and are not necessarily those of the Consumer Financial Protection Bureau or the U.S. Government.

Abstract

The Real Estate Settlement Procedures Act (RESPA) requires that consumers who apply for a mortgage to purchase a home receive certain information on real estate settlement services and related costs. Before 2015, the U.S. Department of Housing and Urban Development (HUD) published a booklet titled ***Shopping for Your Home Loan: Settlement Cost Booklet*** *(Booklet), which lenders provided to mortgage applicants. Starting in October 2015, coincident with the implementation of new mortgage loan estimate and closing disclosure forms, lenders began providing the Consumer Financial Protection Bureau's substantially revised publication, the* ***"Your Home Loan Toolkit"*** *(Toolkit) (CFPB, n.d.1).*

This article details the extent to which mortgage applicants recalled getting the Toolkit. Because the Toolkit is required only for purchase mortgages, the analysis considers purchase-mortgage borrowers' recollection of receiving the Booklet or the Toolkit. We find that the percentage of homebuyers who remember receiving the information increased with the introduction of the new Toolkit. The estimated marginal effects are statistically significant and large compared with other factors

Introduction

Getting a mortgage is a complex financial decision and, for many consumers, it is the biggest and most complex financial decision they will make. One of the primary federal regulations that relates to the acquisition of residential mortgage is the Real Estate Settlement Procedures Act (RESPA). RESPA requires that consumers who apply for a mortgage to purchase a home or to refinance an existing loan receive certain information on real estate settlement services and related costs. RESPA requires that purchase-mortgage borrowers get a "special information booklet" that informs them about real estate settlement services (12 U.S.C. 2601 *et seq.* and its implementing Regulation X, 12 CFR 1024.6).

From 1976 to 2015, HUD implemented this RESPA provision by requiring lenders to provide purchase-mortgage applicants with an information booklet titled "Shopping for Your Home Loan: HUD's Settlement Cost Booklet" (Booklet). The Dodd-Frank Wall Street Reform and Consumer Protection Act (Dodd-Frank Act) transferred the responsibility for the information booklet to the Consumer Financial Protection Bureau (CFPB, n.d.1).[1] On October 3, 2015, new RESPA rules were implemented, including for the CFPB special information booklet. The CFPB booklet is titled "Your Home Loan Toolkit: A Step-by-Step Guide" (Toolkit). RESPA requires that lenders and brokers provide homebuyers a copy of the booklet within 3 days of submitting an application. RESPA does not require the booklet to be given to consumers who are refinancing a residential mortgage.[2]

A comparison of both booklets shows clear differences in the content and format. The Booklet walks the applicant through the Good Faith Estimate (GFE) and HUD-1 Settlement Statement line by line. The Toolkit does not walk the applicant through the Loan Estimate line by line[3] and, although it provides detailed information on the Closing Disclosure, it does so in less detail than does the Booklet. Also, the Dodd-Frank Act listed specific topics to be described and explained in the CFPB informational document.[4] Some of those topics, such as an explanation of variable rate mortgages and home equity lines of credit, had been included in the Booklet. Further, the Dodd-Frank Act required that the information booklet include questions that a consumer should ask regarding the loan, including, among other things, whether the consumer would have the ability to pay back the loan. To fulfill that requirement, the Toolkit includes "The Talk" boxes throughout, which include questions and factors that consumers should consider.

In this article, we detail the shares of mortgage applicants who recalled getting the CFPB Toolkit. We compare those results with the shares of mortgage applicants who remembered getting the earlier Booklet to understand how the recollection of receiving this information changed after implementation of the new Toolkit. We look further at how changes in borrowers' recall of those documents varied by borrower and loan characteristics (for example, loan amount, borrower age, and credit score). The analysis uses consumer responses from the National Survey of

[1] Dodd-Frank Act §1450. The Booklet is available at https://www.hud.gov/sites/documents/HUD_SETTLE_COST.PDF, and the Toolkit is available at https://files.consumerfinance.gov/f/201503_cfpb_your-home-loan-toolkit-web.pdf.
[2] Lenders and mortgage brokers do not have to provide the booklet during refinancing transactions; closed-end loans, such as when the lender takes a subordinate lien; reverse mortgages; and any other federally related mortgage loan whose purpose is not the purchase of a one- to four-family residential property (12 CFR §1024.6(3)).
[3] The switch from the Booklet to the Toolkit was done as part of the larger Truth in Lending Act (TILA)-RESPA Integrated Disclosures Rule, which, among other things, combined the GFE and the initial TILA disclosure and replaced them with a single Loan Estimate.
[4] For a comprehensive list of topics, see Dodd-Frank Act §1450(2)(b).

Mortgage Originations (NSMO), a nationally representative survey that provides rich information on the expectations, knowledge, experiences, and loan terms for consumers with a recent mortgage origination.[5]

We find that the percentage of homebuyers who remember receiving information from their lender increased with the introduction of the Toolkit. Specifically, the percentage of homebuyers who remembered receiving the documents increased from 22 percent for the Booklet to 42 percent for the Toolkit. In general, the increase in the share of homebuyers who remember the information document tended to be larger for borrowers with lower credit scores, household income, education, or experiences with mortgages. Using a multinomial logit model to better isolate the role of the Toolkit in the increase in borrower recognition, we find that controlling for borrower and loan characteristics in NSMO does not meaningfully change the estimated effect of the Toolkit on borrower recall. The estimated marginal effects of the Toolkit indicator are statistically significant and large compared with the other covariates in the model.

Literature Review

Choosing and closing on a mortgage is a complex process and one that most consumers do only infrequently. That infrequency, in turn, may contribute to many consumers not understanding the terms and conditions of their mortgage. Bucks and Pence (2008) explore whether borrowers understand the terms and conditions of their adjustable rate mortgages and find that consumers underestimate or do not understand the extent of possible rate increases from year to year or over the life of the loan. Lacko and Pappalardo (2010) found that many borrowers do not understand the terms of their loans, including the total loan amount and whether their loan had restrictions, such as prepayment penalties. Borrowers' lack of understanding of their mortgage rates and terms could, in part, reflect asymmetric information and search costs. Mortgage professionals engage in transactions on a frequent and regular basis but may find information costly to produce, and once produced, it becomes a common good. Further, consumers may engage in only a limited search across mortgage lenders because their perceived expected benefits from such a search are less than their expected costs.[6]

Various ways exist to increase knowledge, reduce asymmetric information, and encourage consumers to search for the rates, terms, and conditions that meet their needs. Those ways include mandatory standardized disclosure and homeownership education and counseling (HEC). Numerous studies have been conducted on prepurchase HEC programs.[7] Those studies usually evaluate specific programs that generally target consumers with low income or low credit scores and focus on whether HEC programs reduce mortgage default rates (see, for example, Agarwal et al. [2010]). Elsewhere in this volume, Argento and his colleagues (2018) evaluated NSMO survey respondents' use of housing counseling and found that those who reported receiving some form of HEC had better mortgage knowledge and higher satisfaction with their mortgage terms and the process.

Standardized disclosures of loan terms and features may also increase consumer knowledge and understanding of mortgage terms and conditions. Lacko and Pappalardo (2010) conducted a

[5] NSMO and the National Mortgage Database® (NMDB) are described in the guest editors' introduction of this Cityscape issue.
[6] See Alexandrov and Koulayev (2017); Woodward and Hall (2010).
[7] Collins and O'Rourke (2010), Moulton (2012), and Myhre and Watson (2017) provide a review of many of these studies.

randomized control experiment to test borrower understanding using either the mandatory TILA and RESPA mortgage disclosures or prototype disclosures developed for the study. They found that borrowers presented with the prototype disclosures had a higher level of understanding of the mortgage terms, but the borrowers could not fully describe important loan terms.[8]

Consumers have other ways to gain knowledge and be encouraged to search, including talking with family and friends, engaging with websites on the homebuying process, and reading homebuying handbooks provided by government agencies, such as the HUD and CFPB booklets.[9]

Virtually no studies focus on the effects of the various types of knowledge acquisition by consumers. One exception, the CFPB's *Know Before You Owe Brief No. 3*, found evidence of "learning by doing" in the mortgage context. The study looked at changes in mortgage knowledge throughout the homebuying process and found evidence that consumers became more knowledgeable through the process, but it does not attribute this gain to any source in particular.[10]

Changes in Responses to Whether Borrower Remembers Booklet or Toolkit

This article contributes to the literature by exploring whether purchase-mortgage borrowers were more likely to recall the Booklet or the Toolkit and whether those documents led borrowers to ask questions about their loans. These comparisons are intended as a simple test of whether the new Toolkit is more salient and thus may be more effective in informing consumers. Further, we consider how the changes in borrowers' ability to recall the booklet varied across groups based on borrower or loan characteristics.

Beginning in the third quarter of 2015, in anticipation of the replacement of the Booklet with the Toolkit, the NSMO questionnaire asked whether the respondent had received a copy of *Shopping for Your Home Loan: HUD's Settlement Cost Booklet* (HUD, n.d.). A year later, when the survey sample was comprised mostly of mortgages taken out in early 2016 and some from late 2015, the question was changed to ask whether the respondent remembered receiving a copy of *Your Home Loan Toolkit: A Step-by-Step Guide* (CFPB, n.d.).[11] The analysis contrasts pooled responses to the four quarterly NSMO surveys before the change to pooled responses for the subsequent five quarters, through the third quarter of 2017.

The analysis is limited to mortgages for home purchase because lenders were only required to provide the Booklet or Toolkit for purchase transactions. The sample excludes refinance loans,

---

[8] In integrating the TILA and RESPA mortgage disclosures, the CFPB used a User Center Design process with a qualitative and a quantitative phase (Kleimann, 2012; 2013).
[9] The CFPB provides a webpage on mortgages that includes tools and resources for homebuyers (https://www.consumerfinance.gov/consumer-tools/mortgages/). State and city governments also provide homebuyer guides, for example, Minnesota (https://www.ag.state.mn.us/consumer/handbooks/HmBuyers/default.asp), New York City (https://cnycn.org/homebuyers-guide/), and Ohio (https://www.com.ohio.gov/documents/real_HomeBuyersGuide.pdf).
[10] For the study analyzed in the CFPB's Know Before You Owe Brief No. 3 (CFPB n.d.2), prospective homebuyers took an initial baseline survey and then were surveyed biweekly for up to 3 months or until they bought a home, whichever occurred first.
[11] The questions differed slightly over time, as the earlier question asked whether the respondent had received the booklet ("Your lender may have given you a *Shopping for Your Home Loan: HUD's Settlement Cost Booklet*, did you receive a copy?"), and the later question whether the borrower remembered receiving it ("Your lender may have given you a booklet, ***Your Home Loan Toolkit: A Step-by-Step Guide***, do you remember receiving a copy?").

loan modifications, and special-purpose mortgages (loans to add or remove a co-borrower, loans to finance a construction loan, and new mortgages on mortgage-free properties).

Exhibit 1 shows that an average of 22 percent of NSMO respondents reportedly recalled receiving the Booklet.[12] After the new Toolkit took effect, the share who recalled getting it jumped to 42 percent on average. The share of respondents who reported that they did not know whether they received the document dropped from 38 percent to 27 percent.

Exhibit 1

Borrower Recall of Mortgage Booklet and Toolkit (Percent)

	Don't Know	Yes	No
Booklet	38	22	39
Toolkit	27	42	30

Notes: The difference in the distribution of responses is statistically significant at a 0.1-percent level. The sample includes 3,083 purchase-mortgage borrowers in the 2015Q3–2016Q2 NSMO surveys and 2,722 purchase-mortgage borrowers in the 2016Q3–2017Q3 surveys.

The NSMO asked a followup question about whether the document led to the homebuyer asking questions about his or her mortgage in hopes of determining how effective the documents were at getting respondents to think about their mortgage. Exhibit 2 tabulates responses to this followup question for respondents who recalled getting the document (top panel) and for all survey respondents (bottom panel).

Exhibit 2

Share of Borrowers Who Asked Questions Based on Booklet or Toolkit (Percent)

	Yes	No
Borrowers Who Recalled Booklet or Toolkit		
Booklet	29	71
Toolkit	30	70
All Borrowers		
Booklet	6	16
Toolkit	13	30

Notes: The sample includes 700 purchase-mortgage borrowers in the 2015Q3–2016Q2 NSMO surveys and 1148 purchase-mortgage borrowers in the 2016Q3–2017Q3 surveys. "Not applicable" responses for borrowers who said "No" or "Don't Know" when asked whether they recalled the Booklet or Toolkit are not shown.

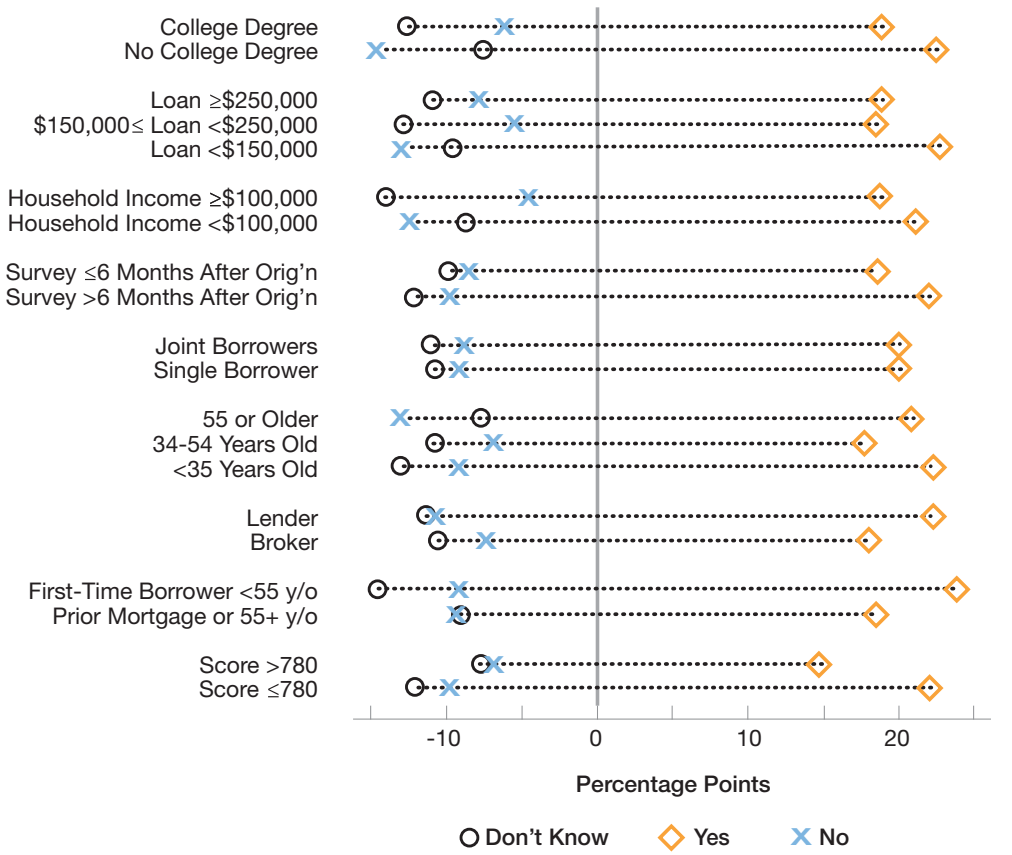
[12] All statistics are weighted using nonresponse-adjusted analysis weights.

Overall, there was a statistically insignificant change in the likelihood borrowers asked questions, conditional on remembering the documents. The share who reportedly asked questions rose from an average of 29 percent before the change to 30 percent after the Toolkit was introduced; roughly 70 percent of respondents did not ask additional questions in either time period. Because the share of borrowers who recalled the Booklet or Toolkit increased, the percentage of all borrowers who asked questions based on these documents roughly doubled, from 6 percent for the Booklet to 13 percent for the Toolkit.

Exhibit 3 breaks down the changes in the estimates shown in Exhibit 1 after the Toolkit was introduced for subgroups defined by borrower or loan characteristics. It shows, specifically, the differences in the shares of respondents within subgroups defined by borrower or loan characteristics who said, "Don't Know," "Yes," or "No." In each row of the exhibit, the likelihood that respondents said they remember the new Toolkit increased by 15 percentage points or more.

Exhibit 3

Changes in Borrower Recall of Receiving Mortgage Disclosure, by Borrower and Loan Characteristics



In general, the increases in the shares of "Yes" tended to be larger for groups of borrowers with lower credit scores, household income, loan amount, education, or experience with mortgages.[13] The increase in "Yes" responses was also greater for respondents who worked directly with a lender rather than a broker. There was little difference in the increases for joint and single borrowers and those who got the survey more than 6 months after they took out the mortgage. For all groups, the percentages of "No" and "Don't Know" responses both declined noticeably after the new Toolkit was introduced, and in most cases the share of "Don't Know" answers dropped by more than that of "No" answers did.[14]

Multivariate Analysis

To better isolate the role of the Toolkit in the increase in borrower recollection shown in exhibits 1 and 3, we estimate a multinomial logit model of whether the respondent recalled the Booklet or the Toolkit ("Don't Know," "Yes," or "No"). The models include the covariates shown in exhibit 3, a time trend, and controls for other factors that might be correlated with borrowers' recall of the Toolkit or Booklet. These factors include debt-to-income ratio, loan-to-value, interest rate spread, loan type, the respondent's labor force status, the respondent's race and ethnicity, and household composition.[15] In addition, the models incorporate information on borrowers' knowledge and views as they began the mortgage process, concern about qualifying, certainty about the type of mortgage they wanted, and knowledge about mortgage terms and the mortgage process.[16] Together, these controls serve to make sure that any measured change in borrower recall is not the result of shifts in risk or in borrower characteristics and attitudes.

Exhibit 4 shows the estimated marginal effects on the probability of each response for selected covariates.[17] Controlling for these additional factors does not meaningfully change the estimated effect of the Toolkit on borrower recall. The Toolkit indicator leads to a significant estimated decline in the likelihood of "Don't Know" (8 percentage points) and "No" (13 percentage points) and a corresponding 22-percentage-point increase for "Yes." These changes are similar to the unadjusted differences of 11, 9, and 20 percentage points (see exhibit 1), respectively.

---

[13] The score categories are based on the VantageScore® 3.0 for the survey respondent or his or her spouse or partner and the average of these scores when both are available. Income is self-reported household income. Age and education are those for the respondent and are self-reported values from the survey. First-time borrowers are people younger than 55 years of age who did not have any record of having a mortgage in the Experian data in the 7 years before the mortgage in consideration and who were buying a house in which they will primarily live. The category "Survey >6 months after orig'n" comprises loans with an origination date 180 days or more before the initial survey mailing for the respective survey wave.

[14] Exhibit A.2 presents percentage point estimates for the differences shown in exhibit 3; the distributions of responses are significantly different before and after the introduction of the Toolkit, at a 0.1-percent level for all 20 groups shown.

[15] See Exhibit A.1 for additional detail on the model covariates.

[16] The NSMO also asks several questions about borrowers' knowledge after they took out the mortgage; those responses are excluded from the models because that knowledge might have been affected by receipt of the Booklet or Toolkit.

[17] Exhibit A.3 presents marginal effects and corresponding standard errors for all covariates.

Exhibit 4

Marginal Effects on Probability of Responses to Recall of Booklet or Toolkit, Selected Covariates (Percentage Points)

	Don't Know	Yes	No
Toolkit	– 8.1**	21.6***	– 13.4***
Credit Score	0.03	– 0.02	– 0.007
Age	– 0.3***	0.3***	0.07
Days Since Origination	0.008	– 0.002	– 0.006
Loan Amount <$50K	– 13.7**	5.8	7.9
Loan Amount $50K–$99.9K	– 4.1	0.9	3.2
Loan Amount $100K–$149.9K	– 2.8	2.2	0.6
Loan Amount $200K–$249.9K	2.3	– 1.0	– 1.3
Loan Amount $250K–$299.9K	1.4	– 3.9	2.5
Loan Amount $300K–$349.9K	– 1.6	– 2.3	3.9
Loan Amount $350K–$399.9K	4.2	– 1.0	– 3.1
Loan Amount $400K or More	0.5	2.8	– 3.3
Rate Spread	1.3	– 4.3***	3.0*
Loan-to-Value	– 0.08	0.2**	– 0.07
First-Time Borrower <55 Years Old	– 2.4	4.8*	– 2.4
Not Owner-Occupied	– 1.9	– 1.3	3.2
Broker	– 1.9	0.9	1.0
Joint Borrowers	– 0.5	0.9	– 0.4
Non-Hispanic White	2.8	– 6.9***	4.1*
Somewhat Concerned Qualified	6.9***	– 3.2	– 3.6
Not at All Concerned Qualified	4.5*	– 4.6*	0.07
Some Idea of Desired Mortgage	– 1.0	– 1.2	2.2
No Idea of Desired Mortgage	5.3	_ 7.9**	2.6

** p<0.05, ** p<0.01, *** p<0.001*

The Toolkit indicator is the only variable with a statistically significant effect on all three response outcomes, and those effects are generally larger than the estimated effects for other covariates. The effects, for example, on the indicator for whether the respondent was a non-Hispanic White are about one-third the magnitude of the effects for the Toolkit, and the marginal effects for the indicator for the respondent having no firm idea about the type of mortgage he or she wanted are two-thirds to one-fifth. The estimated nearly 14-percentage-point reduction in the likelihood of a "Don't Know" response for borrowers with a loan amount of less than $50,000 compared with a loan amount of $150,000–$199,999 (the omitted category) is the exception in having a larger effect than the Toolkit. Qualitatively, the sign of the marginal effects for the Toolkit is comparable to the effects for age. In contrast, non-Hispanic White borrowers, those who were less concerned about qualifying or had "no idea" about the desired mortgage, or those who had mortgage loans with a larger rate spread were more likely to say they did not recall or didn't know if they remembered getting the Booklet or Toolkit.

Conclusion

This article uses the rich information and quarterly data from the NSMO to examine whether purchase-mortgage borrowers were more likely to recall the Toolkit than the Booklet, which the Toolkit replaced in October 2015. The comparison of borrowers' reported ability to recall and ask questions based on this disclosure is intended as a simple test of whether the new Toolkit is more salient and thus may be more effective in informing consumers.

Borrowers were about 20 percentage points more likely to recall the Toolkit than they were the Booklet. There was no meaningful difference in the likelihood that borrowers who remembered the disclosures asked followup questions of their lender, but the increased share of borrowers who recall the Toolkit implies a significant increase in the overall share of borrowers prompted to ask questions as a result. The differences in recall hold within subsets of borrowers classified based on demographic, socioeconomic, and loan characteristics. Similarly, controlling for a variety of borrower and loan characteristics does not meaningfully change the estimated effect of the Toolkit on borrower recall, and the marginal effect of the Toolkit flag is statistically significant and large compared with the estimated marginal effects for other covariates.

On the whole, these findings suggest that the design or content of the Toolkit made the information more salient to borrowers. The large effect—both in absolute terms and compared with other covariates—of the switch to the Toolkit on borrower recall suggests that testing of informational and educational materials for consumer financial products may improve borrower engagement with the content. Although the NSMO questions cannot shed light on which specific aspects of information in the Toolkit were especially effective, more in-depth surveys or focus groups may be able to distinguish these key features.

Appendix

Exhibit A.1

Sample Statistics (1 of 2)

Variable	Description	Mean	Min	Max
1.toolkit	Loan sampled in 2016Q3 or later	0.506	0	1
survey_wave	NSMO survey wave	10.5	7	15
mnscore	Avg. credit score for respondent & spouse/ partner (if applic.)	731.1	502.5	839
x74r	Respondent age	42.3	19	93
daystomail	Days from loan orig'n to 1st survey mailing	174.1	55	495
1.loan_amount_cat	Loan amt at orig'n <$50K	0.021	0	1
2.loan_amount_cat	Loan amt at orig'n $50K–$99.9K	0.129	0	1
3.loan_amount_cat	Loan amt at orig'n $100K–$149.9K	0.195	0	1
5.loan_amount_cat	Loan amt at orig'n $200K–$249.9K	0.142	0	1
6.loan_amount_cat	Loan amt at orig'n $250K–$299.9K	0.106	0	1
7.loan_amount_cat	Loan amt at orig'n $300K–$349.9K	0.069	0	1
8.loan_amount_cat	Loan amt at orig'n $350K–$399.9K	0.048	0	1
9.loan_amount_cat	Loan amt at orig'n $400K or more	0.120	0	1
rate_spread	Interest rate spread at origination	0.263	–1.5	1.5
ltv	Loan-to-value at origination	83.8	11	125
dtirecode	Debt-to-income at orig'n (missing set to 0)	35.5	0	100
1.missdti	Missing debt-to-income (DTI)	0.007	0	1
1.firsttime	Owner-occ,1st-time borrow <55yr	0.373	0	1
4.firsttime	Not owner-occupied	0.093	0	1
1.broker	Loan through broker (vs. lender or builder)	0.430	0	1
1.joint	Two or more borrowers	0.453	0	1
1.x83	Total household income less than $35K	0.059	0	1
2.x83	Total household income $35K–$49.9K	0.111	0	1
3.x83	Total household income $50K–$74.9K	0.202	0	1
5.x83	Total household income $100K–$174.9K	0.277	0	1
6.x83	Total household income $175K or more	0.160	0	1
2.x76rrecode	Respondent high school grad or less	0.104	0	1
3.x76rrecode	Respondent technical school	0.054	0	1
4.x76rrecode	Respondent some college	0.179	0	1
6.x76rrecode	Respondent postgraduate studies	0.280	0	1
1.rwhitenh	Respondent non-Hispanic White	0.765	0	1
2.loan_type	FHA insured	0.221	0	1
3.loan_type	VA guaranteed	0.081	0	1
4.loan_type	FSA/RHS insured	0.035	0	1
2.x05a	Somewhat familiar: interest rates available	0.390	0	1

Exhibit A.1

Sample Statistics (2 of 2)

Variable	Description	Mean	Min	Max
3.x05a	Not at all familiar: interest rates available	0.094	0	1
2.x05b	Somewhat familiar: types of mortgages avail.	0.423	0	1
3.x05b	Not at all familiar: types of mortgages avail.	0.152	0	1
2.x05c	Somewhat familiar: mortgage process	0.407	0	1
3.x05c	Not at all familiar: mortgage process	0.188	0	1
2.x05d	Somewhat familiar: downpayment to qualify	0.357	0	1
3.x05d	Not at all familiar: downpayment to qualify	0.085	0	1
2.x05e	Somewhat familiar: income needed to qualify	0.382	0	1
3.x05e	Not at all familiar: income needed to qualify	0.100	0	1
2.x05f	Somewhat familiar: own credit history, score	0.215	0	1
3.x05f	Not at all familiar: own credit history, score	0.030	0	1
2.x05g	Somewhat familiar: money needed at closing	0.388	0	1
3.x05g	Not at all familiar: money needed at closing	0.159	0	1
2.x06	Somewhat concerned about qualifying	0.339	0	1
3.x06	Not at all concerned about qualifying	0.470	0	1
2.x07	Some idea of mortgage wanted	0.393	0	1
3.x07	Little idea of mortgage wanted	0.096	0	1
1.rworkstatus	Respondent self-employed	0.096	0	1
3.rworkstatus	Respondent retired	0.084	0	1
4.rworkstatus	Respondent not working	0.045	0	1
1.x81a	(Grand)Children <18 in household	0.452	0	1
1.x81b	(Grand)Children 18–22 in household	0.067	0	1
1.x81c	(Grand)Children 23+ in household	0.047	0	1
1.x81def	Other relatives or non-relatives in household	0.089	0	1

Exhibit A.2

Changes in Borrower Recall of Receiving Mortgage Disclosure by Borrower and Loan Characteristics (Percentage Points)

	Don't Know	Yes	No
Score ≤780	–12	22	–10
Score >780	– 8	15	– 7
Not First-Time or 55+	– 9	18	– 9
First-Time <55	– 15	24	– 9
Lender	– 11	18	– 7
Broker	– 11	22	– 11
<35 Years Old	– 13	22	– 9
35–54 Years Old	– 11	18	– 7
55 or Older	– 8	21	– 13
Single	– 11	20	– 9
Joint	– 11	20	– 9
Orig'n >6 Mos.	– 12	22	– 10
Orig'n ≤6 Mos.	– 10	19	– 9
Income <100K	– 9	21	– 13
Income ≥100K	– 14	19	– 5
Loan <$150K	– 10	23	– 13
<35 Years Old	– 13	22	– 9
35–54 Years Old	– 11	18	– 7
55 or Older	– 8	21	– 13
Less Than College	– 8	22	– 15
College or More	– 13	19	– 6

Exhibit A.3

Marginal Effects (ME) and Standard Errors (SE) from Multinomial Logit Model of Responses to Recall of Booklet or Toolkit (Percentage Points)

Variable	Don't Know		Yes		No	
	ME	SE	ME	SE	ME	SE
1.toolkit	− 8.1	2.8**	21.6	2.7***	− 13.4	2.8***
survey_wave	− 0.8	0.6	− 0.5	0.6	1.2	0.6
mnscore	0.03	0.01	− 0.02	0.01	− 0.007	0.01
x74r	− 0.3	0.08***	0.3	0.07***	0.07	0.08
daystomail	0.008	0.01	− 0.002	0.01	− 0.006	0.01
1.loan_amount	− 13.7	4.5**	5.8	6	7.9	5.8
2.loan_amount	− 4.1	2.7	0.9	2.7	3.2	2.8
3.loan_amount	− 2.8	2.3	2.2	2.3	0.6	2.4
5.loan_amount	2.3	2.5	− 1	2.4	− 1.3	2.5
6.loan_amount	1.4	2.7	− 3.9	2.5	2.5	2.8
7.loan_amount	− 1.6	3.2	− 2.3	3	3.9	3.3
8.loan_amount	4.2	3.7	− 1	3.4	− 3.1	3.6
9.loan_amount	0.5	3	2.8	3	− 3.3	3
rate_spread	1.3	1.2	− 4.3	1.2***	3	1.2*
ltv	− 0.08	0.05	0.2	0.05**	− 0.07	0.05
dtirecode	− 0.09	0.07	0.05	0.06	0.04	0.07
1.missdti	− 16.4	6.1**	3.9	8.7	12.4	9.1
1.firsttime	− 2.4	2	4.8	2*	− 2.4	2.1
4.firsttime	− 1.9	2.7	− 1.3	2.6	3.2	2.8
1.broker	− 1.9	1.4	0.9	1.4	1	1.5
1.joint	− 0.5	1.5	0.9	1.4	− 0.4	1.5
1.x83	3.9	3.7	− 0.2	3.4	− 3.7	3.4
2.x83	0.3	2.9	− 0.4	2.7	0.2	2.8
3.x83	− 2.5	2.2	0.1	2.2	2.3	2.3
5.x83	− 0.9	2.1	0.9	2.1	− 0.03	2.1
6.x83	− 1.3	2.7	− 1.8	2.8	3.1	2.9
2.x76rrecode	2.1	2.6	− 2.7	2.4	0.6	2.5
3.x76rrecode	2.4	3.3	− 0.6	3.1	− 1.9	3.2
4.x76rrecode	2	2	− 1.1	2	− 1	2
6.x76rrecode	0.4	1.7	0.2	1.7	− 0.6	1.7
1.rwhitenh	2.8	1.7	− 6.9	1.7***	4.1	1.7*
2.loan_type	1.4	2.2	0.2	2.2	− 1.6	2.2
3.loan_type	− 3.9	2.7	− 3.9	2.6	7.8	3**
4.loan_type	0.9	4.4	0.1	4.2	− 1	4.3
2.x05a	0.7	1.9	1.4	1.9	− 2.1	1.9
3.x05a	1.4	3.6	− 0.8	3.5	− 0.6	3.5

Exhibit A.3

Marginal Effects (ME) and Standard Errors (SE) from Multinomial Logit Model of Responses to Recall of Booklet or Toolkit (Percentage Points)

Variable	Don't Know		Yes		No	
	ME	SE	ME	SE	ME	SE
2.x05b	2.9	2	– 6.6	1.9***	3.7	1.9
3.x05b	0.4	3	– 7.9	3**	7.5	3.2*
2.x05c	– 1.7	2	– 1.7	1.9	3.4	2
3.x05c	– 2.3	3.1	– 1.5	3.1	3.8	3.2
2.x05d	1.7	2.2	– 1.2	2.1	– 0.5	2.2
3.x05d	– 7.6	3.4*	5.5	4	2.1	3.9
2.x05e	3.3	2.1	– 1	2.1	– 2.3	2.2
3.x05e	6.3	3.8	– 6.7	3.6	0.4	3.8
2.x05f	– 0.9	2	0.5	2	0.4	2
3.x05f	3.9	4.8	1.9	4.7	– 5.8	4
2.x05g	– 2.7	2	0.5	1.9	2.2	1.9
3.x05g	– 1	3	– 0.6	3	1.6	3
2.x06	6.9	2.1***	– 3.2	2.1	– 3.6	2.1
3.x06	4.5	2.2*	– 4.6	2.2*	0.07	2.3
2.x07	– 1	1.6	– 1.2	1.6	2.2	1.6
3.x07	5.3	3.1	– 7.9	2.8**	2.6	3
1.rworksta~s	0.7	2.4	– 1.8	2.4	1.1	2.5
3.rworksta~s	– 2.6	2.7	– 1.1	2.5	3.6	2.8
4.rworksta~s	1.6	3.2	– 2.3	3.1	0.7	3.3
1.x81a	0.05	1.5	– 3.3	1.5*	3.3	1.6*
1.x81b	– 2.5	2.6	2	2.6	0.5	2.7
1.x81c	– 3.8	3	1.5	3	2.3	3.1
1.x81def	– 2	2.4	– 3.3	2.4	5.3	2.6*

** p<0.05, ** p<0.01, *** p<0.001*

Acknowledgments

The authors would like to thank Mary Bilinski, Julie Vore, Robert B. Avery, Ron Borzekowski, Misha Davies, and Leslie Parrish for their comments and assistance.

Authors

Brian Bucks is a supervisory economist with the Consumer Financial Protection Bureau.

Tim Critchfield is an interdisciplinary statistician in the Office of Research at the Consumer Financial Protection Bureau.

Susan Singer is the deputy assistant director in the Office of Research with the Consumer Financial Protection Bureau.

References

Agarwal, Sumit, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, and Douglas D. Evanoff. 2010. "Learning to Cope: Voluntary Financial Education and Loan Performance During a Housing Crisis," *American Economic Review* 100 (2): 495–500.

Alexandrov, Alexei, and Sergei Koulayev. 2017. No Shopping in the U.S. Mortgage Market: Direct and Strategic Effects of Providing Information. CFPB Office of Research Working paper No. 2017-01. Washington, DC: Consumer Financial Protection Bureau. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2948491.

Argento, Robert B., Lariece M. Brown, Sergei Koulayev, Grace Li, Marina Myhre, Forrest Pafenberg, and Saty Patrabansh. 2019. "First-Time Homebuyer Counseling and the Mortgage Selection Experience in the United States: Evidence from the National Survey of Mortgage Originations," *Cityscape* 21 (2): 51–74.

Bucks, Brian, and Karen Pence. 2008. "Do Borrowers Know Their Mortgage Terms?" *Journal of Urban Economics* 64 (2): 218–233.

Collins, J. Michael, and Collin M. O'Rourke. 2010. "Financial Education and Counseling—Still Holding Promise," *Journal of Consumer Affairs* 44 (3): 483–498.

Consumer Financial Protection Bureau (CFPB). n.d.1. *Your Home Loan Toolkit: A Step-by-Step Guide*. Washington, DC: Consumer Financial Protection Bureau. https://files.consumerfinance.gov/f/201503_cfpb_your-home-loan-toolkit-web.pdf.

———. 2018. n.d.2. *What Do Prospective Homebuyers Know About Mortgages and Mortgage Shopping? Know Before You Owe: Mortgage Shopping Study—Brief #3*. Washington, DC: Consumer Financial Protection Bureau. https://files.consumerfinance.gov/f/documents/bcfp_mortgages_shopping-study_brief-3-knowledge.pdf.

Kleimann Communication Group, Inc. 2013. *Know Before You Owe: Quantitative Study of the Current and Integrated TILA-RESPA Disclosures*. Presented to the Consumer Financial Protection Bureau, November 20. Washington, DC: Kleimann.

———. 2012. *Know Before You Owe: Evolution of the Integrated* TILA-RESPA Disclosures. Presented to the Consumer Financial Protection Bureau, July 9. Washington, DC: Kleimann.

Lacko, James, M., and Janis K. Pappalardo. 2010. "The Failure and Promise of Mandated Consumer Mortgage Disclosures: Evidence from Qualitative Interviews and a Controlled Experiment with Mortgage Borrowers," *American Economic Review* 100 (2): 516–521.

Moulton, Stephanie. 2012. "Pre-purchase Homebuyer Education and Counseling: Diverse Strategies for Diverse Homebuyers." In *Consumer Knowledge and Financial Decisions*, edited by Douglas J. Lamdin. New York: Springer, 145–164.

Myhre, Marina, L., and Nicole Elsasser Watson. 2017. Housing Counseling Works. Washington, DC: U.S. Department of Housing and Urban Development, Office of Policy Development and Research. https://www.huduser.gov/portal/sites/default/files/pdf/Housing-Counseling-Works.pdf.

U.S. Department of Housing and Urban Development (HUD). n.d. *Shopping for Your Home Loan. HUD's Settlement Cost Booklet*. Washington, DC: HUD. https://www.hud.gov/sites/documents/HUD_SETTLE_COST.PDF.

Woodward, Susan, E., and Robert E. Hall. 2010. "Consumer Confusion in the Mortgage Market: Evidence of Less than a Perfectly Transparent and Competitive Market," *American Economic Review* 100 (2): 511–515.

Refereed Papers

Refereed papers that appear in Cityscape *have undergone a thorough and timely double-blind review by highly qualified referees. The managing editor reviews submitted manuscripts or outlines of proposed papers to determine their suitability for inclusion in this section. To submit a manuscript or outline, send an e-mail to cityscape@hud.gov.*

Road Map to a Unified Measure of Housing Insecurity

Robynn Cox
Benjamin Henwood
Seva Rodnyansky
Eric Rice
Suzanne Wenzel
University of Southern California

Abstract

This article discusses the evolution of the concept and measurement of housing insecurity. Our survey of the literature uncovers that multiple terms and definitions are used to describe housing insecurity. Based on our analysis, we argue for one operational term, ***housing insecurity****, and we put forth an operational definition that captures the various dimensions of this issue. We also argue for the development of an instrument that allows researchers to accurately measure the problem. We provide a road map for how this can be achieved based on the precedent set forth by the U.S. Food Security Survey Module.*

Introduction

Housing insecurity is often a reality for individuals experiencing poverty and/or social marginalization. Although it is multidimensional by nature, typically only a few of these dimensions are reliably measured and included in national estimates. Based on a thorough cross-field literature review of 106 quantitative, qualitative, and mixed-methods studies,[1] we argue in this article that current measures of housing insecurity are incomplete, which may result in inadequate policy responses and funding.

The importance of housing goes beyond the material infrastructure that serves as protection from the elements (Shaw, 2004); housing is interrelated with physical, social, and psychological well-being (Padgett, 2007). Research has primarily focused on either the health and education threats associated with substandard housing and neighborhoods (Bashir, 2002; Katz, Kling, and Liebman, 2001; Leventhal and Brooks-Gunn, 2003; Ludwig et al., 2013; Marsh et al., 2000; Sanbonmatsu et al., 2006) or the psychosocial benefits of housing as a home (Dupuis and Thorns, 1998; Low and Lawrence-Zuniga, 2003; Padgett, 2007; Shaw, 2004; Somerville, 1992). Also a focus of research, lack of housing altogether—or homelessness—brings into stark relief the fundamental importance of housing security as a prerequisite for health, employment, and various other aspects of daily functioning (Henwood et al., 2013).

In the United States, over half a million people experience homelessness on any given day (Henry et al., 2018). Efforts to address this extreme form of housing insecurity receive dedicated infrastructure and oversight. Congress provides direct funds to address homelessness through the McKinney-Vento Homeless Assistance Act, and communities across the United States are required to maintain homeless management information systems to receive federal funding to address homelessness. Each year, communities across the country also conduct a homeless count to monitor the scope of the problem. An Annual Homeless Assessment Report (AHAR) is submitted to Congress each year that includes monitoring from both the Homeless Management Information System (HMIS) and homeless counts. This type of monitoring has helped direct resources and has enabled strategies that have reduced the overall number of chronically homeless adults and homeless veterans since the inception of AHAR in 2007. Although some interventions to address homelessness promote housing stability and security (Padgett, Henwood, and Tsemberis, 2015),[2] the definition of what constitutes homelessness has been a moving target[3] and is not considered to be part of a unified construct of housing insecurity that can be measured on a continuum.

As compared with the extreme of homelessness or even housing affordability, there is much less research exploring other facets of housing insecurity. This could partially be because individuals

[1] See appendix B for a list of articles included in our literature review.
[2] Housing First is an example of an effective intervention for homelessness, with research consistently demonstrating its impact on housing stability. Despite, or perhaps because of, overwhelmingly positive housing outcomes, there has been limited discussion of differences in the operationalization of housing stability and retention in studies of Housing First, which reflects inconsistencies in the definition of these concepts in the broader body of research on homelessness and housing (Byrne, Henwood, and Scriber, 2018).
[3] Beginning in 2009, federal definitions of homelessness were expanded to permit inclusion of persons at "imminent risk of homelessness" to expand access to homelessness assistance benefits. Such persons include individuals or families whose residence will be lost within 14 days of application for homelessness assistance, for whom no subsequent residence has been identified, and for whom resources are lacking to obtain other permanent housing. Another group for which homeless assistance has been expanded is those individuals and families fleeing or attempting to flee domestic violence (HUD, 2012).

and families experiencing housing insecurity are less visible to the public. The limited focus on the more hidden aspects of housing insecurity, such as housing stability, housing quality, and behavioral responses to housing affordability (for example, trading housing safety, housing quality, neighborhood safety, and/or neighborhood quality for affordability), and the failure to take into consideration all the domains of housing insecurity (that is, housing affordability, housing stability, housing safety, housing quality, neighborhood safety, neighborhood quality, and homelessness) have resulted in much less being known about its true prevalence and the actual costs it imposes on society. In addition, in domains with more research focus, such as neighborhood effects, the link has not been adequately made to the issue of housing insecurity. Developing a common language and uniform measurement tool would help society mobilize resources, improve its understanding of the importance of this invisible problem, and generate solutions.

In this article, we propose a uniform instrument to assess housing insecurity that can capture the multidimensional aspects of housing, such as access and quality. A comprehensive measure would be able to encapsulate both access to and quality of housing and the behavioral tradeoffs made to secure housing. In addition to examining how housing insecurity has been conceptualized across the literature, we propose a road map for the creation of a new measure of housing insecurity, based on the development of the U.S. Food Security Survey Module that has been incorporated into the Current Population Survey (CPS) annually since 1995 (National Research Council, 2006).

A Brief History of Housing Insecurity

The concept of adequate housing as a human right was adopted by the United Nations (U.N.) General Assembly in Article 25 of the 1948 Universal Declaration of Human Rights and Article 11 of the International Covenant on Economic, Social and Cultural Rights in 1966 (U.N., 2014). The U.N. Committee on Economic, Social and Cultural Rights did not define "adequate housing," however, until the Committee's general comments in 1991 and 1997, at which point the U.N. characterized adequate housing as meeting the following minimum criteria:

(a) tenure security that guarantees legal protection against forced evictions,[4] harassment, and other threats;
(b) availability of materials and infrastructure, such as safe drinking water; adequate sanitation; energy for cooking, heating, and lighting; food storage; and refuse disposal;
(c) affordability such that paying for housing does not compromise other human rights;
(d) habitability that includes protection against the cold, damp, heat, rain, wind, other threats to health, and structural hazards;
(e) location that is not polluted or dangerous and that does not cut off access to employment opportunities, healthcare services, schools, or other critical social institutions; and
(f) accessibility that can meet the specific needs of disadvantaged and marginalized groups and does not compromise the expression of cultural identity (U.N., 2014).

[4] Note that forced evictions do not include evictions that are justifiable (for example, repeatedly not paying rent or damaging property without justifiable cause) and carried out in accordance with international human rights law (U.N., 2014).

The U.N. provides a framework to monitor human rights, including the right to housing, with several suggestions for housing indicators; those include the share of public expenditure on subsidized or public housing, reported cases of forced evictions, and rates of homelessness. Nonetheless, they do not operationalize a uniform measure of adequate housing or housing security (U.N., 2014).

Likewise, in 1949, the U.S. Congress passed the Housing Act of 1949. This controversial law[5]

> …declare[d] that the general welfare and security of the Nation and the health and living standards of its people require housing production and related community development sufficient to remedy the serious housing shortage, the elimination of substandard and other inadequate housing through the clearance of slums and blighted areas, and the realization as soon as feasible of the goal of a decent home and suitable living environment for every American family, thus contributing to the development and redevelopment of the communities and to the advancement of the growth, wealth, and security of the Nation.[6]

The law explicitly recognizes the importance of affordable adequate housing and a "suitable living environment," or neighborhood, not only for the health and well-being of the people (more than 40 years before the 1991 and 1997 U.N. criteria) but also for macroeconomic growth and stability.

In 1969, the U.S. Department of Health and Human Services (HHS) went a step further in explicitly defining housing instability using five indicators:

(1) exorbitant housing costs relative to income (greater than 50 percent);
(2) inferior housing quality (for example, inadequate plumbing, heat, or electricity; leaks; holes; and so on);
(3) neighborhood instability (for example, high rates of poverty, crime, and unemployment; poor city services; litter; noise; pollution; and so on);
(4) overcrowding; and,
(5) at the extreme, the condition of homelessness (HHS, 1969).

The U.N. Housing Act of 1949 and HHS definitions take a broad view of what constitutes housing security. The Housing Act of 1949 establishes the importance of housing to the welfare and security of the people and the macroeconomic growth and stability of the nation, the U.N describes the minimum conditions necessary for housing security, and HHS provides a foundation for a comprehensive housing insecurity definition. Although the definitions are, in general, mirror images of one another, the HHS definition is more narrowly defined and leaves out certain aspects of the U.N. definition, such as expression of cultural identity, forced evictions, and accessibility to services required by disadvantage and marginalized groups. Nonetheless, the HHS definition incorporates the dimension of housing affordability and is defined in a way that more readily permits the operationalization of a housing insecurity measure: All five of the categories of housing insecurity in this definition are quantifiable. The HHS characterization of housing insecurity also illustrates that, while homelessness is a sufficient condition for housing insecurity, it is not a necessary one.

[5] Critics of this law argue that it led to increases in housing insecurity by cultivating the slums and ghettos it sought to eliminate (von Hoffman, 2000).
[6] U.S. Congress. 1949. Housing Act of 1949. Pub. L. 81-171. 63 Stat. 413-444. http://legisworks.org/congress/81/publaw-171.pdf.

These definitions provide a good starting point to create a fully operational definition of housing insecurity. In the next section, we conduct an exhaustive cross-field literature review, which will form the basis for the fully operational definition of housing insecurity that we set forth in the section after that.

Current Approaches to Housing Insecurity

Despite the fact that the HHS set forth a definition of housing insecurity as early as 1969, and both the U.N. and HHS depictions of this problem correctly outlined the concept, to date, working definitions have not captured the multidimensional aspects of this issue. Most of the research has focused on only three dimensions (homelessness, stability, and affordability), giving a potentially incomplete view of the extent of housing insecurity. To better understand how housing insecurity has been defined both conceptually and operationally, we conducted a thorough cross-field literature review to determine how housing insecurity is defined, measured, and used in research.

We sought out academic studies, government reports, and technical reports released from 1991 to 2017 on the topic of housing insecurity and its allied definitions. We operationalized this review by gathering papers on the Google Scholar and Web of Science search engines for the following terms: "housing security," "housing insecurity," "housing affordability," "housing stability," "housing instability," "homelessness," "housing hardship," "housing quality," "housing safety," and "neighborhood safety." Given the focus of this paper on the global concept of housing insecurity, we restricted our search to the macrolevel terms listed above. Relevant papers (see appendix A, exhibit A-1) either (1) included housing security as an explanatory variable measuring outcomes on households, individuals, or populations; (2) sought to understand reasons for housing security as a dependent variable; or (3) assessed the prevalence of housing insecurity. Our search focused on the U.S. context; however, several articles from other developed countries were relevant to the development of a measure of housing insecurity. Our literature review yielded 106 quantitative, qualitative, and mixed methods studies and reports from a variety of fields and outlets (see appendix B for a complete list of articles reviewed); 55 of those studies used some form of housing insecurity as an explanatory variable measuring a different individual-level or household-level outcome, 44 studies used housing insecurity as a dependent variable, seeking to understand its causes, and 17 studies measured the prevalence of some form of housing insecurity. Of these studies, 7 were counted as both explanatory and dependent variable studies, 2 were counted as both prevalence and dependent variable, and 2 were counted as both prevalence and explanatory variable. Eleven studies were from an international context (see appendix table 1 of Cox et al., 2017a for a full list of all papers, with short descriptions).

Based on the 106 quantitative, qualitative, and mixed methods studies and reports reviewed, we find three major concerns with current approaches to measuring housing insecurity: (1) lack of a uniform definition, (2) underdeveloped concept, and (3) inconsistent measurement. Specifically, we find that the definition of what constitutes housing insecurity varies widely. In our scan of the literature, certain search terms—housing instability (43 percent), homelessness (34 percent), and housing insecurity (16 percent)—had many more associated papers on this topic than did other key words[7] (see appendix A, exhibit A-2). Moreover, most papers did not refer to "housing (in)security" explicitly;

[7] Certain papers came up under multiple search terms.

rather, they examined a particular subcategory, such as housing instability or affordability. Definitions of what constitutes housing insecurity also varied widely. For example, Heflin, London, and Scott (2011) assessed how women in three disadvantaged Cleveland neighborhoods coped with housing hardships, with the dependent variable—housing hardships—including housing stability, housing quality, housing safety, and neighborhood safety. In contrast, Rourke et al. (2012) used a whole battery of housing-insecurity categories (stability, affordability, quality, neighborhood quality, and homelessness) in assessing physical and mental health outcomes in adults with HIV.

Even when definitions varied in consistency, measurements of the same condition often differed in the literature (see appendix A, exhibit A-3). For example, 30 of the papers considered housing (un)affordability as housing insecurity, but affordability was assessed using five different measurements: (1) "difficulty/inability to make payments on housing" (83 percent), (2) "housing cost burden" (37 percent), (3) "foreclosure" (10 percent), (4) "legal housing issues" (7 percent), and (5) "having rental assistance" (3 percent), with some 30 percent of papers testing for multiple measurements but no studies testing for all five measures. There is variability even within measures. Some researchers use a 30-percent housing burden cutoff (for example, Mimura, 2008), while others use 50 percent (see Capps, 2001), and some use both (see Coulton et al., 2001). Assessments of difficulty in making payments vary as well: most are solicited via survey responses, and questions are not standardized. The measurement of housing instability varies even more than affordability. Eight different measures were employed in the 49 papers that used housing instability as a measure of housing insecurity: (1) "multiple moves" (59 percent), (2) "doubling up" (35 percent), (3) "eviction" (18 percent), (4) "overcrowding" (18 percent), (5) "duration of stay" (14 percent), (6) "forced moves" (8 percent), (7) "living in unstable conditions" (2 percent), and (8) "living in multiple subsidized units over time" (2 percent). As with affordability, 39 percent of the papers used multiple definitions, although no study used all eight measures. There was also considerable within-measure variability in instability, with overlapping definitions of overcrowding and doubling up, and overlapping definitions of multiple moves and duration of stay. For example, studies differed on how many moves were multiple: in two studies of mothers and children, using the same Fragile Families and Child Wellbeing Study, Ziol-Guest and McKenna (2014) defined it as three moves over a 5-year timeframe, while Suglia, Duarte, and Sandel (2011) defined it as two or more moves in 2 years.

Most of the papers reviewed define housing insecurity incompletely or narrowly. While some papers we surveyed do focus only on a specific housing issue (such as homelessness), its determinants, or its outcomes, many others claim to speak for housing insecurity in general. Yet, it is rarely the case that all available dimensions of housing insecurity are combined into one indicator, even when more than one measure is available in a survey. Of the 106 surveyed papers, none addresses all seven of our domains of housing insecurity, and only 8 studies address four or more facets. Of the 73 papers with a focus on homelessness, 35 percent focused only on this one subcategory of insecurity, and 38 percent focused on just two.

Data availability and specificity have also created challenges in studying the full continuum of housing insecurity. In the reviewed literature, 98 of the 106 studies analyzed used data-driven approaches to ascertain some dimension of housing insecurity. Twenty-nine studies collected primary data; the rest relied on 38 different secondary sources. Among these secondary data sources, the U.S. Center for Disease Control's Behavioral Risk Factor Surveillance System survey;

the Children's Health Watch survey; HUD's National Survey of Homeless Assistance Providers and Clients; the Worcester, MA Family Research Program; and Australia's Journeys Home data sets were used in three papers each; the Fragile Families and Child Wellbeing Study, the National Survey of America's Families, and the National Longitudinal Survey of Youth data sets were used in two papers each. Such data set variation contributes to the different permutations in housing insecurity measurement.

As a result of varying definitions, comparing housing insecurity across studies is difficult. Moreover, many scholars define the measurement of their version of housing insecurity to a specific subpopulation of interest, making comparisons impossible. Such a narrow definition of housing may also lead to undercounting and partial measurement. Take, for example, the affordability dimension of housing insecurity. If we were only to define housing insecurity by affordability, a household could be considered housing secure if their housing cost burden is below a certain threshold (or if they do not report difficulty with making payments), even if they made tradeoffs between affordability and housing quality such that they chose to live in a low-quality dwelling that poses threats to the household members' health (containing lead paint, mold on the walls, and so on) or safety (for example, active electrical wires sticking out, severely dilapidated stairs, living in a dangerous neighborhood) in order to obtain more affordable housing.

Toward a New Approach to Measuring Housing Insecurity

A comprehensive definition of housing insecurity, as set forth by HHS and the U.N. General Committee, has yet to be put into use.[8] Housing insecurity impacts many well-being outcomes for adults and children, yet many aspects of housing insecurity are understudied, at least compared to homelessness and housing affordability (possibly because they are less visible to the public or more difficult to measure), and their prevalence and cost to society are not fully understood. Developing a uniform measurement tool would help society to mobilize resources, improve its understanding of the importance of this problem, and generate solutions.

One possible definition (based on the overlapping descriptions set forth by HHS and the United Nations) for housing security is as follows:

Availability of and access to stable, safe, adequate, and affordable housing and neighborhoods regardless of gender, race, ethnicity, or sexual orientation.

Likewise, housing insecurity would be defined as follows:

Limited or uncertain availability of stable, safe, adequate, and affordable housing and neighborhoods; limited or uncertain access to stable, safe, adequate, and affordable housing and neighborhoods; or the inability to acquire stable, safe, adequate, and affordable housing and neighborhoods in socially acceptable ways.

[8] Note that the Office of Policy Development and Research within the U.S. Department of Housing and Urban Development (HUD), inspired by an earlier version of this paper first released as a working paper in 2016 (see Cox et al., 2017a), has developed and recently published a 60-day notice to pilot a housing insecurity module within the 2019 American Housing Survey. Published in the *Federal Register* as a notice on September 11, 2018. 83 Fed. Reg. 45,955-45,956. https://www.govinfo.gov/content/pkg/FR-2018-09-11/pdf/2018-19707.pdf.

From this definition, it becomes clear that housing insecurity can be characterized by the following dimensions:

(1) Housing Stability
(2) Housing Affordability
(3) Housing Quality
(4) Housing Safety
(5) Neighborhood Safety
(6) Neighborhood Quality
(7) Homelessness

Where housing stability is the ability of a household to stay in a housing unit of its choosing, for a duration of its choosing, without interruption or complication. Common manifestations of instability include living in overcrowded conditions, doubling up with relatives or friends, frequent moves, forced moves such as eviction, and others. Housing affordability is a household being reasonably able to pay for adequate housing, on time, within its budget; unaffordability includes having a high housing cost-to-income ratio (burden), the difficulty or inability to pay for the full costs of housing on time or fully, and legal issues related to paying for housing. Housing costs include not only the rent or mortgage but also relevant property taxes and utilities and routine maintenance bills.

Housing quality and safety deal with structural characteristics of a household's housing unit. Unsafe housing poses direct health risks to inhabitants. Manifestations include open wires; lack of insulation or heating; holes in exteriors, roofs, or floors; lack of water access; and pests in the unit. Low-quality housing, while not directly harming the household, nevertheless shifts the members' quality of life below an expected level; common manifestations include dilapidated exterior of building, peeling paint in the interior of the unit, lack of access to (working) appliances, not having bathroom fixtures and/or flushing toilets, and having utilities that experience frequent breaking or stopping, including sewage, water, electricity, heat, and so on. In addition to the health concerns, it is important to include low-quality and unsafe housing in housing insecurity because of the tradeoffs individuals and households may choose to make between housing affordability, housing quality, and housing safety.

Neighborhood-level characteristics such as safety and quality are not often thought of as directly housing-related; however, they are certainly tied to households' ability to live in a safe environment in particular and to its well-being in general. Therefore, we believe, neighborhood characteristics form part of a measure of housing insecurity. This is especially true if households are making tradeoffs in neighborhood quality or safety for greater housing affordability. For example, if a household can only afford to live in a high-crime neighborhood, there could be at least two possible policy interventions: give the household a subsidy to move elsewhere (an issue of housing affordability) or make the current neighborhood safer (an issue of neighborhood safety). Lack of neighborhood safety entails the presence of high crime, many abandoned buildings, the proximity of environmental hazards, and excessive noise and traffic, among others. Low-quality neighborhoods include those with poor services, poor infrastructure, low access to amenities, and others. Both neighborhood quality and safety can be measured subjectively by self-reports or objectively, perhaps, through administrative data. For example, a neighborhood audit tool could be used as a more "objective" measure of neighborhood quality, such as The Revised Residential Environment Assessment Tool (REAT 2.0), which includes measures of neighborhood condition, natural surveillance, and natural elements (Poortinga et al., 2017). Note that these tools may still suffer from observer bias. In terms

of neighborhood safety, these measures could be obtained by working with other government agencies, such as the Federal Bureau of Investigation and the Environmental Protection Agency, that collect more objective administrative data on neighborhood safety. These measures might include crime statistics from the FBI's Uniform Crime Reports,[9] neighborhood-level industrial hazard data from EPA's Toxic Release Inventory (for example, see Crowder and Downey, 2010), and EPA's Safe Drinking Water Information System Federal Reporting Services. One interesting implication of the inclusion of neighborhood safety into a housing insecurity measure is that we could consider a scale that might place more weight on variables that pose an immediate danger to the safety of the inhabitant. This could potentially lead to housing in areas with extreme violence being marked as uninhabitable locations that need immediate assistance. The authors are not aware of a standard of neighborhood safety that would lead to this conclusion, but they note that epidemiological models of contagion might shed some light on such a threshold. These models could help to determine when there might be an epidemic (in certain neighborhoods and communities), warranting the designation of dwellings in those areas as uninhabitable.

Finally, homelessness can have multiple forms (HUD, 2012), such as literal homelessness (which includes sleeping on the street), temporary housing (such as homeless shelters), and living in dwellings not meant for human habitation (living in a vehicle, railroad car, abandoned building, encampment, RV, and so on); while examples of more hidden forms of homelessness might include couch surfing and doubling up.

This definition makes room for both a categorical and a continuous measure (both of which have value) of housing insecurity. Such a measure would be a function of the number and depth of housing issues along the aforementioned seven dimensions. While we have included homelessness as part of our housing insecurity definition, as in the case of the condition of hunger in the food security module, it is conceivable that homelessness could be removed from the definition. As previously mentioned, homelessness is a sufficient—but not necessary—condition for housing insecurity; therefore, it could be modeled as the severest form of housing insecurity along a continuum, or possibly as an outcome of housing insecurity. One argument for including homelessness within the definition of housing insecurity is, unlike hunger, homelessness is experienced at the household level; therefore, we have opted to keep it in our definition. What will prove to be more difficult when measuring homelessness in an index of housing insecurity is the ability to capture homeless households that are doubling up with households that are not homeless.[10]

Developing a uniform definition does not mean that research or policy should not separately focus on the different dimensions of housing insecurity when appropriate; rather, it suggests that to precisely estimate the prevalence of this issue and accurately account for its social costs, we need a comprehensive definition so that we can develop an instrument that will jointly measure all of its dimensions. Specifically, a common definition of housing insecurity would facilitate the development of a validated national instrument of housing insecurity that can be assessed at the household level and, possibly, the individual level. To capture the multidimensionality of housing insecurity, this measure should be defined as a scale that can quantify housing security along a continuum, from the most housing secure to the most extreme cases of housing insecurity, as well as categorically.

---

[9] Another source of crime data is the Centers for Disease Control and Prevention's (CDC) National Vital Statistics System (NVSS), which has information on deaths by homicide, as well as the National Crime Victimization Survey (NCVS), which, although less objective, provides self-reports of crime victimization.

[10] Specifically, some households may choose to double up, while others may do so because of the loss of housing.

We provide several guidelines to aid in the development of such an instrument, to ensure that this measure is useful in a variety of settings and contexts. A uniform multidimensional construct captures the presence of an underlying constellation of issues rather than just one problem, an approach that may be appropriate for many studies, but not for all. Measurement frequency should depend on the needs of the particular study, so survey question wording should be easily adaptable to different time lengths but will also have to take into consideration the burden of the survey on the respondent. Housing insecurity prevalence measures at national, state, regional, or local levels should be at frequent enough intervals (at least annually or biennially) to measure the evolution of housing insecurity. For example, similar to the U.S. Household Food Security Module, a housing security module could survey respondents' about their housing security over the past 12 months; and because this survey is collected biennially, to understand the housing characteristics of the U.S. population, population estimates of housing insecurity would be obtained every 2 years. In addition, the timeline for the baseline module questionnaire should cover the lifetime history of housing security of the household for the inaugural year in order to understand the past experience of housing insecurity in American households.

Households should be the units of analysis for housing insecurity, unless study designs necessitate otherwise. While housing insecurity takes neighborhood-level characteristics into account, relevant outcomes are at the household level, recognizing that different processes may affect household-level and neighborhood-level insecurity. Measurement modules should be flexible to include local context. For example, housing cost burden cutoffs may differ by metropolitan area, or regions in warm climates do not need questions about the existence and performance of home heating equipment. Similarly, rural contexts, Native American reservations, and other unique areas may require slight modifications. While this article approaches housing insecurity from a U.S. perspective, the measure should be easily transferrable to other developed countries and, with some modification, to the developing world context.

Current Measures of Housing Insecurity

To date, no research has captured all the proposed seven domains of housing insecurity among a nationally representative population. Tsui et al. (2011) have the most expansive measure of insecurity, including unaffordable or unsafe housing, unsafe neighborhoods, homelessness, experiencing foreclosure, or having been in housing court. In contrast, Curtis and Geller (2010), Geller and Franklin (2014),[11] Goldrick-Rab et al. (2015), and Warren and Font (2015) view insecurity as a combination of housing instability, unaffordability, and homelessness. Routhier (2018) restricts her sample to renters and measures housing insecurity only among the dimensions of housing affordability, housing stability, and poor physical unit conditions. Bailey et al. (2016) see insecurity as a housing stability and affordability issue. Buffardi et al. (2008) and Diette and Ribar (2018) see it as a combination of housing instability and homelessness. Others define insecurity as a mix of homelessness and poor housing quality (Western, Braga, and Kohl, 2015), and yet others think of it as a mix of homelessness and unaffordability (Surratt et al., 2015). Campbell et al. (2014), Greder et al. (2008), Liu et al. (2014), and Stahre et al. (2015) equate housing insecurity with unaffordability only; Cutts et al. (2011) and Frank et al. (2010) associate it with only housing instability; and Thurston et al. (2013) view it solely as unsafe housing.

[11] Geller and Franklin (2014) also mentioned eviction as a component of housing insecurity.

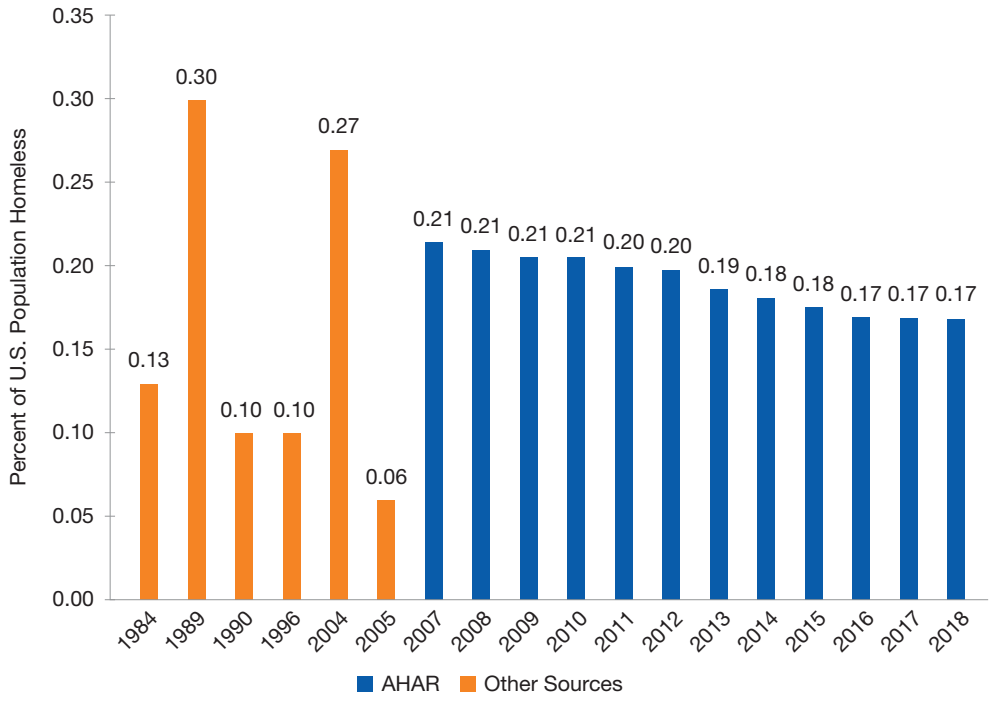
Although a comprehensive housing insecurity measure including all seven insecurity dimensions is heretofore unavailable, some studies do document the prevalence of specific dimensions at the national level for the United States. We next present the two dimensions that have national prevalence statistics (homelessness and housing affordability), followed by a view of currently existing multidimensional measures.

Homelessness

Of our seven housing insecurity dimensions, homelessness appears to have the most standardized measurement in the United States, even though there is still disagreement on what constitutes homelessness. Homelessness is typically categorized in one of three ways: being homeless at the time of the survey, having been homeless at least one night throughout the survey year, or having ever been homeless. Since 2007, the U.S. Department of Housing and Urban Development (HUD) has conducted the Annual Homeless Assessment Report (AHAR) for Congress, which is an annual point-in-time (PIT) homeless count, measuring the number of literal homeless at the time of survey, throughout the country (Solari et al., 2016). The AHAR also documents homeless counts as performed by homeless services agencies, shelters, and transition centers (Solari et al., 2016). Exhibit 1 shows national estimates from AHAR and other sources prior to AHAR's inception in 2007.

Exhibit 1

U.S. Homelessness Rate, Point-in-Time Count



AHAR = Annual Homeless Assessment Report.
Sources: U.S. HUD AHAR 2017 Report, Exhibit 1.1; Honig and Filer (1993); Drake et al. (1989); Lee et al. (2010); National Law Center on Homelessness and Poverty (2004); Pearson, Montgomery, and Locke (2009)

The AHAR process standardized measurements and enabled adequate comparisons. This figure is a great illustration of how the estimation of prevalence rates can improve once definitions are standardized. Prior to AHAR's introduction, prevalence rates ranged from a high of roughly 0.3 percent to a low of about 0.06 percent. After the introduction of AHAR, however, prevalence rates stabilized, ranging from roughly .22 percent to .2 percent from 2007 to 2012, with a continual drop in the prevalence rates from 2013 to 2016 to about .17 percent. Prior to AHAR, it is likely that differences in national estimates of homelessness were influenced by definitional differences versus actual differences in the prevalence of homelessness. Also, while some may argue that AHAR only captures one type of homelessness, literal homelessness, it is clear from Exhibit 1 that AHAR has helped to improve the estimation of homelessness over time.

The prevalence of homelessness can also vary by the time period over which it is defined and the unit of measurement (that is, individual or household). For example, the point-in-time prevalence rate will be vastly different from prevalence rates measuring homelessness over 12 months or one's lifetime. Specifically, prevalence estimates of individuals homeless at least once during a given year were 1.08 percent, using the 1999 National Survey of Homeless Assistance Providers and Clients (NSHAPC), and 0.85–1.19 percent in 2004, using estimates from the Urban Institute and the Harvard Joint Center for Housing Studies (Kushel et al., 2006; National Law Center for Homelessness and Policy, 2004). These studies occurred before AHAR's inception, and their definitions of homelessness may differ, possibly explaining the differences. Widening the measurement time period further, Link et al. (1994) estimated the population of the United States that has ever been homeless as approximately 7–14 percent. Thus, the measurement time period is an important consideration in designing standardized measures.

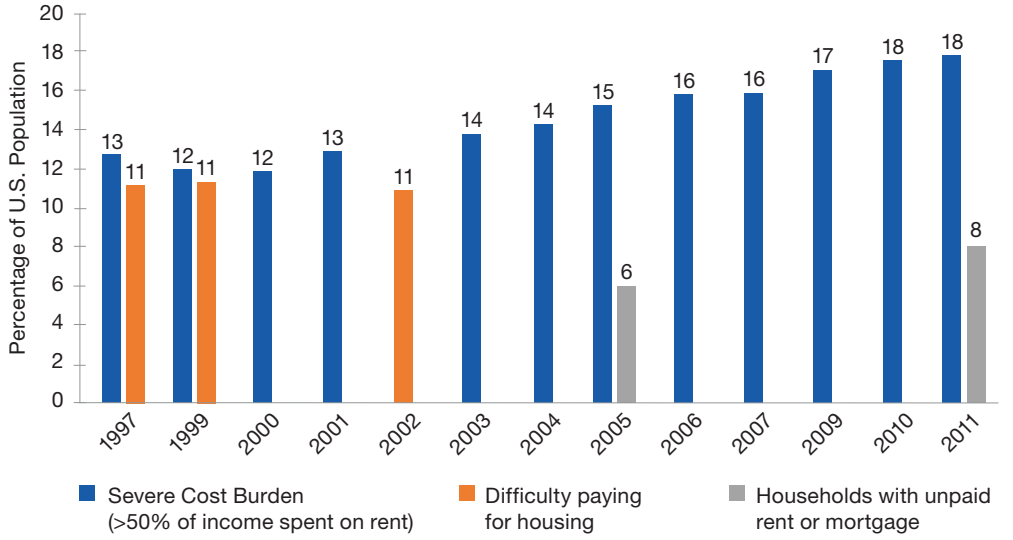
Housing Affordability

Measures, and hence prevalence rates, vary among scholars documenting housing affordability issues. The most commonly used are whether a household demonstrates difficulty paying for housing-related expenses such as rent, mortgage, and utilities; whether rent or mortgage goes unpaid over a given time period; and whether housing cost presents a burden, defined as spending more than a certain percentage (usually between 30 and 50 percent) of household income per month on housing. Other less common measures include experiencing legal issues related to housing finance, foreclosure, or the receipt of housing assistance through a government program. Exhibit 2 shows the national estimates for the three most frequently utilized housing affordability measures over time. Depending on the statistic used, recent estimates vary from 6 to 18 percent of U.S. households that experience difficulties with housing affordability. Each statistic uses a different nationally representative data set. Severe cost burden and difficulty paying for housing seem to trend closely in years when comparable data are available. On the other hand, the statistic for unpaid rent or mortgage, measured using the U.S. Census Bureau's Survey of Income and Program Participation (SIPP), show rates twice as low as severe housing cost burden for the same years (Siebens, 2013).

[12] Nationally reliable statistics on difficulty paying rent come from the National Survey of America's Families, undertaken by the Urban Institute three times in 1997, 1999, and 2002 and not since.

Exhibit 2

Affordability-Related Housing Insecurity Estimates for U.S. Households 1997–2011



Sources: Joint Center for Housing Studies (2002, 2003, 2005, 2006, 2007, 2008, 2011, 2012, 2013); Urban Institute and Child Trends (NSAF) (1997, 1999, 2002); Siebens (SIPP) (2013)

National Multidimensional Measures of Housing Insecurity

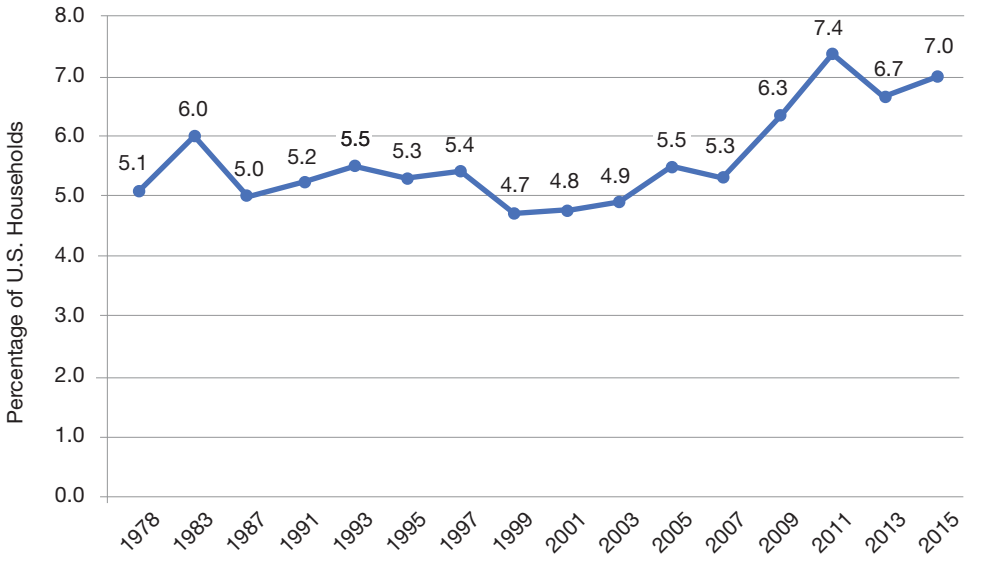
Few studies mention the dimensions of housing safety or quality and neighborhood safety or quality, and national prevalence statistics for these are likewise difficult to find. Leopold et al. (2017) describe the difficulties in obtaining reliable and representative data on overcrowding, housing instability, housing quality, and neighborhood quality. They also point out the dearth of longitudinal data on these topics (Leopold et al., 2017). Prevalence of neighborhood issues is often measured for only one neighborhood, city, or region, and definitions and measurements vary.

One alternative for housing quality and safety is the American Housing Survey (AHS), which provides a variety of statistics that could be used to measure housing quality and safety (see Cox et al., 2017b). Since 1991 HUD has released a biennial report to Congress documenting the "Worst Case Housing Needs" based on the biennial AHS data. Eligible U.S. households for this designation are those who are renters, have incomes below 50 percent of area median income, and do not receive housing assistance. Of these, households are counted as having Worst Case Needs if they have severe rent burdens (paying more than 50 percent of income toward rent) or severely inadequate housing quality in terms of heating, plumbing, electrical systems, or maintenance (HUD, 2017). Worst Case Housing Needs at the national level have hovered between 4.5 and 6 percent for most years, rising to 7.4 percent in 2011 in the aftermath of the Great Recession (exhibit 3). The Worst Case Housing Needs reports provide a start to measuring housing insecurity nationally, including the housing

affordability and quality dimensions. This measure leaves out multiple important dimensions of housing insecurity, however, (such as neighborhood quality and safety, housing safety, instability, and homelessness) and is defined based on very low-income renters.

Exhibit 3

Percentage of U.S. Households with Worst Case Needs: 1978–2015



Notes: 1981, 1985, and 1989 linearly interpolated.
Sources: U.S. HUD Worst Case Needs Report 2017, Exhibit 1-4; U.S. HUD Worst Case Needs Report 2005; U.S. HUD Trends in Worst Case Needs for Housing, 1978–1999

Aside from HUD's "Worst Case Housing Needs" reports, Siebens' U.S. Census Bureau report (2013) takes steps toward a national snapshot of housing insecurity prevalence based on 2011 data. The author shows that 3.4 percent, 2.6 percent, and 6.7 percent of U.S. households had poor housing quality, lived in unsafe housing, and lived in unsafe neighborhoods, respectively. Moreover, the author recommends summarizing a measure of hardship across nine indicators, the first five of which could qualify as housing-related: difficulty meeting essential expenses, not paying rent or mortgage, getting evicted, not paying utilities, having utilities cut off, having phone service cut, not seeing a doctor when needed, not seeing a dentist when needed, or not always having enough food (Siebens, 2013). According to Siebens' measure, in 2011, 78 percent of households faced zero hardships, 9 percent faced one hardship, and 3 percent faced three or more hardships (Siebens, 2013). This measure bears some similarity to our proposed measure of housing insecurity but focuses on overall material well-being instead of developing a comprehensive measure of housing insecurity, and while the report looks at various dimensions of housing insecurity that we have incorporated in our definition, the focus of the report is not solely on the development of a measure of housing insecurity. Therefore, Siebens' scale captures both housing and nonhousing measures of well-being of U.S. households.

A different approach, which looks at the effects of housing and neighborhood quality on mental health, was undertaken by Wright and Kloos (2007). This approach measured three levels of housing and environment variables among residents of supportive housing programs for the mentally ill across 34 housing sites in 10 cities in one U.S. state, for a total sample of 249 (Wright and Kloos, 2007). Self-reported data on apartment quality, neighborhood quality, and neighborhood social climate were obtained using the Housing Environment Survey. Housing data were validated by an observer using the Housing Environment Rating Scale, and neighborhood-level data were supplemented by census tract demographic and socioeconomic data from the 2000 census (Wright and Kloos, 2007). Well-being was measured using four outcome variables: psychiatric distress, orientation to recovery, residential satisfaction, and adaptive functioning. The study finds that neighborhood-level variables explained more variance in well-being than either apartment- or census-tract-level variables. Wright and Kloos (2007) present an important finding and buttress the case for including neighborhood-level variables in a measure of housing security. The study, however, does not take into consideration housing affordability and stability, which are key components of housing insecurity present in many other studies. In addition, it is a local sample for a subpopulation, which may not readily generalize to the U.S. population as a whole.

As previously mentioned, data availability could pose a challenge to studying the full continuum of housing insecurity. However, the AHS currently provides readily available data on many of the housing insecurity dimensions for conducting a proof-of-concept analysis for a uniform instrument based on the seven dimensions proposed in this paper (see Cox et al., 2017b for an analysis). As previously mentioned, the AHS is administered biennially by HUD and contains questions pertinent to housing affordability, quality, and safety, as well as neighborhood safety and quality. The AHS also contains some information on housing instability, although measurement is incomplete and inconsistent across survey waves for the variables that could capture housing instability. For example, overcrowding and eviction were added in 1997, while other variables related to housing instability, such as doubling up, foreclosure, and frequent moves, are only included in the 2013 survey. The AHS is representative at a national level, for certain cities, and at the metropolitan area level, making it a useful tool for national, regional, and major metropolitan area studies. One drawback of the AHS is it follows the housing unit and not the household as the unit of analysis, which could potentially cause additional complications when constructing an externally valid instrument that measures population estimates of housing insecurity at the household level. Specifically, the AHS has a unique longitudinal study design, in which housing units sampled in a specific year are followed over time. This allows for stakeholders to obtain information on how housing stock and its occupants change over time but may require more thought about implementing a housing security module and, more importantly, what such a module would measure in this type of survey design. This is especially true for capturing the housing insecurity domain of homelessness: AHS is not a reliable instrument to measure current homelessness because only individuals with housing are surveyed.

Discussion

Current working definitions of housing insecurity are inconsistent and incomplete, which may lead to households that are actually experiencing housing insecurity to be counted as housing secure. Different measures for the same phenomenon have significant potential to cause comparability

issues and, potentially, omitted variable bias if certain facets within subcategories are omitted from measurement. Moreover, having multiple terms to describe a general social problem could lead to confusion (not just among researchers), making it hard to mobilize resources and efforts to solve the problem. Therefore, based on the conceptual framework put forth by the U.N. General Committee and the U.S. Department of Health and Human Services (HHS), we propose an operational definition of housing insecurity that includes seven dimensions: housing stability, housing affordability, housing quality, housing safety, neighborhood safety, neighborhood quality, and homelessness.

This is not to say that research or policy should not individually focus on the different facets of housing insecurity when appropriate; rather, it suggests that to precisely estimate the prevalence of this issue and accurately account for its social costs, we need a comprehensive approach to measure the multiple dimensions of housing insecurity.

Housing insecurity measurement is a global problem that will require the deployment of resources from not only the research community but also policymakers and practitioners to solve. Without a comprehensive instrument, many aspects of housing insecurity will continue to be understudied, and their prevalence and cost to society will not be fully understood. Consequently, it will be difficult to create policies that can impact the plethora of well-being outcomes for adults and children associated with housing, which could lead to greater inefficiency and waste of resources. Given the potentially large social costs of housing insecurity, it is in the national interest of the United States to ensure a uniform, functional definition of housing insecurity that can capture all its dimensions. That means the government should mobilize and coordinate the resources needed for the development of a national housing insecurity measure. Given its knowledge base and skill set, the research community has a comparative advantage in creating, legitimizing, and validating any instrument developed through this process.

A paradigm for the creation of a uniform housing insecurity instrument already exists with the U.S. Food Security Survey Module, which has been incorporated into the U.S. Census Bureau's Current Population Survey (CPS) annually since 1995 (National Research Council, 2006). In 1984, the President's Task Force on Food Assistance determined that the lack of a valid indicator of hunger constrained public policy concerning this issue. In October 1990, Congress ratified the National Nutrition Monitoring and Related Research Act, authorizing the preparation and implementation of a 10-year plan that incorporated as one of its goals to "establish and improve the quality of national nutritional and health status data and related databases and networks, and stimulate research necessary to develop uniform indicators, standards, methodologies, technologies, and procedures for nutrition monitoring."[13] This led to the creation of the first federal food security instrument in 1995, which was included in the CPS.

The Food Security Module has now been included in the CPS for more than 20 years; as a result, society's understanding of food insecurity has grown tremendously. Specifically, we know more about the risk factors that cause food insecurity and whether food insecurity is a chronic or temporary state. Additionally, food security is now considered a key indicator of the well-being of households

[13] Nutrition Monitoring and Related Research Act of 1990. H.R. 1068, 101st Congress (1989–1990). https://www.congress.gov/bill/101st-congress/house-bill/1608/text.

and children, along with being considered an important outcome in the program evaluation of food assistance programs. All these accomplishments stemmed from a transdisciplinary effort comprising practitioners, policymakers, and academics devoted to developing a validated scale over 20 years that has been used to understand and help solve one of society's most intractable problems (see Coleman-Jensen, 2015 and appendix exhibit A-4 for a full timeline of the creation of the food security module). The time has now come to do the same for housing.

Acknowledgments

The authors would like to thank Christine Sabuni and Anna Sporrong for their assistance with this project.

Authors

Robynn Cox, is an assistant professor at the USC Suzanne Dworak-Peck School of Social Work and can be emailed at robynnco@usc.edu.

Benjamin Henwood, is an associate professor at the USC Suzanne Dworak-Peck School of Social Work and can be emailed at bhenwood@usc.edu.

Seva Rodnyansky, was a graduate student at the USC Sol Price School of Public Policy when this article was written. He is currently a post-doctoral scholar at the Goldman School of Public Policy at the University of California, Berkeley and can be emailed at seva@berkeley.edu.

Eric Rice, is an associate professor at the USC Suzanne Dworak-Peck School of Social Work and can be emailed at ericr@usc.edu.

Suzanne Wenzel, is the Richard and Ann Thor Professorship in Urban Development and professor of social work at the USC Suzanne Dworak-Peck School of Social Work and can be emailed at swenzel@usc.edu.

Appendix A

Exhibit A.1

Overall Motivation of Research on Housing Insecurity

Reason	Percent of Papers (N=106)
Explanatory Variable	52
Outcome	42
Prevalence	16
Outcome and Explanatory Variables	7
Outcome and Prevalence	2
Prevalence and Explanatory Variable	2
International Context	10

Exhibit A.2

Housing Insecurity Definition by Search Term, Based on Literature Review

Search Term	Percent of Papers Mentioning Search Term (N=102)	Percent of Papers Mentioning Search Term (Excluding Homelessness (N=74)
Housing Insecurity	16.0	16.4
Housing Affordability	5.7	5.7
Housing Instability	42.5	37.0
Homelessness	34.0	46.6
Housing Quality	2.8	1.4
Housing Satisfaction	0.9	0
Housing Needs	0.9	0
Housing Hardship	4.7	6.8
Neighborhood Safety	1.9	2.7
Housing Safety	0.9	1.4

Exhibit A.3

Measures Used to Capture Housing Insecurity Within the Domains of Housing Affordability and Housing Instability

Measure	Percent of Papers Using Measure
Housing Affordability	N=30
Difficulty/Inability to Make Payments	83
Housing Cost Burden	37
Foreclosure	10
Legal Housing Issues	7
Rental Assistance	3
Multiple Measures	30
Housing Instability	N=49
Multiple Moves	59
Doubling Up	35
Eviction	18
Overcrowding	18
Duration of Stay	14
Forced Moves	8
Living in Unstable Conditions	2
Living in Multiple Subsidized Units	2
Multiple Measures	39

Exhibit A.4

Timeline of the Creation of a Food Security Module

A History of Food Security Measurement in the United States

January 1984

The President's Task Force on Food Assistance concluded that the absence of any credible indicator of the extent of hunger "contributes to a climate in which policy discussions become unhelpfully heated, and unsubstantiated assertions are then substituted for hard information."

October 1990

The National Nutrition Monitoring and Related Research Act of 1990 became law and included the Ten-Year Comprehensive Plan to "Recommend a standardized mechanism and instrument(s) for defining and obtaining data on the prevalence of "food insecurity" or "food insufficiency" in the United States.

1992

USDA staff began reviewing existing research on the conceptual basis for measuring food insecurity and the practical problems of developing a food security questionnaire for surveys.

January 1984

USDA's Food and Nutrition Service and the U.S. Department of Health and Human Services' National Center for Health Statistics sponsored the *National Conference on Food Security Measurement and Research*. The conference identified the conceptual basis for a national measure of food insecurity and for implementing it in national surveys.

April 1985

The first Federal food security survey was administered as part of the U.S. Census Bureau's annual Current Population Survey. It has been fielded each year since.

1998

ERS assumed sponsorship of the Census Bureau's annual food security survey, along with responsibility for analyzing and reporting the data and for coordinating ongoing USDA research on food security.

2003

At USDA's request, an expert panel was convened by the Committee on National Statistics of the National Academies to conduct a thorough review of the food security measurement methods and the language used to describe those conditions.

November 2006

USDA introduced new language to describe ranges of severity of food insecurity in response to recommendations by the expert panel. "Food insecurity without hunger" was renamed "low food security," and "food insecurity with hunger" was renamed "very low food security."

2011

The food security survey questions were added to the National Health Interview Survey, expanding the list of Federal surveys that had previously incorporated the questions, including the Panel Study of Income Dynamics, National Health and Nutrition Examination Survey, Early Childhood Longitudinal Studies, Survey of Income and Program Participation, and others.

2014

The Food and Agriculture Organization of the United Nations created and fielded the *Food Insecurity Experience Scale* (FIES) in 200 languages in 150 countries as part of the Gallup World Poll. The FIES is based in part on USDA's food security measure.

September 2015

ERS released the 20th year of annual food security survey data.

Source: USDA, Economic Research Service

Source: Reprinted from Coleman-Jensen (2015). https://www.ers.usda.gov/webdocs/charts/62517/oct15_feature_jensen_fig05.png?v=42276.

Appendix B

Below is a list of readings not cited in the text but incorporated in our literature review. Together, the additional readings and the references comprise the 106 studies we consulted to generate our literature review. These papers were selected by conducting literature searches in Google Scholar and Web of Science using the following terms: "housing security," "housing insecurity," "housing affordability," "housing stability," "housing instability," "homelessness," "housing hardship," "housing quality," "housing safety," and "neighborhood safety." Relevant papers either (1) included housing security as an explanatory variable measuring outcomes on households, individuals, or populations; (2) sought to understand reasons for housing security as a dependent variable; or (3) assessed the prevalence of housing insecurity. Our search focused on the U.S. context; however, several articles from other developed countries were relevant to the development of a measure of housing insecurity. A table summarizing all of the articles can be found in Cox et al. (2017a).

Additional Reading

Allgood, S., and R.S. Warren, Jr. 2003. "The duration of homelessness: Evidence from a national survey," *Journal of Housing Economics* 12: 273–290

Aubry, T., F. Klodawsky, and D. Coulombe. 2011. "Comparing the housing trajectories of different classes within a diverse homeless population," *American Journal of Community Psychology* 49: 142–155. https://doi.org/10.1007/s10464-011-9444-z.

Baker, E., K. Mason, and R. Bentley. 2015. "Measuring housing affordability: A longitudinal approach," *Urban Policy and Research* 33: 275–290. https://www.tandfonline.com/doi/abs/10.1080/08111146.2015.1034853.

Bassuk, E.L., J.C. Buckner, L.F. Weinreb, A. Browne, S.S. Bassuk, R. Dawson, and J.N. Perloff. 1997. "Homelessness in female-headed families: Childhood and adult risk and protective factors," *American Journal of Public Health* 87: 241–248. https://ajph.aphapublications.org/doi/10.2105/AJPH.87.2.241.

Bassuk, E.L., J.N. Perloff, and R. Dawson. 2001. "Multiply homeless families: The insidious impact of violence," *Housing Policy Debate* 12: 299–320. https://www.tandfonline.com/doi/abs/10.1080/10511482.2001.9521407.

Bebout, R.R., R.E. Drake, H. Xie, G. McHugo, and M. Harris. 1997. "Housing status among formerly homeless dually diagnosed adults," *Psychiatric Services* 48: 936–941. https://ps.psychiatryonline.org/doi/abs/10.1176/ps.48.7.936.

Bennett, G.G., L.H. McNeill, K.Y. Wolin, D.T. Duncan, E. Puleo, and K.M. Emmons. 2007. "Safe to walk? Neighborhood safety and physical activity among public housing residents," *PLoS Med* 4 (10): e306.

Berzin, S.C., A.M. Rhodes, and M.A. Curtis. 2011. "Housing experiences of former foster youth: How do they fare in comparison to other youth?" *Children and Youth Services Review* 33: 2119–2126. https://psycnet.apa.org/record/2011-14192-001.

Bossarte, R.M., J.R. Blosnich, R.I. Piegari, L.L. Hill, and V. Kane. 2013. "Housing instability and mental distress among US veterans," *American Journal of Public Health* 103: S213–S216. https://ajph.aphapublications.org/doi/10.2105/AJPH.2013.301277.

Bradley, K.H., R.B. Oliver, N.C. Richardson, and E.M. Slayter. 2001. *No Place Like Home: Housing and the Ex-Prisoner*. Issue brief. Boston, MA: Community Resources for Justice. http://www.crj.org/cji/entry/publication_noplacelikehome.

Brisson, D., and J. Covert. 2015. "Housing instability risk among subsidized housing recipients: Characteristics associated with late or nonpayment of rent," *Social Work Research* 39: 119–128. https://academic.oup.com/swr/article-abstract/39/2/119/1649409?redirectedFrom=fulltext.

Brueckner, M., M. Green, and S. Saggers. 2011. "The trappings of home: Young homeless people's transitions towards independent living," *Housing Studies* 26: 1–16. https://www.tandfonline.com/doi/abs/10.1080/02673037.2010.512751.

Buckner, J. C., E.L. Bassuk, and L.F. Weinreb. 2001. "Predictors of academic achievement among homeless and low-income housed children," *Journal of School Psychology* 39 (1): 45–69.

Burgard, S.A., K.S. Seefeldt, and S. Zelner. 2012. "Housing instability and health: Findings from the Michigan Recession and Recovery Study," *Social Science & Medicine* 75: 2215–2224. https://www.ncbi.nlm.nih.gov/pubmed/22981839.

Burt, M.R. 1991. "Causes of the growth of homelessness during the 1980s," Housing Policy Debate 2: 901–936. https://www.tandfonline.com/doi/abs/10.1080/10511482.1991.9521077.

Cancian, M., and D.R. Meyer. 2004. "Alternative measures of economic success among TANF participants: Avoiding poverty, hardship, and dependence on public assistance," *Journal of Policy Analysis and Management* 23 (3): 531–548.

Carrion, B.V., V.A. Earnshaw, T. Kershaw, J.B. Lewis, E.C. Stasko, J.N. Tobin, and J.R. Ickovics. 2014. "Housing instability and birth weight among young urban mothers," *Journal of Urban Health* 92: 1–9. https://www.ncbi.nlm.nih.gov/pubmed/25344356.

Caton, C.L.M., B. Dominguez, B. Schanzer, D.S. Hasin, P.E. Shrout, A. Felix, E. Hsu, et al. 2005. "Risk factors for long-term homelessness: Findings from a longitudinal study of first-time homeless single adults," *American Journal of Public Health* 95: 1753–1759. https://ajph.aphapublications.org/doi/full/10.2105/AJPH.2005.063321.

Clark, S.L. 2010. *Housing instability: Toward a better understanding of frequent residential mobility among America's urban poor*. Washington, DC: Center for Housing Policy.

Cobb-Clark, D.A., N. Herault, R. Scutella, and Y-P. Tseng. 2015. "A journey home: What drives how long people are homeless?" *Journal of Urban Economics* 91: 57–72.

Coley, R.L., T. Leventhal, A.D. Lynch, and M. Kull. 2013. "Relations between housing characteristics and the well-being of low-income children and adolescents," *Developmental Psychology* 49: 1775–1789. https://www.ncbi.nlm.nih.gov/pubmed/23244408.

Cook-Craig, P., and L. Koehly. 2011. "Stability in the social support networks of homeless families in shelter: Findings from a study of families in a faith-based shelter program," *Journal of Family Social Work* 14: 191–207. https://www.tandfonline.com/doi/abs/10.1080/10522158.2011.571553.

Curtis, M.A., H. Corman, K. Noonan, and N.E. Reichman. 2013. "Life shocks and homelessness," *Demography* 50: 2227–2253. https://www.ncbi.nlm.nih.gov/pubmed/23868747.

Dickson-Gomez, J., M. Convey, H. Hilario, A.M. Corbett, and M. Weeks. 2008. "Structural and personal factors related to access to housing and housing stability among urban drug users in Hartford, Connecticut," *Contemporary Drug Problems* 35: 115–152. https://journals.sagepub.com/doi/10.1177/009145090803500106.

Dornbusch, S.M. 1994. "Additional perspectives on homeless families," *American Behavioral Scientist* 37: 404–411. https://journals.sagepub.com/doi/abs/10.1177/0002764294037003005.

Eastwood, E.A., and J.M. Birnbaum. 2007. "Physical and sexual abuse and unstable housing among adolescents with HIV," *AIDS and Behavior* 11: 116. https://link.springer.com/article/10.1007/s10461-007-9236-5.

Fowler, P.J., D.B. Henry, and K.E. Marcal. 2015. "Family and housing instability: Longitudinal impact on adolescent emotional and behavioral well-being," *Social Science Research* 53: 364–374. https://www.ncbi.nlm.nih.gov/pubmed/26188460.

Frederick, T.J., M. Chwalek, J. Hughes, J. Karabanow, and S. Kidd. 2014. "How stable is stable? Defining and measuring housing stability," *Journal of Community Psychology* 42: 964–979. https://onlinelibrary.wiley.com/doi/abs/10.1002/jcop.21665.

German, D., M.A. Davey, and C.A. Latkin. 2007. "Residential transience and HIV risk behaviors among injection drug users," *AIDS and Behavior* 11: 21. https://link.springer.com/article/10.1007/s10461-007-9238-3.

Glomm, G., and A. John. 2002. "Homelessness and labor markets," *Regional Science and Urban Economics* 32: 591–606. https://www.sciencedirect.com/science/article/pii/S0166046201000977?via%3Dihub.

Gould, T.E., and A.R. Williams. 2010. "Family homelessness: An investigation of structural effects," *Journal of Human Behavior in the Social Environment* 20: 170–192. https://www.tandfonline.com/doi/abs/10.1080/10911350903269765.

Greenberg, G.A., and R.A. Rosenheck. 2008a. "Homelessness in the state and federal prison population," *Criminal Behaviour and Mental Health* 18 (2): 88–103.

———. 2008b. "Jail incarceration, homelessness, and mental health: A national study," *Psychiatric Services* 59 (2): 170–177.

Harris, C., and D. Wagner. 2016. "The Navajo Nation accepted more than $1 billion for houses. So, where did it go?" *The Arizona Republic*. https://www.azcentral.com/story/news/local/arizona-investigations/2016/12/14/navajo-housing-federal-funds/94563354/.

Heflin, C., J. Sandberg, and P. Rafail. 2009. "The structure of material hardship in US households: An examination of the coherence behind common measures of well-being," *Social Problems* 56 (4): 746–764.

Herman, D.B., E.S. Susser, E.L. Struening, and B.L. Link. 1997. "Adverse childhood experiences: Are they risk factors for adult homelessness?" *American Journal of Public Health* 87: 249–255. https://ajph.aphapublications.org/doi/abs/10.2105/AJPH.87.2.249.

Hodgson, K.J., K.H. Shelton, M.B.M. van den Bree, and F.J. Los. 2013. "Psychopathology in young people experiencing homelessness: A systematic review," *American Journal of Public Health* 103: e24–e37. https://ajph.aphapublications.org/doi/10.2105/AJPH.2013.301318.

Hurlbut, M.S., R.L. Hough, and P.A. Wood. 1996. "Effects of substance abuse on housing stability of homeless mentally ill persons in supported housing," *Psychiatric Services* 47: 731–736. https://ps.psychiatryonline.org/doi/10.1176/ps.47.7.731.

Jones, L. 2011. "The impact of transitional housing on the post-discharge functioning of former foster youth," *Residential Treatment for Children & Youth* 28: 17–38. https://www.tandfonline.com/doi/abs/10.1080/0886571X.2011.541843.

Keall, M., M.G. Baker, P. Howden-Chapman, M. Cunningham, and D. Ormandy. 2010. "Assessing housing quality and its impact on health, safety and sustainability," *Journal of Epidemiology & Community Health*, 64 (9): 754–771.

Khodyakov, D., A. Ochoa, B.L. Olivieri-Mui, C. Bouwmeester, B.J. Zarowitz, M. Patel, B. Briesacher, et al. 2017. "Screening tool of older person's prescriptions/screening tools to alert doctors to right treatment medication criteria modified for U.S. nursing home setting," *Journal of the American Geriatrics Society* 65 (3): 586–591.

Kirkpatrick, S.I., and V. Tarasuk. 2011. "Housing circumstances are associated with household food access among low-income urban families," *Journal of Urban Health* 88: 284–296. https://www.ncbi.nlm.nih.gov/pubmed/21286826.

Knopf-Amelung, Sarah. 2013. "Aging and housing instability: Homelessness among older and elderly adults." *In Focus: A Quarterly Research Review of the National HCH Council* 2 (1). Nashville, TN: National Health Care for the Homeless Council. https://www.nhchc.org/wp-content/uploads/2011/09/infocus_september2013.pdf.

Koegel, P., E. Melamid, and M.A. Burnam. 1995. "Childhood risk factors for homelessness among homeless adults," *American Journal of Public Health* 85: 1642–1649. https://ajph.aphapublications.org/doi/abs/10.2105/AJPH.85.12.1642.

Landree, E., R. Silberglitt, B.G. Chow, L. Sherry, and M.S. Tseng. 2009. A delicate balance: Portfolio analysis and management for intelligence information dissemination programs. Santa Monica, CA: RAND Corporation. https://www.rand.org/pubs/monographs/MG939.html.

Los Angeles County Department of Public Health. 2015. *Social determinants of health: Housing and health in Los Angeles County*. Los Angeles, CA: Los Angeles County Department of Public Health.

Ma, C.T., Lauren Gee, and M.B. Kushel, M.D. 2008. "Associations between housing instability and food insecurity with health care access in low-income children," *Ambulatory Pediatrics* 8 (1): 50–57.

Margolis, A.D., R.J. MacGowan, O. Grinstead, J. Sosman, I. Kashif, T.P. Flanigan, and Project START Study Group. 2006. "Unprotected sex with multiple partners: Implications for HIV prevention among young men with a history of incarceration," *Sexually Transmitted Diseases* 33 (3): 175–180.

Martinez, J., D. Bell, R. Camacho, L.M. Henry-Reid, M. Bell, C. Watson, and F. Rodriguez. 2000. "Adherence to antiviral drug regimens in HIV-infected adolescent patients engaged in care in a comprehensive adolescent and young adult clinic," *Journal of the National Medical Association* 92: 55–61.

Mayberry, L.S., M. Shinn, J.G. Benton, and J. Wise. 2014. "Families experiencing housing instability: The effects of housing programs on family routines and rituals," *American Journal of Orthopsychiatry* 84: 95–109. https://www.ncbi.nlm.nih.gov/pubmed/24826832.

McVicar, D., J. Moschion, and J.C. van Ours. 2015. "From substance use to homelessness or vice versa?" *Social Science & Medicine* 136–137: 89–98. https://www.ncbi.nlm.nih.gov/pubmed/25989002.

Murphy, R.D., S.E. Zemore, and N. Mulia. 2013. "Housing instability and alcohol problems during the 2007–2009 US recession: The moderating role of perceived family support," *Journal of Urban Health* 91: 17–32. https://link.springer.com/article/10.1007/s11524-013-9813-z.

Nakazato, N., U. Schimmack, and S. Oishi. 2010. "Effect of changes in living conditions on well-being: A prospective top–down bottom–up model," *Social Indicators Research* 100: 115–135. https://link.springer.com/article/10.1007/s11205-010-9607-6.

Nemiroff, R., T. Aubry, and F. Klodawsky. 2010. "Factors contributing to becoming housed for women who have experienced homelessness," *Canadian Journal of Urban Research* 19: 23–45.

Obradović, J., J.D. Long, J.J. Cutuli, C.K. Chan, E. Hinz, D. Heistad, and A.S. Masten. 2009. "Academic achievement of homeless and highly mobile children in an urban school district: Longitudinal evidence on risk, growth, and resilience," *Development and Psychopathology* 21 (2): 493–518.

Orwin, R.G., C.K. Scott, and C. Arieira. 2005. "Transitions through homelessness and factors that predict them," *Journal of Substance Abuse Treatment* 28: S23–S39. https://www.journalofsubstanceabusetreatment.com/article/S0740-5472(04)00133-3/abstract.

Pavao, J., J. Alvarez, N. Baumrind, M. Induni, and R. Kimerling. 2007. "Intimate partner violence and housing instability," *American Journal of Preventive Medicine* 32: 143–146. https://www.ajpmonline.org/article/S0749-3797(06)00434-X/abstract.

Phinney, R., S. Danziger, H.A. Pollack, and K. Seefeldt. 2007. "Housing instability among current and former welfare recipients," *American Journal of Public Health* 97: 832–837. ajph.aphapublications.org/doi/10.2105/AJPH.2005.082677.

Piliavin, I., B.R. Entner Wright, R.D. Mare, and A.H. Westerfelt. 1996. "Exits from and returns to homelessness," *Social Service Review* 70: 33–57. https://journals.sagepub.com/doi/abs/10.1177/0044118X11402851.

Pindus, N., T.G. Kingsley, J. Biess, D. Levy, J. Simington, and C. Hayes. 2017. *Housing needs of American Indians and Alaska Natives in tribal areas: A report from the Assessment of American Indian, Alaska Native, and Native Hawaiian housing needs*. Washington, DC: U.S. Department of Housing and Urban Development: Office of Policy Development and Research.

Quigley, J.M., and S. Raphael. 2001. "The economics of homelessness: The evidence from North America," *International Journal of Housing Policy* 1: 323–336. https://www.tandfonline.com/doi/abs/10.1080/14616710110091525.

Reid, K.W., E. Vittinghoff, and M.B. Kushel. 2008. "Association between the level of housing instability, economic standing and health care access: A meta-regression," *Journal of Health Care for the Poor and Underserved* 19: 1212–1228. https://www.ncbi.nlm.nih.gov/pubmed/19029747.

Rog, D.J., and J.C. Buckner. 2007. "Homeless families and children." In *Toward Understanding Homelessness: The 2007 National Symposium on Homelessness Research* 4: 1–33. Washington, DC: U.S. Department of Housing and Urban Development. https://www.huduser.gov/portal/publications/homeless/p5.html.

Rollins, C., N.E. Glass, N.A. Perrin, K.A. Billhardt, A. Clough, J. Barnes, G.C. Hanson, and T.L. Bloom. 2012. "Housing instability is as strong a predictor of poor health outcomes as level of danger in an abusive relationship: Findings from the SHARE Study," *Journal of Interpersonal Violence* 27: 623–643. https://journals.sagepub.com/doi/abs/10.1177/0886260511423241?journalCode=jiva.

Rosenthal, D., M.J. Rotheram-Borus, P. Batterham, S. Mallett, E. Rice, and N.G. Milburn. 2007. "Housing stability over two years and HIV risk among newly homeless youth," AIDS and Behavior 11: 831–841. https://europepmc.org/abstract/med/17497219.

Rubin, D.H., C.J. Erickson, M. San Agustin, S.D. Cleary, J.K. Allen, and P. Cohen. 1996. "Cognitive and academic functioning of homeless children compared with housed children," *Pediatrics* 97 (3): 289–294.

Shelton, K., P. Taylor, A. Bonner, and M. van den Bree. 2009. "Risk factors for homelessness: Evidence from a population-based study," *Psychiatric Services* 60: 465–472. https://ps.psychiatryonline.org/doi/full/10.1176/ps.2009.60.4.465.

Stone, Kimberly E., Emmanuella M. Eastman, Andrea C. Gielen, Barbara Squires, Glenda Hicks, Dana Kaplin, and Janet R. Serwint. 2007. "Home safety in inner cities: Prevalence and feasibility of home safety-product use in inner-city housing," *Pediatrics* 120 (2): e346–e353. https://pediatrics.aappublications.org/content/120/2/e346..info.

Taylor, M.P., D.J. Pevalin, and J. Todd. 2007. "The psychological costs of unsustainable housing commitments," *Psychological Medicine* 37 (7): 1027–1036.

Toro, P.A., J.M.P. Rabideau, C.W. Bellavia, C.V. Daeschler, D.D. Wall, D.M. Thomas, and S.J. Smith. 1997. "Evaluating an intervention for homeless persons: Results of a field experiment," *Journal of Consulting and Clinical Psychology* 65: 476–484. https://psycnet.apa.org/record/1997-06366-013.

Tsemberis, S., G. McHugo, V. Williams, P. Hanrahan, and A. Stefancic. 2007. "Measuring homelessness and residential stability: The Residential Time-Line Follow-Back Inventory," *Journal of Community Psychology* 35: 29–42. https://onlinelibrary.wiley.com/doi/abs/10.1002/jcop.20132.

U.S. Commission on Human Rights (USCHR). 2003. *A quiet crisis: Federal funding and unmet needs in Indian country*. http://www.usccr.gov/pubs/na0703/na0204.pdf.

U.S. Department of Housing and Urban Development (HUD). 2016. *Obama administration announces nearly 50 percent decline in veteran homelessness*. Press release.

U.S. General Accounting Office (GAO). 1994. *Elementary school children: Many change schools frequently, harming their education*. http://archive.gao.gov/t2pbat4/150724.pdf.

U.S. Government Accountability Office (GAO). 2010. *Native American housing: Tribes generally view block grant program as effective, but tracking of infrastructure plans and investments needs improvement*. http://www.gao.gov/assets/310/301157.pdf.

U.S. Interagency Council on Homelessness (USICH). 2011. *Opening doors. Federal strategic plan to prevent and end homelessness. Update 2011*. https://www.usich.gov/resources/uploads/asset_library/USICH_FSPUpdate_2012_12312.pdf.

Vega, W.A., and S.P. Wallace. 2016. "Affordable housing: A key lever to community health for older Americans," *American Journal of Public Health* 106: 635–636. https://ajph.aphapublications.org/doi/abs/10.2105/AJPH.2015.303034.

Yoo, J.P., K.S. Slack, and J.L. Holl. 2009. "Material hardship and the physical health of school-aged children in low-income households," *American Journal of Public Health* 99 (5): 829–836.

Vijayaraghavan, M., E.A. Jacobs, H. Seligman, and A. Fernandez. 2011. "The association between housing instability, food insecurity, and diabetes self-efficacy in low-income adults," *Journal of Health Care for the Poor and Underserved* 22: 1279–1291. https://www.ncbi.nlm.nih.gov/pubmed/22080709.

References

Bailey, Kathryn T., John T. Cook, Stephanie Ettinger de Cuba, Patrick H. Casey, Mariana Chilton, Sharon M. Coleman, Diana Becker Cutts, Timothy C. Heeren, Ruth Rose-Jacobs, Maureen M. Black, and Deborah A. Frank. 2016. "Development of an Index of Subsidized Housing Availability and its Relationship to Housing Insecurity," *Housing Policy Debate* 26 (1): 172–187. https://www.tandfonline.com/doi/abs/10.1080/10511482.2015.1015042.

Bashir, Samiya A. 2002. "Home is Where the Harm is: Inadequate Housing as a Public Health Crisis," *American Journal of Public Health* 92 (5): 733–738. DOI: 10.2105/ajph.92.5.733.

Buffardi, Anne L., Kathy K. Thomas, King K. Holmes, and Lisa E. Manhart. 2008. "Moving Upstream: Ecosocial and Psychosocial Correlates of Sexually Transmitted Infections Among Young Adults in the United States," *American Journal of Public Health* 98 (6): 1128–1136. DOI: 10.2105/AJPH.2007.120451.

Byrne, Thomas, Benjamin F. Henwood, and Brynn Scriber. 2018. "Residential moves among housing first participants," *Journal of Behavioral Health Services & Research* 45 (1): 124–132. DOI: 10.1007/s11414-016-9537-4.

Campbell, Iain, Sharon Parkinson, and Gavin Wood. 2014. *Underemployment and housing insecurity: An empirical analysis of HILDA data*. Melbourne, Australia: Australian Housing and Urban Research Institute.

Capps, Randy. 2001. *Hardship Among Children of Immigrants: Findings from the 1999 National Survey of American Families*. Washington, DC: Urban Institute. http://webarchive.urban.org/publications/310096.html.

Coleman-Jensen, Alisha. 2015. "Commemorating 20 Years of U.S. Food Security Measurement," *Amber Waves*, October 5. Washington, DC: U.S. Department of Agriculture Economic Research Service. https://www.ers.usda.gov/amber-waves/2015/october/commemorating-20-years-of-us-food-security-measurement/.

Coulton, Claudia, Cara J. Pasqualone, Toby Martin, Neil Bania, Nina Lalich, and Lisa Nelson. 2001. *Issues of Housing Affordability and Hardship Among Cuyahoga County Families Leaving Welfare: Quarter 4, 1998–Quarter 3, 1999*, Special Topics in Welfare Reform Report No. 1. Cleveland, OH: Center on Urban Poverty and Social Change.

Cox, Robynn, Benjamin Henwood, Seva Rodnyansky, Suzanne Wenzel, and Eric Rice. 2017a. Roadmap to a Unified Measure of Housing Insecurity. Working paper 07112017. Washington, DC: Washington Center for Equitable Growth. https://equitablegrowth.org/working-papers/unified-measure-of-housing-insecurity/.

Cox, Robynn, Seva Rodnyansky, Benjamin Henwood, and Suzanne Wenzel. 2017b. Measuring Population Estimates of Housing Insecurity in the United States: A Comprehensive Approach. Working paper 12192017. Washington, DC: Washington Center for Equitable Growth. https://equitablegrowth.org/working-papers/measuring-housing-insecurity/.

Crowder, Kyle, and Liam Downey. 2010. "Inter-Neighborhood Migration, Race, and Environmental Hazards: Modeling Micro-Level Processes of Environmental Inequality," *American Journal of Sociology* 115 (4): 1110–1149. DOI: 10.1086/649576.

Curtis, Marah A., and Amanda Geller. 2010. Housing Insecurity Among Urban Fathers. Working paper no. 10-08. New York: Columbia Population Research Center. DOI: 10.7916/D8WH2W9T.

Cutts, Diana Becker, Alan F. Meyers, Maureen M. Black, Patrick H. Casey, Mariana Chilton, John T. Cook, Joni Geppert, Stephanie Ettinger de Cuba, Timothy Heeren, Sharon Coleman, Ruth Rose-Jacobs, and Deborah A. Frank. 2011. "US Housing Insecurity and the Health of Very Young Children," *American Journal of Public Health* 101 (8): 1508–1514. DOI: 10.2105/AJPH.2011.300139.

Diette, Timothy M., and David C. Ribar. 2018. "A Longitudinal Analysis of Violence and Housing Insecurity," *Economic Inquiry* 56 (3): 1602–1621. DOI: 10.1111/ecin.12571.

Drake, Robert E., Michael A. Wallach, and J. Schuyler Hoffman. 1989. "Housing Instability and Homelessness Among Aftercare Patients of an Urban State Hospital," *Psychiatric Services* 40 (1): 46–51. DOI: 10.1176/ps.40.1.46.

Dupuis, Ann, and David C. Thorns. 1998. "Home, Home Ownership and the Search for Ontological Security," *The Sociological Review* 46 (1): 24–47. DOI: 10.1111/1467-954X.00088.

Dworsky, Amy. 2008. *Educating Homeless Children in Chicago: A Case Study of Children in the Family Regeneration Program*. Chicago, IL: Chapin Hall at the University of Chicago. http://www.niusileadscape.org/docs/FINAL_PRODUCTS/LearningCarousel/EducatingHomelessChicago.pdf.

Early, Dirk W. 2005. "An Empirical Investigation of the Determinants of Street Homelessness," *Journal of Housing Economics* 14 (1): 27–47. DOI: 10.1016/j.jhe.2005.03.001.

Fowler, Patrick J., Paul A. Toro, and Bart W. Miles. 2011. "Emerging Adulthood and Leaving Foster Care: Settings Associated with Mental Health," *American Journal of Community Psychology* 47 (3-4): 335–348. DOI: 10.1007/s10464-010-9401-2.

Frank, Deborah A., Patrick H. Casey, Maureen M. Black, Ruth Rose-Jacobs, Mariana Chilton, Diana Cutts, Elizabeth March, Timothy Heeren, Sharon Coleman, Stephanie Ettinger de Cuba, and John T. Cook. 2010. "Cumulative Hardship and Wellness of Low-Income, Young Children: Multisite Surveillance Study," *Pediatrics* 125 (5): e1115–e1123. DOI: 10.1542/peds.2009-1078.

Geller, Amanda, and Allyson Walker Franklin. 2014. "Paternal Incarceration and the Housing Security of Urban Mothers," *Journal of Marriage and Family* 76 (2): 411–427. DOI: 10.1111/jomf.12098.

Goldrick-Rab, Sara, Katharine Broton, and Daniel Eisenberg. 2015. "Hungry to Learn: Addressing Food and Housing Insecurity Among Undergraduates." Madison, WI: Wisconsin Hope Lab. https://postsecondary.gatesfoundation.org/wp-content/uploads/2016/11/Hungry-to-Learn.pdf.

Greder, Kimberly, Christine Cook, Steven Garasky, and Lizmelia Ortiz. 2008. *Latino Immigrants: Food and Housing Insecurity*. Ames, IA: Iowa State University.

Heflin, Colleen, Andrew S. London, and Ellen K. Scott. 2011. "Mitigating Material Hardship: The Strategies Low-Income Families Employ to Reduce the Consequences of Poverty," *Sociological Inquiry* 81 (2): 223–246. DOI: 10.1111/j.1475-682X.2011.00369.x.

Henry, Meghan, Anna Mahathey, Tyler Morrill, Anna Robinson, Azim Shivji, and Rian Watt. 2018. *The 2018 Annual Homeless Assessment Report (AHAR) to Congress: Part 1: Point-in-Time Estimates of Homelessness*. Washington, DC: U.S. Department of Housing and Urban Development.

Henwood, Benjamin F., Leopoldo J. Cabassa, Catherine M. Craig, and Deborah K. Padgett. 2013. "Permanent Supportive Housing: Addressing Homelessness and Health Disparities?" *American Journal of Public Health* 103 (S2): S188–S192. DOI: 10.2105/AJPH.2013.301490.

Honig, Marjorie, and Randall K. Filer. 1993. "Causes of Intercity Variation in Homelessness," *The American Economic Review* 83 (1): 248–255.

Joint Center for Housing Studies of Harvard University. 2013. *The State of the Nation's Housing 2013*. Cambridge, MA. https://www.jchs.harvard.edu/sites/default/files/son2013_0.pdf.

———. 2012. *The State of the Nation's Housing 2012*. Cambridge, MA. https://www.jchs.harvard.edu/sites/default/files/son2012.pdf.

———. 2011. *The State of the Nation's Housing 2011*. Cambridge, MA. https://www.jchs.harvard.edu/sites/default/files/son2011.pdf.

———. 2008. *The State of the Nation's Housing 2008*. Cambridge, MA. https://www.jchs.harvard.edu/sites/default/files/son2008.pdf.

———. 2007. *The State of the Nation's Housing 2007*. Cambridge, MA. https://www.jchs.harvard.edu/sites/default/files/son2007.pdf.

———. 2006. *The State of the Nation's Housing 2006*. Cambridge, MA. https://www.jchs.harvard.edu/sites/default/files/son2006.pdf.

———. 2005. *The State of the Nation's Housing 2005*. Cambridge, MA. https://www.jchs.harvard.edu/sites/default/files/son2005.pdf.

———. 2003. *The State of the Nation's Housing 2003*. Cambridge, MA. https://www.jchs.harvard.edu/sites/default/files/son2003.pdf.

———. 2002. *The State of the Nation's Housing 2002*. Cambridge, MA. https://www.jchs.harvard.edu/sites/default/files/son2002.pdf.

Katz, Lawrence F., Jeffrey R. Kling, and Jeffrey B. Liebman. 2001. "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment," *The Quarterly Journal of Economics* 116 (2): 607–654. DOI: 10.1162/00335530151144113.

Kushel, Margot B., Reena Gupta, Lauren Gee, and Jennifer S. Haas. 2006. "Housing Instability and Food Insecurity as Barriers to Health Care Among Low-Income Americans," *Journal of General Internal Medicine* 21 (1): 71–77. DOI: 10.1111/j.1525-1497.2005.00278.x.

Lee, Barrett A., Kimberly A. Tyler, and James D. Wright. 2010. "The New Homelessness Revisited," *Annual Review of Sociology* 36: 501–521. DOI: 10.1146/annurev-soc-070308-115940.

Leopold, Josh, Mary Cunningham, Lily Posey, and Tiffany Manuel. 2017. Improving measures of housing insecurity: a path forward. Washington, D.C.: The Urban Institute.

Leventhal, Tama, and Jeanne Brooks-Gunn. 2003. "Moving to Opportunity: An Experimental Study of Neighborhood Effects on Mental Health," *American Journal of Public Health* 93 (9): 1576–1582.

Link, Bruce G., Ezra Susser, Ann Stueve, Joe Phelan, Robert E. Moore, and Elmer Struening. 1994. "Lifetime and Five-Year Prevalence of Homelessness in the United States," American Journal of Public Health 84 (12): 1907–1912.

Liu, Yong, Rashid S. Njai, Kurt J. Greenlund, Daniel P. Chapman, and Janet B. Croft. 2014. "Relationships Between Housing and Food Insecurity, Frequent Mental Distress, and Insufficient Sleep Among Adults in 12 US States, 2009," *Preventing Chronic Disease* 11: E37. DOI: 10.5888/pcd11.130334.

Low, Setha M., and Denise Lawrence-Zúñiga. 2003. *The Anthropology of Space and Place: Locating Culture*. Boston, MA: Wiley-Blackwell.

Ludwig, Jens, Greg J. Duncan, Lisa A. Gennetian, Lawrence F. Katz, Ronald C. Kessler, Jeffrey R. Kling, and Lisa Sanbonmatsu. 2013. "Long-Term Neighborhood Effects on Low-Income Families: Evidence from Moving to Opportunity," *The American Economic Review* 103 (3): 226–231. DOI: 10.1257/aer.103.3.226.

Marsh, Alex, David Gordon, Pauline Heslop, and Christina Pantazis. 2000. "Housing Deprivation and Health: A Longitudinal Analysis," *Housing Studies* 15 (3), 411–428. DOI: 10.1080/02673030050009258.

Mimura, Yoko. 2008. "Housing Cost Burden, Poverty Status, and Economic Hardship Among Low-Income Families," *Journal of Family and Economic Issues* 29 (1): 152–165.

National Law Center on Homelessness & Poverty (NLCHP). 2004. *Homelessness in the United States and the Human Right to Housing*. Washington, DC: National Law Center on Homelessness & Poverty. http://www.mplp.org/Resources/mplpresource.2006-06-13.0349156065/file0.

National Research Council. 2006. *Food Insecurity and Hunger in the United States: An Assessment of the Measure*. Washington, DC: The National Academies Press. http://www.nap.edu/catalog/11578/food-insecurity-and-hunger-in-the-united-states-an-assessment.

Padgett, Deborah K. 2007. "There's no place like (a) home: Ontological security among persons with serious mental illness in the United States," *Social Science & Medicine* 64: 1925–1936. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1934341/.

Padgett, Deborah K., Benjamin F. Henwood, and Sam J. Tsemberis. 2015. *Housing First: Ending Homelessness, Transforming Systems, and Changing Lives*. Oxford, UK: Oxford University Press.

Pearson, Carol, Ann Elizabeth Montgomery, and Gretchen Locke. 2009. "Housing Stability Among Homeless Individuals with Serious Mental Illness Participating in Housing First Programs," *Journal of Community Psychology* 37 (3): 404–417. DOI: 10.1002/jcop.20303.

Poortinga, Wouter, Tatiana Calve, Nikki Jones, Simon Lannon, Tabitha Rees, Sarah E. Rodgers, Ronan A. Lyons, and Rhodri Johnson. 2017. "Neighborhood Quality and Attachment: Validation of the Revised Residential Environment Assessment Tool," *Environment and Behavior* 49 (3): 255–282. DOI: 10.1177/0013916516634403.

Rourke, Sean B., Tsegaye Bekele, Ruthann Tucker, Saara Greene, Michael Sobota, Jay Koornstra, LaVerne Monette, Jean Bacon, Shafi Bhuiyan, Sergio Rueda, James Watson, Stephen W. Hwang, James Dunn, and Keith Hambly. 2012. "Housing Characteristics and Their Influence on Health-Related Quality of Life in Persons Living with HIV in Ontario, Canada: Results from the Positive Spaces, Healthy Places Study," *AIDS and Behavior* 16 (8): 2361–2373. DOI: 10.1007/s10461-012-0284-0.

Routhier, Giselle. 2019. "Beyond Worst Case Needs: Measuring the Breadth and Severity of Housing Insecurity Among Urban Renters," *Housing Policy Debate* 29 (2): 235–249. DOI: 10.1080/10511482.2018.1509228.

Sanbonmatsu, Lisa, Jeffrey R. Kling, Greg J. Duncan, and Jeanne Brooks-Gunn. 2006. "Neighborhoods and Academic Achievement Results from the Moving to Opportunity Experiment," *Journal of Human Resources* 41 (4): 649–691. DOI: 10.3368/jhr.XLI.4.649.

Shaw, Mary. 2004. "Housing and Public Health," *Annual Review of Public Health* 25: 397–418. DOI: 10.1146/annurev.publhealth.25.101802.123036.

Shinn, Marybeth, Beth C. Weitzman, Daniela Stojanovic, James R. Knickman, Lucila Jiménez, Lisa Duchon, Susan James, and David H. Krantz. 1998. "Predictors of Homelessness Among Families in New York City: From Shelter Request to Housing Stability," *American Journal of Public Health* 88 (11): 1651–1657. DOI: 10.2105/AJPH.88.11.1651.

Siebens, Julie. 2013. *Extended Measures of Well-Being: Living Conditions in the United States: 2011* Report No. P70-136. Washington, DC: U.S. Census Bureau. https://www.census.gov/library/publications/2013/demo/p70-136.html.

Solari, Claudia D., Sean Morris, Azim Shivji, and Tanya de Souza. 2016. *The 2015 Annual Homeless Assessment Report (AHAR) to Congress: PART 2: Estimates of Homelessness in the United States*. Washington, DC: U.S. Department of Housing and Urban Development.

Somerville, Peter. 1992. "Homelessness and the Meaning of Home: Rooflessness or Rootlessness?" *International Journal of Urban and Regional Research* 16 (4): 529–539. DOI: 10.1111/j.1468-2427.1992.tb00194.x.

Stahre, Mandy, Juliet VanEenwyk, Paul Siegel, and Rashid Njai. 2015. "Housing Insecurity and the Association with Health Outcomes and Unhealthy Behaviors, Washington State, 2011," *Preventing Chronic Disease* 12: 140511. DOI: 10.5888/pcd12.140511.

Suglia, Shakira Franco, Cristiane S. Duarte, and Megan T. Sandel. 2011. "Housing Quality, Housing Instability, and Maternal Mental Health," *Journal of Urban Health* 88 (6): 1105–1116. DOI: 10.1007/s11524-011-9587-0.

Surratt, Hilary L., Catherine L. O'Grady, Maria A. Levi-Minzi, and Steven P. Kurtz. 2015. "Medication Adherence Challenges Among HIV Positive Substance Abusers: The Role of Food and Housing Insecurity," *AIDS Care: Psychological and Socio-Medical Aspects of AIDS/HIV* 27 (3): 307–314. DOI: 10.1080/09540121.2014.967656.

Thurston, Wilfreda E., Amrita Roy, Barbara Clow, David Este, Tess Gordey, Margaret Haworth-Brockman, Liza McCoy, Rachel Rapaport Beck, Christine Saulnier, and Lesley Carruthers. 2013. "Pathways Into and Out of Homelessness: Domestic Violence and Housing Security for Immigrant Women," *Journal of Immigrant & Refugee Studies* 11 (3): 278–298. DOI: 10.1080/15562948.2013.801734.

Topp, Libby, Jenny Iversen, Eileen Baldry, and Lisa Maher. 2013. "Housing Instability Among People Who Inject Drugs: Results from the Australian Needle and Syringe Program Survey," *Journal of Urban Health* 90 (4): 699–716. DOI: 10.1007/s11524-012-9730-6.

Tsui, Emma, Nicholas Freudenberg, Luis Manzo, Hollie Jones, Amy Kwan, and Monica Gagnon. 2011. *Housing Instability at CUNY: Results from a Survey of CUNY Undergraduate Students*. New York: Healthy CUNY Initiative, City University of New York. https://www.gc.cuny.edu/CUNY_GC/media/CUNY-Graduate-Center/PDF/Centers/Center%20for%20Human%20Environments/cunyhousinginstability.pdf.

United Nations. 2014. *The Right to Adequate Housing*. Fact Sheet No. 21. http://www.ohchr.org/Documents/Publications/FS21_rev_1_Housing_en.pdf.

U.S. Department of Health and Human Services (HHS). 1969. *Ancillary Services to Support Welfare to Work: Housing Instability*. Washington, DC: U.S. Department of Health & Human Services, Office of the Assistant Secretary for Planning and Evaluation.

U.S. Department of Housing and Urban Development (HUD). 2018. *The 2018 Annual Homeless Assessment Report (AHAR) to Congress*. Washington, DC: Office of Community Planning and Development.

———. 2017. *The 2017 Annual Homeless Assessment Report (AHAR) to Congress*. Washington, DC: Office of Community Planning and Development.

———. 2017. *Worst Case Housing Needs: 2015 Report to Congress*. Washington, DC: Office of Policy Development and Research.

———. 2015. *Worst Case Housing Needs: 2015 Report to Congress*. Washington, DC: Office of Policy Development and Research.

———. 2012. *The 2012 Annual Homeless Assessment Report (AHAR) to Congress*. Washington, DC: Office of Community Planning and Development.

———. 2005. *Worst Case Housing Needs: 2005 Report to Congress*. Washington, DC: Office of Policy Development and Research.

———. 2003. *Trends in Worst Case Needs for Housing, 1978–1999. A Report to Congress on Worst Case Housing Needs*. Washington, DC: Office of Policy Development and Research.

U.S. Interagency Council on Homelessness (USICH). 2017. *Mayors Challenge to End Homelessness*. https://www.usich.gov/solutions/collaborative-leadership/mayors-challenge.

Urban Institute, and Child Trends. 2002. National Survey of America's Families (NSAF), 2002. Ann Arbor, MI: Inter-University Consortium for Political and Social Research [distributor], 2007-10-03. https://www.researchconnections.org/childcare/studies/4582/version/1.

———. 1999. National Survey of America's Families (NSAF), 1999. Ann Arbor, MI: Inter-University Consortium for Political and Social Research [distributor], 2007-10-03. https://www.researchconnections.org/childcare/studies/3927/version/1.

———. 1997. National Survey of America's Families (NSAF), 1997. Ann Arbor, MI: Inter-University Consortium for Political and Social Research [distributor], 2007-10-04. https://www.researchconnections.org/childcare/studies/4581/version/1.

von Hoffman, Alexander. 2000. "A Study in Contradictions: The Origins and Legacy of the Housing Act of 1949," *Housing Policy Debate* 11 (2): 299–326.

Warren, Emily J., and Sarah A. Font. 2015. "Housing Insecurity, Maternal Stress, and Child Maltreatment: An Application of the Family Stress Model," *Social Service Review* 89 (1): 9–39. DOI: 10.1086/680043.

Western, Bruce, Anthony Braga, and Rhiana Kohl. 2015. *The Boston Reentry Study: Housing Insecurity After Prison*. Cambridge, MA: Rappaport Institute for Greater Boston. http://scholar.harvard.edu/files/brucewestern/files/housingvfinal.pdf?m=1441125268.

Wright, Patricia Ann, and Bret Kloos. 2007. "Housing Environment and Mental Health Outcomes: A Levels of Analysis Perspective." *Journal of Environmental Psychology* 27 (1): 79–89. DOI: 10.1016/j.jenvp.2006.12.001.

Ziol-Guest, Kathleen M., and Claire C. McKenna. 2014. "Early Childhood Housing Instability and School Readiness," *Child Development* 85 (1): 103–113. DOI: 10.1111/cdev.12105.

The Effects of Flood Insurance on Housing Markets

Agustín Indaco
Carnegie Mellon University in Qatar

Francesc Ortega
Süleyman Taşpınar
Queens College, City University of New York

Abstract

In this article, we analyze the role of flood insurance on the housing markets of coastal areas. To do so, we assembled a parcel-level dataset of the universe of residential sales for two coastal urban areas in the United States—Miami-Dade County (2008–15) and Virginia Beach (2000–16)—matched with their Federal Emergency Management Agency (FEMA) flood maps, which characterize the flood-risk level for each property. First, we compare trends in housing values and sales activity among properties on the floodplain, as defined by the National Flood Insurance Program (NFIP), relative to properties located elsewhere within the same area. Despite the heightened flood risk in the past two decades, we do not find evidence of divergent trends. Second, we analyze the effects of the recent reforms to the NFIP. In 2012 and 2014, Congress passed legislation announcing important increases in insurance premiums and flood map updates. We find robust evidence of large price reductions for properties that were drawn into the flood zone of the new FEMA flood maps. We estimate that, as a result of the mandatory insurance requirement in the flood zone, NFIP insurance costs for such properties in Virginia Beach will increase by an average of about $3,500 per year and lead to a reduction in housing values of about $64,000.

Introduction

Rising sea levels have increased the risk of coastal flooding over the past few decades, becoming the most economically damaging effect of climate change for many coastal locations (Buchanan, Oppenheimer, and Kopp, 2017). The increasing flood risk has been shown to arise from increased likelihood of tidal flooding in low-lying areas (Cooper et al., 2013; Sweet and Park, 2014; Sweet et al., 2014), increased frequency and severity of coastal flooding (Vitousek et al., 2017), and intensification of extreme flooding events (Hunter, 2010; Park et al., 2011; Tebaldi, Strauss, and Zervas, 2012).

At the same time, the population in shoreline areas in the United States has substantially increased (Hauer, Evans, and Mishra, 2016). One factor that may have contributed to this trend is the availability of subsidized flood insurance through the National Flood Insurance Program (NFIP; Conte and Kelly, 2017). The program, managed by the Federal Emergency Management Agency (FEMA), was created through the 1968 National Flood Insurance Act and has since risen to more than 5 million policyholders in 2016. Over the past two decades, NFIP has become burdened with enormous amounts of debt following hurricanes Katrina (2005) and Sandy (2012), which prompted a profound reform in 2012 and 2014, respectively. A major goal of the reform was to increase insurance rates to reflect the increase in flood risk between now and the inception of the program. As a result, flood insurance premiums increased substantially for many floodplain properties.[1]

The main goal of this article is to examine whether flood insurance affects the housing markets of coastal areas. NFIP is the key program at the disposal of the U.S. Government to try to influence individuals' housing location decisions in areas close to bodies of water. In the context of rising sea levels, assessing whether the key elements of NFIP—that is, the new flood maps and the resulting changed schedule of insurance premiums—are effective in shaping location decisions is important.

To carry out the analysis, we assembled a parcel-level dataset including the universe of residential sales for two coastal urban areas in the United States—Miami-Dade County (2008–15) and Virginia Beach (2000–16)—matched with their FEMA flood maps, which characterize the flood-risk level for each property.[2] Parts of these areas are prone to flooding events, and flood insurance is a first-order policy issue in both.

The first question we ask is whether there has been any divergence in trends regarding housing values and sales activity among properties under high risk of flooding, as defined by NFIP, relative to properties located elsewhere within the same area. Evidence of divergence in trends would suggest that housing markets have already been incorporating the increases in coastal flooding risk. We find that these variables behaved very similarly in the floodplain of these coastal areas, relative to the rest of the respective areas, and mostly traced the economic cycle. The lack of evidence of a generalized

[1] Flood insurance premiums vary as a function of flood risk and elevation. To understand the consequences of flood insurance reform on housing values in a simple manner, it is helpful to consider an example based on Kousky (2017). Based on the 2012 reform, the new flood insurance rate corresponding to a single-family home within the 100-year floodplain and 4 feet below the base flood elevation increases from $2,644 to $10,263. Given that the prereform rates did not vary by elevation, however, some properties sitting above the base flood elevation may experience a decrease in premiums. Properties outside the 100-year floodplain pay much lower rates. As a result, the median premium across all single-family policies is $512.

[2] Lacking a national sales registry, obtaining data from local sources and homogenizing individual sales data is a time-consuming endeavor.

differential trajectory for housing values on the floodplain of the coastal areas we analyze may be explained by the role of flood insurance. After all, we would not expect an increase in flood risk to have a large impact on housing values in a setting where properties are fully insured and insurance premiums remain constant over time.

Second, we focus on the recent reforms to NFIP in order to identify the effects of flood insurance on the housing market. In 2012 and 2014, Congress passed legislation to consider the updated (higher) estimated flood risk and ensure the financial viability of the flood insurance program. We focus on two important features of the reform: the increase in flood insurance premiums scheduled to take place from July 2012 onward and the release of updated flood maps, which impose changes to a set of properties that are now subject to mandatory flood insurance requirements. We hypothesize that the increases in premiums mandated by the reforms should have negatively affected the values of properties located in the floodplain, many of which are subject to mandatory flood insurance requirements. Specifically, we ask whether the increase in premiums affecting properties sold after July 2012 affected sales prices. We conduct a difference-in-difference analysis on the change in sales prices around the date of the reform. Our estimates do not reveal a differential effect from the enactment of the Biggert-Waters Act on previously established floodplain properties for the areas of Virginia Beach and Miami-Dade. In addition, we study the changes in the values of properties whose idiosyncratic flood risk has been updated in the new flood map. In the case of Virginia Beach, where updated maps came into effect in 2015, we find clear evidence of a large reduction in sales prices (of 25–28 log points) for properties that were drawn into the new flood maps. Hence, this finding is evidence that flood risk revisions affect housing values within 2 to 3 years, primarily through the effect on insurance premiums. As grandfathering of premium subsidies is phased out, these effects are likely to become more salient.

The rest of this article is organized as follows. The next section presents a brief review of the related literature. The following section provides background information on the history of flood insurance and recent policy changes. Then, we describe the main data sources and the merging process, followed by a section with descriptive statistics on the main variables. Next, we study overall trends of housing prices in flood risk areas relative to the overall area. The next two sections present the main results, followed by our conclusion. Some details of the legislation are relegated to an appendix.

Literature

A large literature estimates the price discount on sales of houses within the 100-year floodplain (Atreya, Ferreira, and Kriesel, 2013; Bin, Kruse, and Landry, 2008; Bin and Landry, 2013; Harrison, Bin, and Polasky, 2004; Smersh and Schwartz, 2001). Most studies use highly localized data and find evidence of a discount, but the magnitude ranges widely (from about 4 percent to 12 percent). In recent studies, Zhang (2016) and Zhang and Leonard (2018) estimated spatial quantile regression models using data for the Fargo-Moorhead metropolitan statistical area. They found evidence of a 3- to 6-percent floodplain penalty, larger among lower-priced homes. The authors argued that this heterogeneous effect may reflect the existence of upper bounds on coverage limits. These studies also show that the penalty rose sharply following a large flooding event in 2009 but fell back to its normal

value 2 years later. Compared with these studies, our dataset is much larger, and we focus on the role of flood insurance rather than the evolution of housing markets in the aftermath of large storms or flooding episodes.

Several recent papers also study flood risk and its implications. Gallagher (2014) studied how economic agents update flood risk beliefs after floods occur, as measured by insurance take-up rates. Using a national dataset with information on all flood insurance policies and whether a community is hit by a flood, he found strong evidence of an immediate increase in the fraction of homeowners covered by flood insurance in the affected communities, although the effect vanishes after a few years. Conte and Kelly (2017) documented that the distribution of aggregate hurricane damage in the United States is fat-tailed. They argued that this may be because the distribution of coastal property is also fat-tailed; that is, storms tend to intersect geographic areas with little or no property but also intersect areas with large amounts of property more often than expected, given a normal distribution. The authors then went on to present a model of homeowner behavior for buying insurance, which includes a disaster relief agency that reimburses a fraction of household losses, and showed that moral hazard[3] leads to an increase in the number of households living on the coast and in the size (value) of these properties, resulting in more damage during catastrophic storms.

Our analysis is particularly related to four recent studies. Bakkensen and Barrage (2017) estimated that coastal housing prices exceed fundamentals[4] by about 10 percent. Their analysis suggests that residents of flood-prone areas may have below-average fear of flooding, which might explain the sluggishness in the response of home prices in flood-prone areas to new information of increased flood risk.

Our use of changes to the flood maps as a source of identification is reminiscent of Votsis and Perrels (2016). These authors studied the effects of flood risk on housing values in Finland, exploiting the introduction of flood maps following the European Union Water Directive. These flood risk maps were made publicly available in 2006–07 for several flood-prone areas in Finland. The authors employed a difference-in-difference approach to capture the price differential of flood risk disclosure. They defined the treatment group as those dwellings located in the flood-prone area and the control group as nearby dwellings outside the flood-prone area. They found a significant price drop after the flood risk maps were released, with values ranging from 6 percent to 30 percent.

Recent work on the relationship between flood risk and housing values has been concentrated on the aftermath of Hurricane Sandy in New York in 2012. Ortega and Taşpınar (2018) argued that the hurricane triggered an increase in the perceived risk of living in the flood zone, persistently lowering housing values. Relatedly, Gibson, Mullins, and Hill (2019) studied the response of New York's housing prices to what they considered three flood risk signals:

[3] Moral hazard is defined as the behavior of increasing exposure to risk while insured, especially when the increased risk is taken because someone other than the policy holder bears the cost of those risks.
[4] Here, "fundamentals" is understood as the basic quantitative and qualitative information that contributes to the financial valuation of a house—for example, number of rooms, year built, neighborhood location, and so on.

(1) the Biggert-Waters Flood Insurance Reform Act; (2) Hurricane Sandy; and (3) revised FEMA flood maps. This study shares with ours the goal of estimating the effects of the Biggert-Waters Act on the housing market. Their analysis solely concerns New York City, however.[5]

Flood Insurance

Background: A Brief History of Flood Insurance

Congress created the National Flood Insurance Program (NFIP) in 1968, which is administered by Federal Emergency Management Agency (FEMA), with the goal of providing affordable flood insurance to homeowners. An integral part of the program is the Flood Insurance Rate Map (FIRM), which establishes risk zones. These zones determine flood risk for each property, building code requirements, and flood insurance requirements. Importantly, properties on the high-risk zone, also known as the 100-year floodplain, are required by law to purchase flood insurance if they have federally backed mortgages (or if they have received FEMA assistance in the past).

Largely because of Hurricane Katrina, NFIP accumulated a large amount of debt—more than $25 billion. To make the program financially stable, Congress passed the Biggert-Waters Flood Insurance Reform Act in 2012, which planned to eliminate subsidies to flood insurance rates and phase out several exemptions. However, as a result of vigorous public opposition in affected areas, Congress passed the 2014 Homeowner Flood Insurance Affordability Act. This act repealed some mandates of the 2012 act and slowed down some rate increases (as of April 1, 2015).[6]

Premium Increases

The 2012 Biggert-Waters Act stated that policies for properties purchased or newly insured after July 2012 would have to pay full-risk rates. The act also mandated annual premium increases of 25 percent until rates reflect full risk for structures with severe repeated flood losses as well as for nonprimary residences and businesses (GAO, 2013). The 2014 act slowed down the premium increases to 18 percent per year but did not repeal them.

New Flood Zone Maps

In addition to the aforementioned changes in flood insurance rates, revised FEMA flood maps were commissioned by the 2012 act. The release of these maps differs across areas of the United States. For instance, the new flood zone map for Virginia Beach went into effect on January 2015, whereas the flood map for Miami-Dade has not yet been revised.[7]

Flood maps are created by FEMA in partnership with states and local communities. Hydrologic and hydraulic studies determine ground elevations, as well as the depth of floodwaters, the amount of water that will be carried during flood events, the width of floodplains, and obstruction of water flows. If FEMA determines that a flood map is required, it consults with state and local officials and

[5] A previous version of our article (Indaco, Ortega, and Taşpınar, 2018) included analysis of the effects of NFIP reform on New York City's housing market. We concluded that the close temporal overlap between Hurricane Sandy and the Biggert-Waters Act made it impossible to discern the effects of the two events. As a result, the current version of the article focuses solely on Virginia Beach and Miami-Dade, which were affected by the flood insurance reform but not by Hurricane Sandy. We were not aware of the study by Gibson, Mullins, and Hill (2019) at the time of conducting our analysis, and it seems that the authors of that study were not aware of our article either. Both studies seem to have been carried out simultaneously.

[6] The appendix contains further details.

[7] A new flood map for New York City was released in early 2013 but has not become effective as of early 2019. Thus, political and affordability concerns can affect the timing of the adoption of revised flood maps.

issues a preliminary map. Local authorities can propose amendments or appeals that are considered before issuing the definitive map and making it publicly available.

The Homeowner Flood Insurance Affordability Act of 2014 directs FEMA to notify members of Congress when constituents in their district are affected by a flood mapping update. Typically, members of Congress then disseminate the news to the homeowners in their districts. In addition, in many states, sellers are required to disclose flood risk to potential buyers.

Data

Our dataset merges two different types of data. The first type of data is the universe of sales in the areas of Virginia Beach and Miami-Dade. These datasets have a similar structure for the two areas, but the coverage in terms of years and the actual property characteristics varies somewhat. The second type of data that we use are the FIRMs produced by FEMA for each of these two areas. These data contain shape files that allow us to overlay the tax parcel from the sales dataset onto the map of the flood risk zones.

The resulting dataset contains all sales that took place for these two areas, along with the flood risk class for each of the properties involved.

Housing Sales Data

We obtained the sales data for Virginia Beach from the Virginia Beach Real Estate Assessor's Office. The dataset contains the sales price and date for the last three transactions for every property in Virginia Beach. Approximately 160,000 properties included in this dataset have been sold at least once. The data contain all transactions between January 2000 and September 2016. We merged the sales data with another dataset provided by the area that contains the precise FEMA code for each property both for the previous map of 2009 as well as the most recent map that came into effect in 2015. The merged dataset contains data on 72,289 properties and 120,551 transactions realized after January 2000.

Our data for Miami-Dade County is based on the Florida Department of Revenue tax data files. We obtained and merged the Sales-Data-Files (SDF), the Name-Address-Legal (NAL), and the Name-Address-Personal (NAP) files. The merged dataset contains detailed information on parcels, including land use code, sales price, sales year, sales month, and assessment value, as well as on the geographic characteristics of the parcels. Our sample corresponds to sales recorded from 2008 to 2015, with a total of 356,672 observations. The flood zone information for the parcels is based on the FEMA flood insurance risk maps (effective since 2009).

Flood Insurance Rate Maps

We obtained the shape files for the FEMA flood insurance maps for our two urban areas of interest. These maps partition the areas into three areas: the 100-year floodplain (also known as Special Flood Hazard Areas), the 500-year floodplain, and the rest of the area. Among these, the highest risk of flooding pertains to the 100-year floodplain.[8] In the past few years, FEMA has produced revised flood maps. The release of the revised maps has been staggered, and for some areas, the new maps

[8] The flood maps include a finer partition of the set of properties, which can then be aggregated. As reported by FEMA (https://www.fema.gov/flood-zones), the 100-year floodplain is the union of the following categories: A, AO, AH, A1-A30, AE, A99, AR, AR/AE, AR/AO, AR/A1–A30, AR/A, V, VE, and V1–V30. The 500-year floodplain (not including the 100-year floodplain) is the union of category B properties and the X-shaded category. Finally, the minimal flood risk zone is the rest of the area, defined as the union of category C properties and the X-unshaded category.

are still under construction. As of early 2019, there exist revised maps for Virginia Beach but not for Miami-Dade. We provide more details on these maps in a later section.

Descriptive Statistics

In order to get rid of entries affected by typos in the sales price, and to eliminate sales of tiny parcels or other transactions that do not refer to proper housing units, we restrict to sales with sales prices above $10,000 and trim observations at the 99th percentile of the sales price. Exhibit 1 provides information on the distribution of sales by housing type in each of our two areas. Our dataset for Virginia Beach (the data cover the 2000–16 period) contains 120,542 sales, of which 3.9 percent correspond to properties located on the floodplain. We see a higher share of single-family and two-family homes on the floodplain (80 percent) than in the rest of the area (74 percent). The data for Miami-Dade cover the 2008–15 period and contain 220,303 sales. In Miami-Dade, the floodplain accounts for a much larger share of sales (51 percent) than for Virginia Beach, and the share of sales corresponding to single-family and two-family homes is similar on the floodplain (67 percent) and in the rest of the area (68 percent).

Exhibit 1

Frequency of Sales by Type of Property (Land Use)

	freq	freq	freq	rel. freq.	rel. freq.	rel. freq.
	All	FP	**Outside FP**	All	FP	**Outside FP**
Virginia Beach 2000–2016						
Single-Family to Two-Family	89,275	3,713	85,562	0.74	0.80	0.74
Multifamily (More Than 2)	173	12	161	<0.01	<0.01	<0.01
Other	31,094	917	30,177	0.26	0.20	0.26
Total	120,542	4,642	115,900	1.00	1.00	1.00
Miami-Dade 2008–2015						
Single-Family to Two-Family	149,118	63,120	86,101	0.68	0.67	0.68
Multifamily (More Than 2)	14,172	6,063	8,126	0.06	0.06	0.06
Other	57,013	25,507	31,593	0.26	0.27	0.25
Total	220,303	94,690	125,820	1.00	1.00	1.00

FP = floodplain.
Notes: Floodplains are defined based on the FEMA flood insurance rate maps. The Virginia Beach and Miami-Dade flood maps are the 2009 versions.
Source: Authors' calculations, using data from the Virginia Beach Real Estate Assessor's Office and the Florida Department of Revenue

We now turn to the timing of sales, which is summarized in exhibit 2. This exhibit only reports sales pertaining to single-family homes in order to focus on a more homogeneous subsample that can be better compared within and across areas. The first three columns refer to Virginia Beach. On average, 199 single-family homes on the floodplain sold, constituting 4 percent of the areawide sales, a rate that has fluctuated very little during the 2000–16 period. The last three columns of the exhibit refer to Miami-Dade. On average, 7,890 single-family homes located on the floodplain sold annually, amounting to 42 percent of areawide sales.[9] In sum, those data suggest that sales activity on the floodplain has kept the same pace as in the rest of the respective areas during our period of analysis.

[9] The frequency of sales for 2015 in Miami-Dade County in our dataset is roughly one-half the size of the number of sales in 2014. That is because the data are missing the sales corresponding to the last two quarters of 2015. Note, however, that the share of sales corresponding to floodplain properties is 43 percent, in line with the share for the previous years.

Exhibit 2

Frequency of Sales by Year: Single-Family Homes Only

Year	VB FP	VB Outside FP	VB FP Share (%)	MD FP	MD Outside FP	MD FP Share (%)
2000	173	4,831	0.03			
2001	213	6,210	0.03			
2002	205	6,376	0.03			
2003	201	5,754	0.03			
2004	299	7,219	0.04			
2005	256	6,304	0.04			
2006	175	5,108	0.03			
2007	204	4,888	0.04			
2008	140	3,821	0.04	4,406	6,182	0.42
2009	174	4,132	0.04	5,917	8,330	0.42
2010	162	4,145	0.04	8,491	11,786	0.42
2011	188	4,059	0.04	7,544	10,333	0.42
2012	209	4,287	0.05	8,980	12,090	0.43
2013	218	4,716	0.04	11,000	14,792	0.43
2014	186	4,478	0.04	10,958	14,552	0.43
2015	224	4,843	0.04	5,824	8,036	0.42
2016	161	3,592	0.04			
Avg.	199	4,986	0.04	7,890	10,763	0.42

FP = floodplain. MD = Miami-Dade. VB = Virginia Beach
Notes: Floodplains are defined based on the FEMA flood insurance rate maps. The Virginia Beach and Miami-Dade flood maps are the 2009 versions. Sales in 2016 for Virginia Beach do not include the data for the last quarter of that year. Likewise, sales for Miami-Dade in 2015 do not include data for the last quarter of that year.
Source: Authors' calculations, using data from the Virginia Beach Real Estate Assessor's Office and the Florida Department of Revenue

Let us now turn to the evolution of median sales prices. Exhibit 3 reports median sales prices for single-family homes. The average sales price for a single-family home in Virginia Beach is substantially higher on the floodplain than elsewhere in the area ($296,000 versus $229,000) by about 29 percent. Column 3 shows that this ratio has also fluctuated substantially between 2000 and 2016 but without a clear trend. The last three columns in the exhibit report the data for Miami-Dade County. In this case, median sales prices are very similar inside and outside the floodplain ($186,000 versus $182,000). The floodplain differential has remained constant at around 3 percent in the 2008–15 period. In sum, we do not observe a reduction in the floodplain price differential associated with houses located on the floodplain taking place from 2013 onward.

Trends in Floodplain Premiums

Housing values in an area's floodplain may be higher or lower than in the rest of the area. On one hand, proximity to the coast has an amenity value; on the other, it exposes those properties to risk of flooding. Regardless of whether the floodplain price differential is positive or negative, the increased

Exhibit 3

Median Sales Price by Year: Single-Family Homes Only

Year	VB FP	VB Outside FP	VB FP/Outside (%)	MD FP	MD Outside FP	MD FP/Outside (%)
2000	172,500	138,000	1.25			
2001	148,000	145,000	1.02			
2002	202,140	159,900	1.26			
2003	225,000	175,500	1.28			
2004	309,900	210,000	1.48			
2005	381,750	265,000	1.44			
2006	399,000	284,900	1.40			
2007	355,950	285,000	1.25			
2008	300,000	267,885	1.12	256,500	255,000	1.01
2009	257,000	250,000	1.03	165,000	165,000	1.00
2010	299,375	250,000	1.20	141,100	140,000	1.01
2011	290,041	240,000	1.21	149,250	145,200	1.03
2012	310,000	245,000	1.27	160,000	154,000	1.04
2013	295,000	245,000	1.20	185,000	180,000	1.03
2014	377,000	247,638	1.52	203,905	199,000	1.02
2015	341,000	250,000	1.36	224,250	220,000	1.02
2016	307,500	259,000	1.19			
Avg.	296,346	229,347	1.29	185,626	182,275	1.02

FP = floodplain. MD = Miami-Dade. VB = Virginia Beach.
Notes: Floodplains are defined based on the FEMA flood insurance rate maps. The Virginia Beach and Miami-Dade flood maps are the 2009 versions. Sales in 2016 for Virginia Beach do not include the data for the last quarter of that year. Likewise, sales for Miami-Dade in 2015 do not include data for the last quarter of that year.
Source: Authors' calculations, using data from the Virginia Beach Real Estate Assessor's Office and the Florida Department of Revenue

flood risk associated with rising sea levels might have reduced the floodplain price premium if it had been incorporated into the beliefs of market agents. One needs to keep in mind, however, that flood insurance may dampen these effects. If insurance premiums remain unchanged, floodplain housing prices may not respond to increased flooding risk. Therefore, whether the rise in risk of flooding has already affected housing markets is an empirical question.

Specifically, this section compares the trends in sales prices and sales activity for properties located on the floodplain of each of the two areas relative to properties located elsewhere within the same area. Importantly, in doing so, we will also account for individual-property or neighborhood-level heterogeneity.

Sales Activity

We begin by analyzing sales activity. One option would be to compare annual counts of sales in the floodplain to sales in the rest of the area. The same goal can be attained, however, while accounting for property-level heterogeneity. Specifically, we build a balanced property-level panel where each

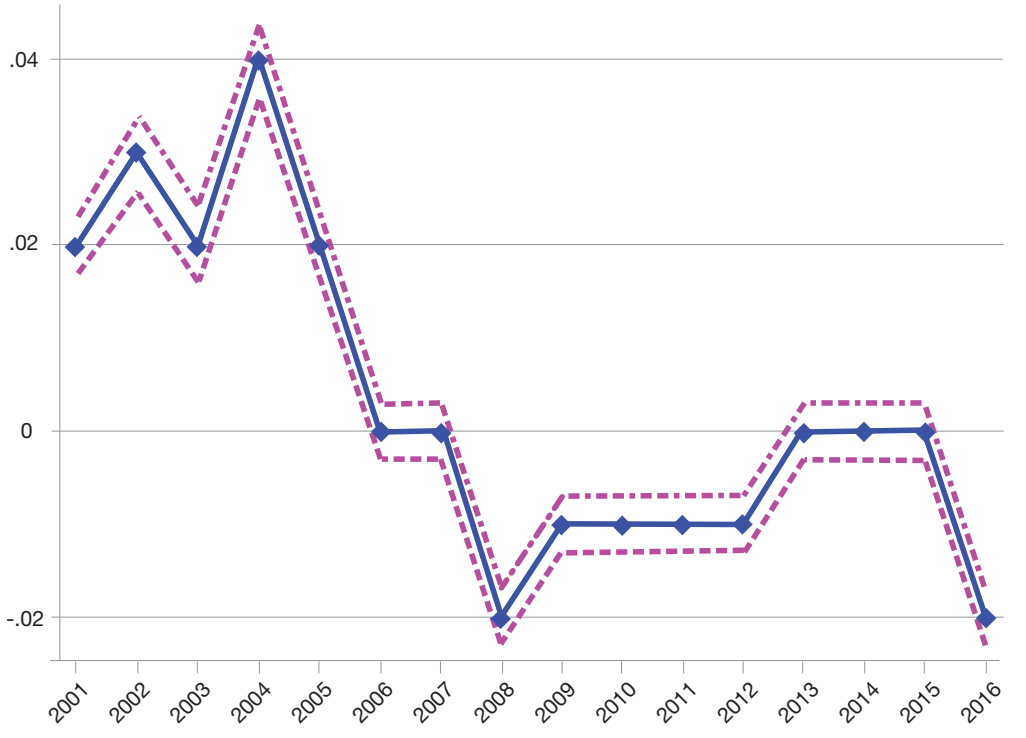
property i is observed in each year.[10] Our data contain the universe of property sales in each of the areas. Hence, we can define an indicator $Sold_{it}$ for whether property i that takes a value of 1 in years when that property was sold and zero otherwise. Then we estimate a linear probability model:

$$Sold_{it} = \alpha_i + \lambda_t + \beta_t FP_i + \varepsilon_{it}, \quad (1)$$

where we are including property fixed-effects (α_i), capturing all time-invariant characteristics of the property, in addition to year fixed-effects (λ_t) that capture the evolution over time of the sales probability outside the floodplain. Time-varying β_t captures the differential in the sales probability associated with being located in the 100-year floodplain (FP).[11] The probability that a property is sold in a given year can be interpreted as a measure of selling activity based on the share of the stock of housing that is sold in any given year.

Exhibit 4a

Virginia Beach, Sales Probability: Areawide



Source: Authors' calculations, using data from the Virginia Beach Real Estate Assessor's Office

[10] For each of the two areas, we obtained a complete roster of all residential properties, identified by their tax identifier. We build a balanced panel at the annual level for all these properties. Because we know all sales over our sample period, we can then create a zero-one sales indicator for each property in each year.

[11] In practice, $\beta t\, FPi$ is shorthand notation for the sum over all periods of interactions between the floodplain indicator, FPi, and the corresponding year dummy, Dt.

We begin with the results for Virginia Beach, shown in exhibit 4a and exhibit 4b. The former figure shows an areawide large and monotonic drop in sales between 2004 and 2008, which then plateaus between 2009 and 2011. Since 2011, sales have been slowly increasing but are still far below the 2004 level. Given that data for 2016 are only through September, the probability of sales for the year is somewhat inferior to that of previous years. This evolution clearly reflects the economic cycle—see also column 1 in exhibit 5. Turning now to the floodplain sales differential depicted in exhibit 4b, we observe an initial drop of the floodplain penalty in 2001 of 1 percentage point, relative to the reference year of 2000, that remains below the reference year until 2012. The floodplain differential eventually oscillates around its year-2000 value between 2012 and 2016.

Exhibit 4b

Virginia Beach, Sales Probability: Floodplain Differential

Notes: Point estimates from the event study at annual frequency on the sample of single-family homes. We are plotting the coefficients of the year dummies (top) and interactions with 100-year floodplain, or FP100 (bottom). Regressions include property fixed-effects, and standard errors are clustered at the block level, which allows for arbitrary correlations across parcels in the same block and over time.
Source: Authors' calculations, using data from the Virginia Beach Real Estate Assessor's Office

Exhibit 5

Estimation Floodplain Differential

Dep. Var.	Virginia Beach 1 Sold	Virginia Beach 2 lnp	Miami-Dade 3 Sold	Miami-Dade 4 lnp
FP x D2001	-0.01	-0.50***		
	[0.008]	[0.101]		
FP x D2002	-0.02	-0.29***		
	[0.009]	[0.081]		
FP x D2003	-0.03***	-0.23**		
	[0.009]	[0.090]		
FP x D2004	-0.02***	-0.04		
	[0.007]	[0.069]		
FP x D2005	-0.01	-0.01		
	[0.009]	[0.061]		
FP x D2006	-0.01	-0.03		
	[0.008]	[0.057]		
FP x D2007	-0.02***	-0.04		
	[0.007]	[0.061]		
FP x D2008	-0.01	0.03		
	[0.007]	[0.061]		
FP x D2009	-0.01	0.00	-0.000	0.052*
	[0.007]	[0.061]	[0.004]	[0.027]
FP x D2010	-0.01	-0.04	-0.000	0.004
	[0.008]	[0.060]	[0.003]	[0.023]
FP x D2011	-0.01	-0.02	0.002	0.020
	[0.006]	[0.069]	[0.004]	[0.024]
FP x D2012	0.00	-0.02	0.005	0.025
	[0.006]	[0.074]	[0.005]	[0.024]
FP x D2013	0.01	-0.08	0.004	0.013
	[0.008]	[0.071]	[0.003]	[0.023]
FP x D2014	0.01	-0.10	0.008**	0.010
	[0.008]	[0.082]	[0.004]	[0.023]
FP x D2015	-0.01	-0.01	0.001	0.025
	[0.008]	[0.077]	[0.003]	[0.025]
FP x D2016	-0.00	0.00		
	[0.007]	[0.053]		
Observations	1,063,928	89,275	815,104	131,230
Number of blocks	54,268	54,268	101,888	101,915
Fixed-effects	parcel	parcel	parcel	parcel
Clustering s.e.	block	block	block	block

Dep. Var. = dependent variable. FP = floodplain. lnp = log of sales price. s.e. = standard errors.
****p < 0.01, **p < 0.05, *p < 0.1*
Notes: All models include property fixed-effects. Sample contains sales of single-family homes only. The sales probability models are estimated on a balanced panel, where each parcel appears in every year. The sample periods are 2000–16 for Virginia Beach and 2008–15 for Miami-Dade County. The coefficients referring to the first year in each area's sample are dropped to avoid perfect collinearity. Standard errors clustered at the block level, which allows for arbitrary correlations across parcels in the same block and over time.
Source: Authors' calculations, using data from the Virginia Beach Real Estate Assessor's Office and the Florida Department of Revenue

The graphs for Miami-Dade County also confirm this pattern. Sales activity clearly traces the economic cycle (exhibit 6a), and sales probabilities for floodplain properties appear to increase slightly during our sample period before going back to their initial value in the reference year 2008 (exhibit 6b).[12]

Exhibit 6a

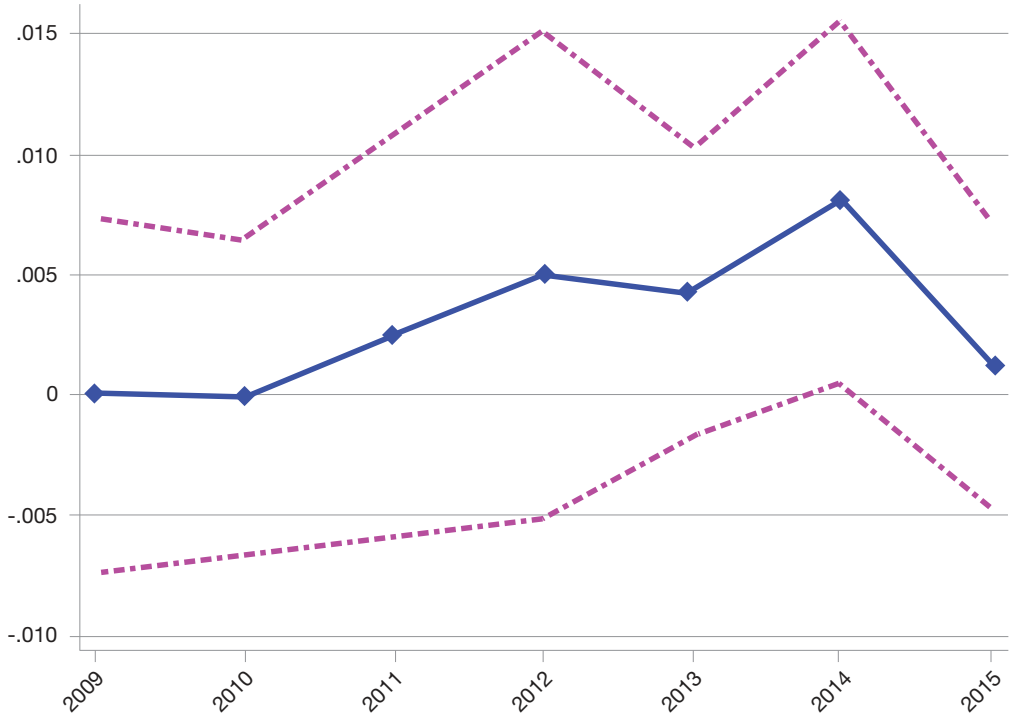
Miami-Dade, Sales Probability: Areawide

Note: Exhibit plots the point estimates along with the corresponding 95-percent confidence interval.
Source: Authors' calculations, using data from the Florida Department of Revenue

[12] The estimates are collected in column 6 of exhibit 5. Note also the sharp drop in exhibit 6a corresponding to 2015. This reflects the fact that our data for 2015 are missing the sales for the last two quarters. However, this seems to affect similarly properties on and off the floodplain. Thus, it probably does not bias the estimated floodplain differentials plotted in exhibit 6b.

Exhibit 6b

Miami-Dade, Sales Probability: Floodplain Differential



Notes: Point estimates from the event study at annual frequency on the sample of single-family homes. We are plotting the coefficients of the year dummies (top) and interactions with 100-year floodplain, or FP100 (bottom). Regressions include property fixed-effects, and standard errors are clustered at the block level, which allows for arbitrary correlations across parcels in the same block and over time.
Source: Authors' calculations, using data from the Florida Department of Revenue

In conclusion, our analysis shows that sales in each area have been largely driven by the housing cycle, and we do not find systematic evidence of a differential evolution of sales activity in the floodplain of the areas we have analyzed.

Sales Prices

Next, we turn to the analysis of sales prices. Let p_{it} denote the sales price of property i in year t. We assume the log of the sales price is given by

$$In\ p_{it} = \alpha_i + \lambda_t + \beta_t FP_i + \varepsilon_{it}, \qquad (2)$$

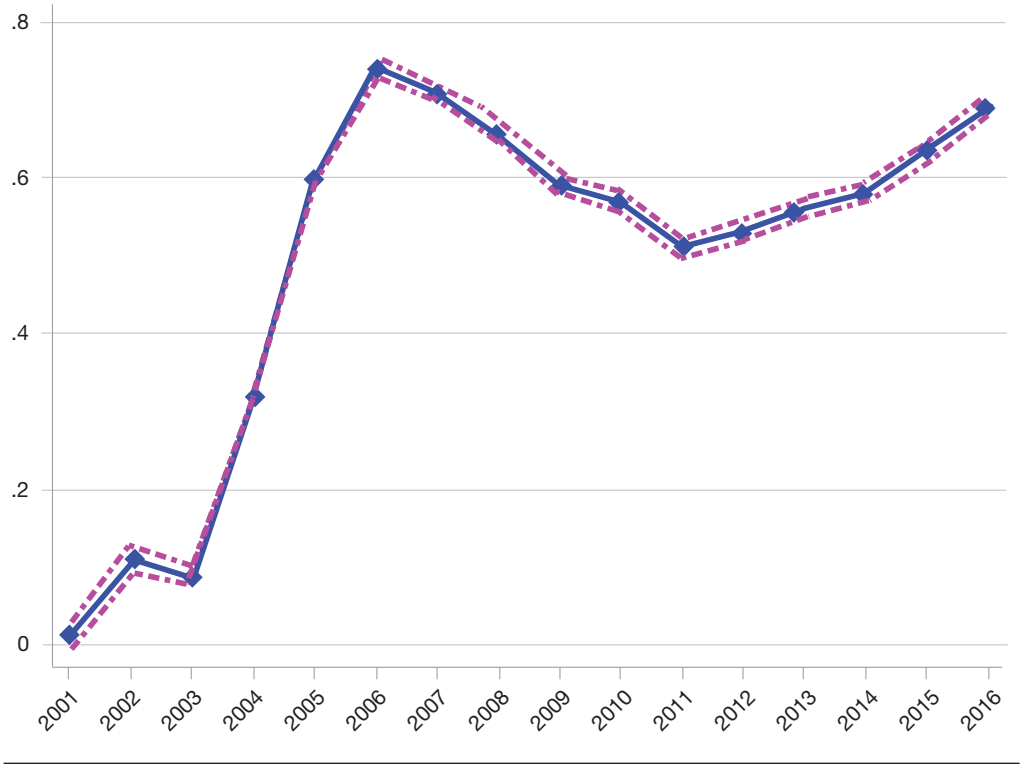
where FP_i is an indicator for being located on the 100-year floodplain (as defined at the beginning of the sample period). The point estimates for year dummies $\{\lambda_t\}$ trace out the average housing prices (for single-family homes) in the area, while $\{\beta_t\}$ captures the floodplain price differential. Clearly,

because of the property fixed-effects, properties that are sold only once do not contribute to the identification of the coefficients. In effect, our sample consists of properties that were sold repeatedly during our sample period.[13]

We report the time-varying coefficients graphically, beginning with the results for Virginia Beach. As shown in exhibit 7a, the housing cycle was also clearly noticeable in this area, with prices peaking in 2006, falling until 2011, and recovering since then. More relevant for our purposes, exhibit 7b suggests that the floodplain price differential remained stable around the 2000 level between 2004 and 2012, when the Biggert-Waters Act was implemented. The detailed estimates for $\{\hat{\beta}_t\}$ can be found in column 2 in exhibit 5. Exhibit 8a illustrates the results for Miami-Dade. As seen in the figure, sales prices outside the floodplain remained depressed between 2009 and 2011 and then increased rapidly from 2012 onward. But, once again, we see no evidence of a differential trend in sales prices on the floodplain relative to that of the rest of the area in exhibit 8b. The estimates for $\{\hat{\beta}_t\}$ for Miami-Dade can be found in column 4 in exhibit 5.

Exhibit 7a

Virginia Beach, Sales Price: Areawide



Note: Exhibit plots the point estimates along with the corresponding 95-percent confidence interval.
Source: Authors' calculations, using data from the Virginia Beach Real Estate Assessor's Office

[13] The point estimates are very similar if we use area-block fixed-effects in place of property fixed-effects. The reason is that houses in the same neighborhood (defined as an area or census block) tend to share many characteristics, such as construction year and size, because they were developed around the same time.

Exhibit 7b

Virginia Beach, Sales Price: Floodplain Differential

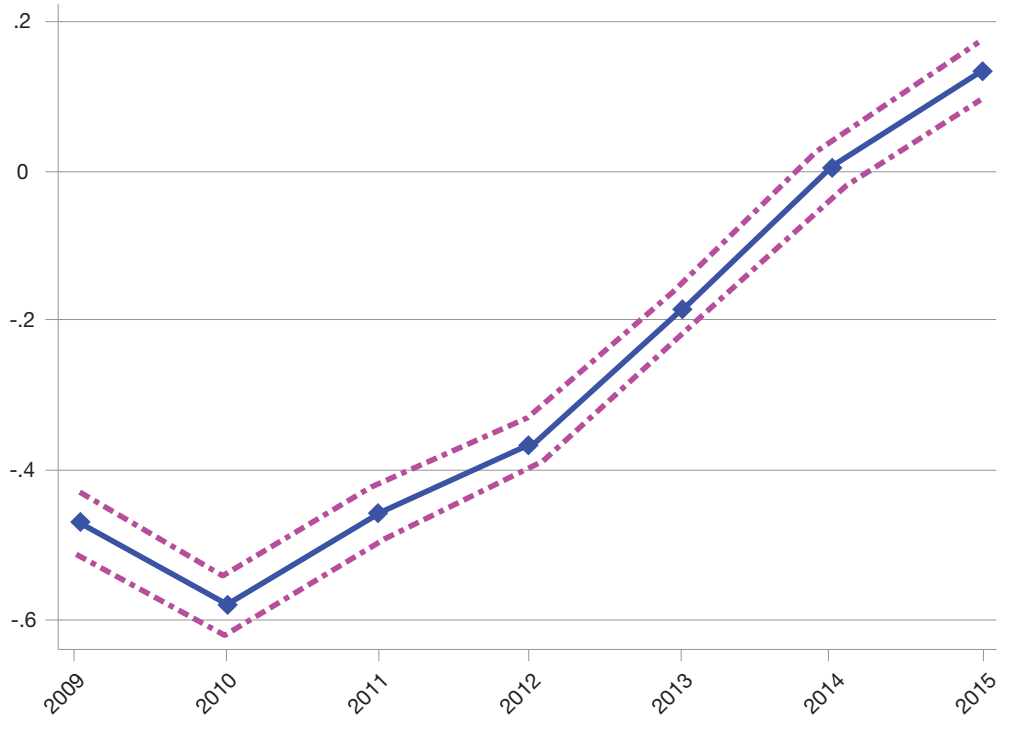
Notes: Point estimates from the event study at annual frequency on the sample of single-family homes. We are plotting the coefficients of the year dummies (top) and interactions with 100-year floodplain, or FP100 (bottom). Regressions include property fixed-effects, and standard errors are clustered at the block level, which allows for arbitrary correlations across parcels in the same block and over time.
Source: Authors' calculations, using data from the Virginia Beach Real Estate Assessor's Office

Summing up, our analysis in this section suggests that sales activity and housing prices have been strongly influenced by the economic cycle. This has been true both for properties located on the floodplain as elsewhere in the area. But, in general, we find very little evidence of differential trends for the floodplain areas. Our interpretation of this finding, particularly for the years prior to the 2012–14 flood insurance reform, is that flood insurance has cushioned the impact of increased flooding risk on housing values. The reason is that premiums remained subsidized and largely unchanged during that period, despite the increased evidence of higher risk.[14] The trends analysis in this section can also be used to assess the assumption that causal interpretation of difference-in-difference estimates requires the identifying assumption of a lack of differential pretreatment trends for flood zone properties. Given that sales activities and sales prices evolved in a very similar fashion in and outside the flood zones in these areas, we understand that a causal interpretation of the estimates for Virginia Beach and Miami-Dade County is justified.

[14] This interpretation is consistent with the findings in Bosker et al. (2018). Using detailed data for the Netherlands, where flood insurance is unavailable, these authors estimate a 1-percent price penalty for residential properties located in the flood zone. The availability of flood insurance in the United States may have been enough to shrink the flood zone price penalty down to almost zero.

Exhibit 8a

Miami-Dade, Sales Price: Areawide



Note: Exhibit plots the point estimates, along with the corresponding 95-percent confidence interval.
Source: Authors' calculations, using data from the Florida Department of Revenue

Flood Insurance Reform: Premium Increases

Difference-in-Difference Estimation

As argued earlier, the 2012 flood reform ushered in increases in flood insurance premiums that may have affected some sales that took place after July 2012. The 2014 act attenuated these increases (to 18 percent per year) but did not eliminate them. As a result, it is natural to consider a difference-in-difference estimator around that date, comparing properties in the floodplain to properties located elsewhere in the area. In order to more cleanly identify the effects of the policy change, we follow Hallstrom and Smith (2005) (who were interested in the effects of flood risk) and focus our analysis on the repeat sales subsample that brackets July 2012.[15] We hypothesize that the price increases between two such sales for properties on the floodplain will be smaller than the price increase experienced by properties located elsewhere in the area. It is helpful to think of our treatment group as containing houses located on the floodplain that were sold before and after

[15] For houses that were sold more than once after July 2012, we keep only the first sale and drop the later ones. This corresponds to 241 sales for Virginia Beach and 6,178 for Miami-Dade.

Exhibit 8b

Miami-Dade, Sales Price: Floodplain Differential

Notes: Point estimates from the event study at annual frequency on the sample of single-family homes. We are plotting the coefficients of the year dummies (top) and interactions with 100-year floodplain, or FP100 (bottom). Regressions include property fixed-effects, and standard errors are clustered at the block level, which allows for arbitrary correlations across parcels in the same block and over time. Exhibit plots the point estimates, along with the corresponding 95-percent confidence interval.
Source: Authors' calculations, using data from the Florida Department of Revenue

July 2012. The control group includes repeat sales that happened before July 2012 and repeat sales that may straddle across 2012 but refer to properties located outside the floodplain.[16]

More specifically, we define a new indicator, $PostBW_t$ (for Biggert-Waters Act), for sales that took place after July 2012, and FP_i denotes the floodplain dummy as before. We then pose that the log of the sales price for property i in year t is given by

$$ln\, p_{it} = \alpha_i + \lambda_t + \beta_t PostBW_t \times FP_i + \varepsilon_{it}, \quad (3)$$

where coefficient β identifies the effects of the 2012 Biggert-Waters Act (effective from July 2012) on houses in the floodplain. We expect this coefficient to be negative because of the increase in flood insurance for houses located on the floodplain and estimate the model separately for each of our two coastal areas.

[16] We can experiment with different variations of the control group, for example, restricting only to repeat sales that bracket July 2012. In that case, the control group is only made up of the repeat sales outside the floodplain.

As before, we restrict to the sample of single-family homes. The data for Virginia Beach contain 89,275 sales for single-family units between January 2000 and September 2016. Roughly 31 percent of those correspond to properties that have been sold only once in this period and are thus dropped. In our sample period, roughly 40 percent of properties were sold exactly twice, and roughly 29 percent were sold at least three times.[17] For Miami-Dade, our data contain 131,099 sales for single-family homes from 2008 to 2015. Approximately 76 percent of the sales refer to properties that were sold only once in this period. Among properties sold more than once, about 20 percent of the sales correspond to properties that were sold twice, and about 2.7 percent were sold three times.[18]

Let us now turn to our estimation results. The key coefficient is the one associated with the interaction between the post-reform and the floodplain indicators, denoted by β. This coefficient is identified by the change in price for houses that were sold both before and after the reform and are located on the floodplain, the net of the price changes between other pairs of sales. Exhibit 9 collects the estimates. We begin with the results for Virginia Beach. In column 1, the point estimate of the coefficient of interest is positive but not statistically different from zero. In the case of Miami-Dade (column 2), it is negative, but, again, we are unable to reject the zero-null hypothesis. Taken together, these estimates suggest that the premium changes associated with the 2012 act have not had, for the time being, any effect on sales prices for properties located on the floodplain on account of the data for the two areas in our sample.

Exhibit 9

Difference-in-Difference Estimator for the Log Sales Price Around July 2012, Repeat Sales Only

Dep. Var. lnp	Virginia Beach 1	Miami-Dade 2
FP x PostBWA	0.04	-0.008***
	[0.023]	[0.020]
Observations	61,281	49,290
R-Squared	0.329	0.222
Number of Parcels	26,428	23,808
Year-Month Dummies	yes	yes
Sample	2000–16m9	2008–15

BWA = Biggert-Waters Act. FP = floodplain. Dep. Var. lnp = Dependent variable is log of sales price.
****p < 0.01, **p < 0.05, *p < 0.1*
Notes: All models include property fixed-effects. PostBWA is an indicator taking a value of one for sales that took place after July 2012. Standard errors clustered at the block level, which allows for arbitrary correlations across parcels in the same block and over time.
Source: Authors' calculations, using data from the Virginia Beach Real Estate Assessor's Office and the Florida Department of Revenue

[17] Recall that the data provided to us by the Virginia Beach Government Real Estate Assessor's Office include only the last three sales for each household.
[18] The remaining 3 percent are units sold four to six times.

Flood Insurance: New Flood Maps

The 2012 Biggert-Waters Act called for updating the flood maps across the country. This is a staggered process. Whereas new maps for Virginia Beach were completed and put in effect in 2015, the revision of the flood map for Miami-Dade is still under construction, with an unknown release date.

Interesting for our purposes, the new flood maps typically change the risk classification for some properties. For instance, when the floodplain is expanded, some units that were located outside the floodplain under the old map now find themselves located inside the new floodplain because of the increased risk of flooding. This has potentially large consequences in terms of flood insurance. First, these properties are now subject to the mandatory flood insurance requirements. In addition, the premiums these properties face are typically substantially higher, in line with the new risk classification. It is important to keep in mind that the 2012 and 2014 reforms allow for "grandfathering." Namely, properties that were already insured can keep their old (lower) premiums. When the property is sold (or the insurance policy lapses), however, the applicable premiums will be those that correspond to the new risk classification.

It is also worth noting that property reclassifications may not always purely reflect an increase in flood risk. Pralle (2019) analyzed the politics behind flood map revisions and documented several instances where FEMA's revisions were likely influenced by local political concerns, rather than being solely driven by objective risk assessments. Nonetheless, a change in a property's risk classification that draws that property in or out of a flood zone will typically have an impact on the flood insurance premium paid by the owner, which should be expected to capitalize in the market value of the property.

Our goal in this section is to exploit the variation in the risk category of the properties that were reclassified in the new flood maps. We hypothesize that the prices of the houses subject to risk reclassification may have been affected by the release of relevant information.

FEMA Flood Maps

FEMA flood maps for Virginia Beach were revised in both 2009 and 2015.[19] The 2015 map happens to contain fewer buildings in the flood zone than its predecessor.[20] Exhibit 10 summarizes the changes in risk classification. In terms of sales, our data contain 520 cases of properties that were not part of the 100-year floodplain under the 2009 map but have been included in the newly defined floodplain. At the same time, 683 sales correspond to properties that were previously in the floodplain but are now excluded from it. The corresponding numbers for single-family homes are 485 and 534, respectively.

[19] The link lists the FEMA flood map revisions scheduled for 2015: https://www.servicelinknationalflood.com/Content/pdfs/2015_Revision_List.pdf.
[20] Individual properties or whole developments can request changes in risk classification to FEMA by providing their own technical assessments and by adopting protective measures.

Exhibit 10

Transition Matrix Flood Zones: Sales Counts: Virginia Beach

Dep. Var. Inp	All Properties Counts	All Properties Shares	Single-Family Counts	Single-Family Shares
Out-Out	115,389	0.957	85,077	0.953
FP-FP	3,959	0.033	3,179	0.036
Out-FP	520	0.004	485	0.005
FP-Out	683	0.006	534	0.006
Total	**120,551**	**1.000**	**89,275**	**1.000**

Dep. Var. Inp = Dependent variable is log of sales price. FP = floodplain.
Note: A property is in the (100-year) floodplain if it has risk classification A, AE, AH, or VE in the FEMA flood map. All these categories are subject to mandatory flood insurance requirements. The specific definitions for the flood zone designations are given in the appendix. The flood maps for Virginia Beach are dated 2009 and 2015.
Source: Authors' calculations, using data from the Virginia Beach Real Estate Assessor's Office

Thus, many individual homeowners received an update on their properties' idiosyncratic flood risk. Clearly, the properties that were not in the flood zone under the old map but were classified as belonging to the flood zone under the new map received the largest and most relevant amount of information. We will exploit this in the next section.

Regression Analysis

In order to analyze the effects of a change in the flood risk zone designation, we restrict to sales of properties located in the floodplain under the old map or the new one (or both). Accordingly, we build indicators for sales corresponding to properties located within the floodplain under both maps ($\boldsymbol{InIn} = \mathbf{1}$), for sales of properties that were initially outside the floodplain but are within it under the new map ($\boldsymbol{OutIn} = \mathbf{1}$), and for sales of properties that were initially within the floodplain but are excluded from it under the new map ($\boldsymbol{InOut} = \mathbf{1}$). The resulting sample size (in terms of sales) is 5,162 (4,198 of which are single-family homes).

As before, we consider regression models for both sales activity and sales prices. The former analysis is based on a balanced panel, and the dependent variable is an indicator for whether a given property i was sold in year t:

$$Sold_{it} = \alpha_i + \lambda_t + \beta_{OI} Post_t \times OutIn_i + \beta_{IO} Post_t \times InOut_i + \varepsilon_{it}. \qquad (4)$$

We note that property fixed-effects and year dummies are included in the regression, and in the omitted category are properties that were located on the floodplain in both maps ($\boldsymbol{InIn} = \mathbf{1}$). The "Post" indicator in the regression for Virginia Beach takes a value of 1 for 2015.

In terms of sales prices, we consider the following (property) fixed-effects specification:

$$In\ p_{it} = \alpha_i + \lambda_t + \beta_{OI} Post_t \times OutIn_i + \beta_{IO} Post_t \times InOut_i + \varepsilon_{it}, \qquad (5)$$

where the dependent variable is the log of the sales price of property i and quarter-year t.

Our hypothesis is that $\beta_{OI} < 0 < \beta_{IO}$, indicating that properties that are being assigned a higher flood risk and are subject to the mandatory insurance requirement should experience a drop in price, while properties whose risk is revised downward should experience a price increase. The predictions in terms of the sales probability are less clear. An increase in the risk assessment could lead to either an increase or a decrease in the probability of being sold, depending on the speed of adjustment of the property's price.

The results are collected in exhibit 11. Column 1 reports the estimates regarding the sales probability. The point estimates for both β_{OI} and β_{IO} are negative and statistically significant. That is, the sales probability for houses experiencing a change in their risk assessment (in either direction) seems to have fallen relative to houses that were in the floodplain under both the old and new maps. In terms of sales prices (columns 2 and 3), we find clear evidence of a large reduction in sales prices (of 25–28 log points) for properties that were drawn into the new flood maps, confirming our expectation. At the mean of our sample, this effect entails a $64,000 reduction in the value of a single-family home.[21]

Exhibit 11

Effects of Risk Reclassification: Virginia Beach

Dep. Var.	1 Sold	2 ln p	3 ln p
Post x Out-In	-0.01*	-0.25***	-0.28***
	[0.006]	[0.076]	[0.078]
Post x In-Out	-0.01**	-0.07***	0.01
	[0.004]	[0.085]	[0.074]
Observations	62,033	5,162	4,198
R-Squared	0.004	0.376	0.407
Number of Parcels	3,380	3,380	2,736
Building Class	All	All	Fam1

Dep. Var. = dependent variable. Fam1 = Single-family. ln p = log of sales price.
****p < 0.01, **p < 0.05, *p < 0.1*
Notes: The "Post" indicator in the regression for Virginia Beach takes a value of 1 for 2016. Property fixed-effects are included in all models. The sample only includes sales referring to properties that are located on the floodplain under the old or new flood maps (or both). The omitted category in the regression models consists of sales of properties located on the floodplain under both the old and new maps. Recall that in column 1, the dataset is a balanced panel, with zero values for property-period cells if the property was not sold. All columns include year dummies. Standard errors are clustered at the area-block level in all columns.
Source: Authors' calculations, using data from the Virginia Beach Real Estate Assessor's Office

Conclusion

The National Flood Insurance Program (NFIP) is the key program at the disposal of the U.S. government to shape the actions of individuals regarding the decision to live in flood-prone areas. In the current context of rising sea levels, it is important to assess the effectiveness of this policy tool. The mandatory insurance requirement derived from the FEMA flood maps and the corresponding premiums are the elements of NFIP that more directly affect consumers through their effect on housing values. This article analyzes the effectiveness of these two aspects of NFIP based on the recent changes to the program, using data for two important urban coastal areas: Virginia Beach and Miami-Dade County.

It is important to distinguish the effects of NFIP reforms from those of objective flood risk. At some level, the flood maps and flood insurance premiums are connected to estimates of flood risk at a high geographic granularity. Undoubtedly, however, other factors also play a role, such as the decision to update flood maps or allow grandfathering rules, or mayors' decisions to appeal a new flood map (Pralle, 2019). These factors can drive a wedge between objective flood risk and market values. From this viewpoint, the usefulness of our analysis to identify the effects of increased flood risk may be limited.

Our first finding is that property sales and sales prices have evolved similarly in the floodplains of Virginia Beach and Miami-Dade and elsewhere in these areas, largely tracing the economic cycle. In a context of increasing flood risk, we interpret this parallel evolution as driven by the availability of flood insurance at constant and subsidized rates up until 2012, which may have cushioned the impact of increased risk on housing values. First, homeowners also may not be fully aware of flood risk, homeowners less concerned about flooding risk may sort into the flood zone (Bakkensen and Barrage, 2017), and the effective difference in flood insurance costs in and out of the flood zone may have been relatively small in the prereform period.

Second, we focus on two specific features of the 2012–14 flood insurance reform to try to identify the effects of flood insurance on the housing market. First, we ask whether the scheduled increase in insurance premiums affecting properties sold after July 2012 affected sales prices. The estimates suggest that has not been the case. Because of the gradual implementation schedule of the new rates, these effects may only materialize over a longer time span. The current high degree of heterogeneity in flood insurance premiums among properties within the flood zone may also attenuate the estimates.

Third, the reform also mandated that FEMA update the flood maps of all areas, which is a staggered process. New maps were publicly released and adopted for Virginia Beach in 2015. The new maps provide an additional source of identification. Specifically, we focus on the

[21] It is also interesting to quantify the increase in flood insurance costs for properties that are drawn into the flood zone. We obtained insurance quotes from The Flood Insurance Agency for a single-family house in Virginia Beach (flood zone A, AE, or AH) with a replacement value of $250,000, content coverage of $100,000, and a $5,000 deductible. The resulting annual premium is $3,506. Using a 5-percent interest rate, the present discounted value of such a payment amounts to $70,120. This is remarkably close to the $64,000 price reduction that we estimated.

properties that have experienced an upward or downward revision in their idiosyncratic flood risk classification. We found a large and robust negative effect on the prices of properties that have suffered an upward revision in flood risk in Virginia Beach (of 25–28 log points). We estimate that, as a result of the mandatory insurance requirement in the flood zone, NFIP insurance costs for these properties will increase by an average of about $3,500 per year and lead to a reduction in housing values of about $64,000.

These findings lead us to draw the following conclusions. First, we believe that the more gradual pace of reform adopted in 2014 has delayed the effects of the more ambitious 2012 NFIP reform on the housing market. We expect flood-zone housing values to reflect the increases in flood insurance premiums over the medium term. However, our analysis has shown that the adoption of new flood maps has a large and immediate effect on the values of properties that are drawn into the new flood zone, probably due to the mandatory flood insurance requirement now affecting these properties.

All in all, our findings reveal that flood insurance reform has the potential to affect housing values and, hence, shape households' location decisions. However, the implementation of the reforms has been slowed down by political factors. At a national level, the Great Recession played an important role in the decision to adopt a more gradual increase in flood insurance premiums in 2014. Likewise, areas like New York City appealed the new flood maps made public in 2013, citing affordability concerns, delaying its implementation for several years.

Appendix

The Flood Insurance Reform Acts in Detail

The 2012 Biggert-Waters Act (BWA)

With BWA, the high-risk zones are expanded to reflect current risk levels. Regarding insurance rates, there are increases for some properties (described in the following). Although rates will not change immediately for about 70 percent of the properties (because of grandfathering rules), once those houses are sold (or experience massive flood damage), the new rates will kick in.

Specific implications:

- Policies will be written or renewed at full-risk rates for properties purchased or newly insured after July 5, 2012.

From October 1, 2013:

- Subsidized policies for (1) non-primary residences, (2) businesses, and (3) structures with severe repeated flood losses will have annual premium increases of 25 percent until rates reflect full risk.

- Primary residences in Special Flood Hazard Areas (SFHAs) can keep their subsidized rates until (1) the property is sold, (2) the policy lapses, (3) the property suffers severe flood losses (and the owner refuses an offer to mitigate), or (4) a new policy is purchased.
- Phasing out of grandfathered rates over 5 years.[22]
- Establishment of a Reserve Fund to help cover costs when claims exceed the annual premiums collected by the National Flood Insurance Program (NFIP).

The 2014 Homeowner Flood Insurance Affordability Act (HFIAA)

This act repealed some mandates of the 2012 act, slowed down some rate increases, and amended most provisions mandating that certain policies transition immediately to full-risk rates.

Specific implications from April 1, 2015:

- Limits increases for individual premiums to 18 percent per year.
- Limits increases for average rate classes to 15 percent per year.
- Increases the Reserve Fund assessments required by BWA.
- Implements annual surcharges on all new and renewed policies. This is meant to compensate for the loss in revenue from slowing down BWA rate increases. These surcharges will be collected until all subsidies are eliminated. The surcharges are flat fees applied to all policies regardless of the flood zone where the building is located. The annual surcharge for primary residences is $25, and for non-primary residences or non-residential properties, it is $250.
- Introduces cost-saving flood insurance coverage for properties newly mapped into the SFHA. They will receive PRP (Preferred Risk Policy) rates for 1 year after the maps become effective. The rates at renewal will increase no more than 18 percent per year.
- Resuscitates grandfathering. Grandfathering remains an option for policyholders when new maps show their buildings in a higher risk area (for example, zone A to zone V or increase in the Base Flood Elevation, or BFE). Available to property owners who (1) have flood insurance policies in effect when the new flood maps become effective, or (2) have built-in compliance with the Flood Insurance Rate Map (FIRM) in effect at the time of construction. These policyholders have the option of using the flood zone on a previous FIRM that was in effect when the building was originally constructed (for those built in compliance) or when coverage was first obtained (for those with continuous coverage).

[22] Prior to BWA, when revised maps showed higher risk zones, policyholders were permitted to grandfather and use the zone and elevation of an older map.

Flood Zone Designations

Zone A is the flood insurance rate zone determined by approximate methods, as no BFEs are available for these areas. Mandatory flood insurance purchase requirements apply, that is, properties with a federally backed mortgage are required to purchase flood insurance.

Zone AE is the flood insurance rate zone that corresponds with flood depths greater than 3 feet. Mandatory flood insurance purchase requirements apply.

Zone AH is the flood insurance rate zone that corresponds to areas of shallow flooding with average depths between 1 and 3 feet. Mandatory flood insurance purchase requirements apply.

Zone VE is the flood insurance rate zone that corresponds to coastal areas that have additional hazards associated with storm waves. Mandatory flood insurance requirements apply.

Zone X and Zone X-500 are flood insurance rate zones that are outside the floodplain or with average flood depths of less than 1 foot. For these properties, it is not mandatory to purchase flood insurance.

Acknowledgments

We thank PSC-CUNY (Award TRADA-47-538) for funding.

Authors

Agustín Indaco is an assistant teaching professor of economics at Carnegie Mellon University in Qatar.

Francesc Ortega, corresponding author, is the Dina Axelard Perry Professor in Economics in the Department of Economics at Queens College CUNY and can be reached at fortega@qc.cuny.edu.

Süleyman Taşpınar is an assistant professor in the Department of Economics at Queens College CUNY.

References

Atreya, Ajita, Susana Ferreira, and Warren Kriesel. 2013. "Forgetting the Flood? An Analysis of the Flood Risk Discount over Time," *Land Economics* 89 (4): 577–596.

Bakkensen, Laura A., and Lint Barrage. 2017. Flood Risk Belief Heterogeneity and Coastal Home Price Dynamics: Going Under Water? NBER Working Paper No. 23854. Cambridge, MA: National Bureau of Economic Research.

Bin, Okmyung, Jamie Brown Kruse, and Craig Landry. 2008. "Flood Hazards, Insurance Rates, and Amenities: Evidence From the Coastal Housing Market," *Journal of Risk & Insurance* 75 (1): 63–82.

Bin, Okmyung, and Craig Landry. 2013. "Changes in Implicit Flood Risk Premiums: Empirical Evidence from the Housing Market," *Journal of Environmental Economics and Management* 65 (3): 361–376.

Bin, Okmyung, and Stephen Polasky. 2004. "Effects of Flood Hazards on Property Values: Evidence Before and After Hurricane Floyd," *Land Economics* 80 (4): 490–500.

Bosker, Maarten, Harry Garretsen, Gerard Marlet, and Clemens van Woerkens. 2018. "Nether Lands: Evidence on the Price and Perception of Rare Natural Disasters," *Journal of the European Economic Association* 17 (2): 413–453.

Buchanan, Maya K., Michael Oppenheimer, and Robert E. Kopp. 2017. "Amplification of Flood Frequencies with Local Sea Level Rise and Emerging Flood Regimes," *Environmental Research Letters* 12 (6): 064009.

Conte, Marc N., and David L. Kelly. 2017. "An Imperfect Storm: Fat-Tailed Tropical Cyclone Damages, Insurance, and Climate Policy," *Journal of Environmental Economics and Management* 92: 677–706.

Cooper, Hannah M., Charles H. Fletcher, Qi Chen, and Matthew M. Barbee. 2013. "Sea-Level Rise Vulnerability Mapping for Adaptation Decisions Using LiDAR DEMs," *Progress in Physical Geography* 37 (6): 745–766.

Gallagher, Justin. 2014. "Learning About an Infrequent Event: Evidence from Flood Insurance Take-Up in the United States," *American Economic Journal: Applied Economics* 6 (3): 206–233.

Gibson, Matthew, Jamie T. Mullins, and Alison Hill. 2019. Climate Risk and Beliefs: Evidence from New York Floodplains. Working paper 2019-02. Williamstown, MA: Williams College, Department of Economics.

Government Accountability Office (GAO). 2013. *Flood Insurance: More Information Needed on Subsidized Properties. Report to Congressional Committees* 13-607. Washington, DC: United States Government Accountability Office.

Hallstrom, Daniel G., and V. Kerry Smith. 2005. "Market Responses to Hurricanes," *Journal of Environmental Economics and Management* 50 (3): 541–561.

Harrison, David M., Greg T. Smersh, and Arthur L. Schwartz, Jr. 2001. "Environmental Determinants of Housing Prices: The Impact of Flood Zone Status," *Journal of Real Estate Research* 21 (1/2): 3–20.

Hauer, Mathew E., Jason M. Evans, and Deepak R. Mishra. 2016. "Millions Projected To Be at Risk From Sea-Level Rise in the Continental United States," *Nature Climate Change* 6: 691–695.

Hunter, John. 2010. "Estimating Sea-Level Extremes Under Conditions of Uncertain Sea-Level Rise," *Climatic Change* 99 (3): 331–350.

Indaco, Agustín, Francesc Ortega, and Süleyman Taşpınar. 2018. The Effects of Flood Insurance on Housing Markets. IZA discussion paper 11810. Bonn, Germany: Institute for the Study of Labor (IZA).

Kousky, Carolyn. 2018. "Financing Flood Losses: A Discussion of the National Flood Insurance Program," *Risk Management and Insurance Review* 21 (1): 11–32.

Ortega, Francesc, and Süleyman Taşpınar. 2018. "Rising Sea Levels and Sinking Property Values: Hurricane Sandy and New York's Housing," *Journal of Urban Economics* 106: 81–100.

Park, Joseph, Jayantha Obeysekera, Michelle Irizarry, Jenifer Barnes, Paul Trimble, and Winifred Park-Said. 2011. "Storm Surge Projections and Implications for Water Management in South Florida," *Climatic Change* 107 (1): 109–128.

Pralle, Sarah. 2019. "Drawing Lines: FEMA and the Politics of Mapping Flood Zones," *Climatic Change* 152 (2): 227–237.

Sweet, William V., and Joseph Park. 2014. "From the Extreme to the Mean: Acceleration and Tipping Points of Coastal Inundation from Sea Level Rise," *Earth's Future* 2 (12): 579–600.

Sweet, William, Joseph Park, John Marra, Chris Zervas, and Stephen Gill. 2014. *The Geography of Development: Evaluating Migration Restrictions and Coastal Flooding*. NOAA Technical Report NOS CO-OPS 073. Silver Spring, MD: National Oceanic and Atmospheric Administration.

Tebaldi, Claudia, Benjamin H. Strauss, and Chris E. Zervas. 2012. "Modelling Sea Level Rise Impacts on Storm Surges Along US Coasts," *Environmental Research Letters* 7 (1): 014032.

Vitousek, Sean, Patrick L. Barnard, Charles H. Fletcher, Neil Frazer, Li Erikson, and Curt D. Storlazzi. 2017. "Doubling of Coastal Flooding Frequency Within Decades due to Sea-Level Rise," *Scientific Reports* 7 (1): 1399.

Votsis, Athanasios, and Adriaan Perrels. 2016. "Housing Prices and the Public Disclosure of Flood Risk: A Difference-in-Differences Analysis in Finland," *Journal of Real Estate Finance and Economics* 53 (4): 450–471.

Zhang, Lei. 2016. "Flood Hazards Impact on Neighborhood House Prices: A Spatial Quantile Regression Analysis," *Regional Science and Urban Economics* 60 (Supplement C): 12–19.

Zhang, Lei, and Tammy Leonard. 2018. "Flood Hazards Impact on Neighborhood House Prices," *Journal of Real Estate Finance and Economics*: 1–19.

Departments

In this issue—

- *Affordable Design*
- *Data Shop*
- *Graphic Detail*
- *Impact*
- *Policy Briefs*
- *SpAM*

Affordable Design

The U.S. Department of Housing and Urban Development sponsors or cosponsors three annual competitions for innovation in affordable design: The Innovation in Affordable Housing Student Design and Planning Competition; the American Institute of Architects – HUD Secretary's Housing Community Design Awards; and the HUD Secretary's Opportunity & Empowerment Award, co-sponsored with the American Planning Association. This Cityscape department reports on the competitions and their winners. Each competition seeks to identify and develop new, forward-looking planning and design solutions for expanding or preserving affordable housing. Professional jurors determine the outcome of these competitions.

Annual HUD Secretary's Opportunity and Empowerment Award: Family Scholar House, Louisville, Kentucky

Regina C. Gray
U.S. Department of Housing and Urban Development

For the past two decades, HUD's Office of Policy Development and Research (PD&R) and the American Planning Association (APA) have partnered to recognize planning development projects that have truly transformed the community into one that is livable and resilient. The HUD Secretary's Opportunity and Empowerment Award recognizes communities that have implemented creative and effective strategies to improve the quality of life for low- and moderate-income residents. The award emphasizes how creative housing, economic development, and private investments can be used in or with a comprehensive community development plan to build social equity and empower individuals and families who reside in that community. Nominees must demonstrate how they respond effectively to those challenges while simultaneously and effectively addressing community needs; this is done by expanding affordable housing and transportation options and enhancing economic competitiveness through sound and equitable development practices. PD&R recently completed a review of the nominations for the annual HUD Secretary's Opportunity and Empowerment Award. The jury received six applications, deliberated in late fall, and recommended the Family Scholar House (FSH) in Louisville, Kentucky, as this year's award recipient.

FSH was selected, in large part, because the HUD-APA jury was particularly impressed with the wide array of programs offered to area residents and their families to help lift them from poverty. FSH is a nonprofit organization that offers programs that range from helping single mothers find

permanent employment to providing education and job opportunities for youth transitioning from foster care. In the past year, 504 families have taken advantage of the various educational and career-oriented services that are offered by FSH, and of that number, 215 families currently reside in one of the five FSH multifamily properties.[1]

The Parkland Family Scholar House in Louisville, Kentucky.

Founded in 1995, the program has a broader mission to end the cycle of poverty and transform the community by empowering families to achieve financial independence. The FSH model has expanded to help participants locate low-cost housing, provide transportation and childcare options, and offer job training and opportunities for the underemployed. The program has added a variety of supportive services ranging from mental health assessments to food assistance and a host of financial services around credit repair and asset building.

The success story of the FSH centers on the continued commitment from the program's public and private partners. FSH and Marian Development established an impressively diverse funding portfolio for the $10.5 million program, successfully leveraging low-income housing tax credits, HOME investment funds, Community Development Block Grant funds, and foundational support. An Affordable Housing Trust Fund, along with state and federal historical tax credits, were used to develop the five FSH multifamily housing developments in Louisville and one affiliate location in Covington, Kentucky. The partnership implemented an ambitious funding campaign to raise an

[1] According to the most recent annual report from FSH, supportive services are available for adults and children, particularly single parents, who are looking for opportunities to complete their education, seek better jobs, and programs that help stabilize housing conditions for themselves and their children. The annual report is accessible at the following link: https://familyscholarhouse.org/wp-content/uploads/2019/01/by-the-numbers.pdf.

additional $1.2 million. The projects met several local planning goals, including the creation of an additional 324 affordable housing units across all its satellite properties, a robust homelessness prevention initiative, and onsite supportive services for residents.

The FSH program's success is further evidenced by the high percentage of participants who have exited the program to stable housing and suitable employment. According to the organization's most recent statistics,[2] for example, nearly 80 percent of participants have found permanent employment, and all have located stable, affordable housing. A recent *Washington Post* article noted that more than 500 families have lived in one of the FSH satellite centers over the course of the past decade. The average tenure is about 3 years, and every family has found permanent, stable housing (Kolodner, 2018).

The education center at the Louisville Family Scholar House location (pictured above) provides an array of educational and job training services.

Each of the five FSH developments uses an operating subsidy in the form of project-based Section 8 vouchers or Moving to Work housing choice vouchers. The jury found that these projects met several local planning goals related to the creation of affordable housing, economic and community development, homelessness prevention, and creation of specialized pre-residential and residential services.

[2] See the "By the Numbers" annual report released by FSH at the following link: https://familyscholarhouse.org/wp-content/uploads/2019/01/by-the-numbers.pdf.

One of the newer FSH centers is housed at Riverport Landings, also located in Louisville, and is an "intergenerational living" community that offers senior housing and services, childcare services, and housing for single moms.

To learn more about the HUD Secretary's Opportunity and Empowerment Award, go to https://www.huduser.gov/portal/publications/pdf/hud_583_2015.pdf.

Author

Regina C. Gray is a social science analyst in the Affordable Housing Research and Technology Division of the U.S. Department of Housing and Urban Development.

Reference

Kolodner, Meredith. 2018. "They Were Young. And Homeless. And Moms. Now, They Have College Degrees," *The Washington Post*, April 22. https://www.washingtonpost.com/local/education/they-were-young-and-homeless-and-moms-now-they-have-college-degrees/2018/04/22/7c9bb5ae-440f-11e8-bba2-0976a82b05a2_story.html?noredirect=on&utm_term=.4221df693aa2.

Data Shop

Data Shop, a department of Cityscape, presents short articles or notes on the uses of data in housing and urban research. Through this department, the Office of Policy Development and Research introduces readers to new and overlooked data sources and to improved techniques in using well-known data. The emphasis is on sources and methods that analysts can use in their own work. Researchers often run into knotty data problems involving data interpretation or manipulation that must be solved before a project can proceed, but they seldom get to focus in detail on the solutions to such problems. If you have an idea for an applied, data-centric note of no more than 3,000 words, please send a one-paragraph abstract to david.a.vandenbroucke@hud.gov for consideration.

Tracking Individuals Pre- and Post-Foreclosure

Christos Makridis
Massachusetts Institute of Technology (MIT) Sloan School of Management

Michael Ohlrogge
New York University School of Law

Disclaimer
All opinions expressed here are strictly those of the authors and do not reflect the views of affiliated organizations or the U.S. Government.

Abstract

In this article, the authors examine how it is possible to construct records that track 1.4 million households that experience foreclosure from their pre-foreclosure to post-foreclosure residences. These records were created by merging two powerful sets of data: county registrar of deeds records (licensed from CoreLogic, Inc.) and consumer mail marketing data (licensed from RefUSA). The article starts with a description of the county registrar of deeds data and how it can be used to create a dataset of mortgages and outcomes (including foreclosure). The authors proceed to describe the nature of the consumer marketing data from RefUSA, how it can be used to track households as they move locations, and then how that data can be merged with the mortgage records constructed from registrar of deeds data. The article also includes discussions of how those combined records can be merged with the Home Mortgage Disclosure Act (HMDA) database, using GIS software, to learn additional demographic information (income, race, and so on) about individuals with mortgages and foreclosures

Introduction

Residential mortgages and their foreclosures played a pivotal role in the recent financial crisis. A host of high-quality data sets exist with millions, billions, even trillions of observations on residential mortgage origination, performance, delinquency, and foreclosure. Those data sets have given rise to hundreds, if not thousands, of academic papers investigating which individuals got residential mortgages, on what terms, how those mortgages performed, and the implications of those findings for the broader economy. For all of this richness in investigation though, data limitations have made it difficult or impossible for researchers to accurately track where specific individuals move after foreclosure.

More than 6 million households and 10 million individuals experienced mortgage foreclosures from 2007 to 2013 (CoreLogic, Inc., 2017). Thus, understanding what happened to those people is important from macro- and microeconomic perspectives and for documenting the human impacts of one of the most significant financial events of the postwar period in the United States. In our companion article (Makridis and Ohlrogge, 2018), we take on one such effort to improve understanding of post-foreclosure outcomes by examining whether individuals move to areas with better or worse labor market opportunities after foreclosures and what predicts the significant heterogeneity among individuals in those outcomes.

In this article, we present details about our new method for constructing data that tracks more than 1.4 million households that experienced mortgage foreclosure, identifying the precise addresses at which they lived before and after the end of their mortgages. Our approach proceeds through several steps.

First, we use parcel-level data, licensed from CoreLogic, Inc. and gathered from county registrars of deeds to identify the precise addresses and dates for which ownership of a piece of land changes hands due to a completed foreclosure action. Second, we use the street addresses (house number, unit number [if applicable], street name, city, state, and ZIP Code) on these foreclosure records to match them to a consumer mail marketing database licensed from RefUSA. Third, we use the unique household identifiers obtained from RefUSA to track individuals to their next address after a foreclosure. Finally, to obtain additional information about the individuals who experience foreclosure (such as their income at time of mortgage origination, race, and so forth), we match the foreclosure records with data from the Home Mortgage Disclosure Act (HMDA).

To our knowledge, this is the first and only data set to be used in academic research that identifies individuals who have lost their houses to foreclosure and then shows where they move after that foreclosure. Furthermore, although our method depends on two types of data that must be licensed for a fee, such licensing is available to any researcher with the requisite budget, and for both types of data, multiple sources are available from which to license the data, thus increasing the options, availability, and affordability of building such data.

The remainder of this article proceeds as follows. In the next section, we review existing mortgage data sets and their limitations with respect to identifying outcomes of individuals post-foreclosure. We then describe how we constructed records of mortgage originations and outcomes from county

registrar of deeds data. In the following section, we describe how we merged this mortgage data with the RefUSA mail marketing database to identify where individuals move post-foreclosure. In the next section, we discuss how we obtained additional information about the individuals subject to foreclosure by merging our mortgage records with the HMDA database. In the final section, we discuss alternative vendors for the two key types of proprietary data we use in our data construction.

Limitations of Existing Mortgage Data Sets

The reason that studying the post-foreclosure outcomes for households has been difficult lies in the nature of how most mortgage data sets used for academic research are generated. Many mortgage data sets, both public and proprietary, are ultimately generated by mortgage servicers and distributed to mortgage investors.[1] After a mortgage foreclosure occurs, those who service or own mortgages have relatively little reason to track where individuals move.[2]

Another type of mortgage data set comes from government records of mortgage originations. The HMDA data set and the registrar of deeds data (described further herein) are examples; however, HMDA does not track any information past a mortgage's origination. Deed registries generally provide the names of individuals who take out mortgages and own land. Then in theory it might be possible to search such data by the names of individuals to discover where they move after foreclosure. First-and-last name combinations are far from unique across the country; in any case, those records will be a small and unrepresentative subsample of the population who are able to buy property again immediately following a mortgage foreclosure.

The only other data set similar to ours for tracking individuals post-foreclosure is that created by the Federal Reserve Bank of New York (FRBNY), using Equifax credit data (see Brevoort and Cooper, 2013; Lee and Van der Klaauw, 2010; Molloy and Shan, 2013). In addition to this data being proprietary to Federal Reserve researchers, it has three significant limitations compared with the data we construct. First, this data set identifies only individuals who have begun the foreclosure process. Given that roughly one-half of foreclosure starts never result in a completed foreclosure (see, for example, Federal Housing Finance Agency, 2018), this introduces a substantial amount of uncertainty to the data. Second, the FRBNY-Equifax data set identifies an individual's location only when a credit report is requested for that person; thus, if an individual moves post-foreclosure but then does not have any credit reports run, that person will seem to have not moved at all. Finally, we track roughly 1.4 million households that experienced foreclosure between 2006 and 2011, compared with only 330,000 households that experienced foreclosure between 2000 and 2009, as tracked in the FRBNY-Equifax data (Brevoort and Cooper, 2013).[3]

[1] Publicly available sets such as these include those from Fannie Mae, Freddie Mac, and the Columbia Collateral File. Private ones that must be licensed for a fee include those from firms such as BlackBox and CoreLogic, Inc.

[2] In particular, in a large number of states, antideficiency statutes explicitly prohibit mortgage owners from trying to collect from borrowers any amounts that remain unpaid on a mortgage loan after foreclosure of a house backing that loan. Even where those statutes do not apply, attempting to collect money beyond the value of the housing collateral from a foreclosed borrower is frequently not economically practicable.

[3] That 330,000 figure is only for foreclosure starts, thus representing perhaps one-half that number of completed foreclosures.

Building Mortgage Records from Registrar of Deeds Data

Throughout the United States, property law generally requires that an interest in land, such as an ownership or mortgage interest, be registered with the county in which the property is located if that interest is to be legally enforceable. Those registries are then made publicly available by counties so that, for instance, a person purchasing property can be assured that the seller has the legal right to dispose of that asset. Commercial companies, such as CoreLogic, Inc. and its predecessors, have in turn gathered this information from the roughly 3,000 U.S. counties, cleaned and standardized it, and licensed it for academic and other research purposes.

County registrars of deeds create unique parcel identification numbers by which records on individual properties can be traced over time.[4] We thus start by identifying in our data when each new loan is taken out on a property, whether that loan is for the initial purchase of the property, a refinance of a purchase loan, or a home equity line of credit. For each such loan, we then observe the set of subsequent records tied to the property's unique identifier.

If, for instance, we observe an initial purchase loan taken out on a property and no subsequent records on that property until 6 years later, at which point we observe a new (non-foreclosure) sale of that property from the prior owners to new owners, then we conclude that the initial loan ended in a prepayment. If a loan is a purchase loan or a refinance of a purchase loan, and we observe a subsequent loan of comparable size (adjusted for amortization) being taken out on the same property, with no interceding changes of that property's ownership, we likewise conclude that the loan was refinanced and thus ended in a prepayment. If, however, we observe a loan registered on a property, and then we observe the property next changing hands through a foreclosure transaction—with no intervening prepayment events—then we conclude that the given loan ended in a foreclosure.

As a practical matter, if our only objective were to track people from before and after foreclosure, our procedure here could be simpler. In particular, we would need to identify only the dates and addresses at which foreclosures occur and then match those addresses to the unique family IDs assigned to them by the RefUSA data, as described later in this article. We also, however, wish to match those loan records to the HDMA data to gain more information about the people whose homes are foreclosed on. We must therefore generate the more precise information on the loans that are being foreclosed on, such as what type of loan (purchase, refinance, home equity, and so on) it was, when the loan was originated, who the lender was, and so forth.

Out of a starting set of roughly 500 million property-level records in the deed data we license, we derive approximately 96 million unique mortgage records through these procedures. In general, multiple property-level records are associated with each mortgage (for instance, one or sometimes

[4] Some of the work of companies such as CoreLogic, Inc. also involves ensuring the uniqueness and continuity of these identifiers.

multiple records when the mortgage is taken out as well as when it is extinguished). We also exclude records on commercial properties, loans for construction and development of subdivisions and large groupings of residential properties, and other records not directly relevant for tracking residential mortgages to individual homeowners.[5]

Matching Deed Mortgage Records to Consumer Marketing Data

Having created mortgage records and identified individuals whose properties were foreclosed on, we next draw on the consumer mail marketing database licensed from RefUSA. RefUSA's primary clients are companies or charitable organizations that target individuals with advertisements and solicitations. The foundation for RefUSA's database comes from proprietary data they license from the U.S. Postal Service. That database incorporates information from individuals who fill out official "change of address" forms when they move, as well as a variety of other government forms through which individuals disclose new addresses. RefUSA then supplements that data through partnerships with credit card companies, magazines, utilities, and similar entities to which people voluntarily report changes in address. Any person who has moved and has subsequently seen catalogs and charitable solicitations follow them, unprompted, to their new address will be familiar with the ability of firms like RefUSA to track individuals, even without any deliberate actions by those individuals to disseminate their new addresses. RefUSA has compiled this information into a yearly panel of roughly 130 million U.S. households.

RefUSA assigns each household a unique identification number to track those households as they move to new locations. Each yearly entry in the panel includes a full street address. Thus, one can readily identify households that move by filtering for those that change street addresses from one year to the next. The specific data we license from RefUSA spans years 2006 to 2012; thus, we are able to identify households that move in years 2006 through 2011.

We match moving households in RefUSA with foreclosed households from the deed data on the basis of the exact correspondence of house number (which includes unit number, if applicable), street name, state, ZIP Code, and year of move or foreclosure. If a foreclosure occurs, for instance, in 2009, then depending on when a household's address is updated in the RefUSA data, they may not be recorded as moving until 2010. Thus, we allow a margin of 1 year when matching on the year of move or foreclosure.

In the mortgage data constructed from deed records, we identify 4,088,248 foreclosures occurring between 2006 and 2011. Of those, we are able to match 1,415,241 to records of households shown as moving in the RefUSA data, giving us a match rate of 34.6 percent. Our best understanding of the reason for this low match rate is that not all of the foreclosed households were initially in the

[5] Overall, this construction of mortgage histories out of these public record filings is similar to the approach taken by Ferreira and Gyourko (2015). See also Diamond and McQuade (2016) and Ohlrogge (2018) for additional uses of such public records deed data.

RefUSA database, and for those that were, RefUSA may have experienced delays before identifying and reporting them as having moved.[6]

A natural concern with a low match rate such as this is that the sample of foreclosures for which we identify where the household moved post-foreclosure may not be representative of the full set of foreclosures. We perform several tests to investigate whether this is likely a concern.[7]

First, we consider the initial mortgage amount (in dollars) for foreclosures that we match to the RefUSA moving records and for those that we did not. We construct a quantile-quantile (Q-Q) plot[8] comparing the quantiles of the two distributions. Exhibit 1 presents the results, showing nearly identical distributions between the two sets of foreclosures.[9] Thus, no systematic bias seems to exist, whereby households with smaller mortgages (presumably, with lower incomes) are less likely to be tracked to their post-foreclosure locations.

Next, we investigate whether geographic differences might be present, in which some areas are more likely to have individuals whom we are able to successfully track via RefUSA. We compute the fraction of total observations that each U.S. ZIP Code accounts for in the matched and total sample, finding a correlation between the two samples of 0.92. Although a higher match rate would undoubtedly be desirable, we are nevertheless able to identify a large number (1.4 million) of foreclosures for which we can track the post-foreclosure residence of the household, and no systematic differences seem to be present in either the individual or the geographic characteristics of the individuals for whom we are and are not able to match.

[6] In Makridis and Ohlrogge (2017), we examined another data product from RefUSA that advertised information on the number of people employed at each U.S. business, with updates on a yearly basis. In that paper, we demonstrated that those employment counts are subject to substantial interpolation that may limit the number of applications for which they can be used. By contrast, determining the address of an individual is far easier than determining the number of people employed by a business. As such, we are less concerned about data errors in RefUSA's consumer marketing database. Although such errors undoubtedly exist and unquestionably account for some portion of our failure to match foreclosure records, the tests we describe herein give us some confidence that match failures do not substantially reduce the usefulness of the data.

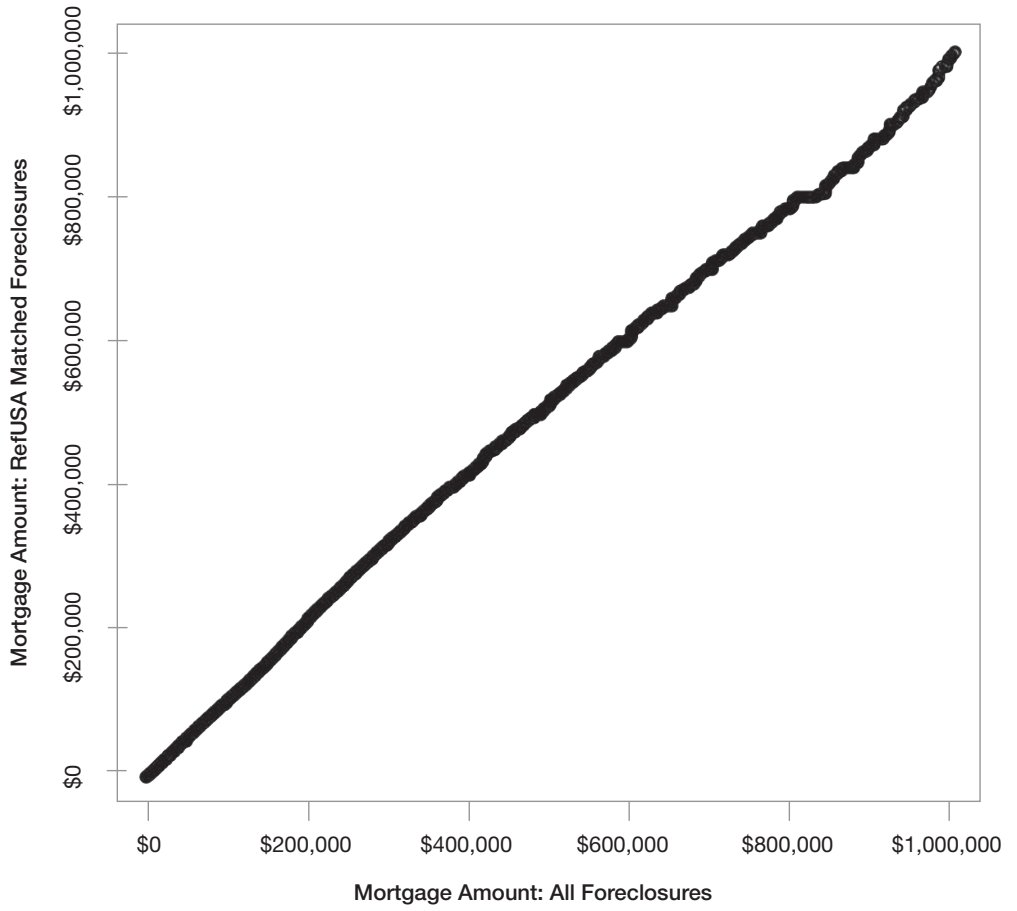
[7] See also Makridis and Ohlrogge (2018) for additional validation tests and discussion.

[8] A Q-Q plot is a graphical method for comparing two probability distributions by plotting their quantiles (for example deciles, quintiles, quartiles, medians) against each other.

[9] In this plot, we restricted to mortgages less than $1 million in initial value. Only very sparse coverage exists for mortgages higher than that value, making comparisons of distributions much less meaningful.

Exhibit 1

Q–Q Plot—Mortgage Amount All Foreclosures vs. RefUSA Matched Foreclosures

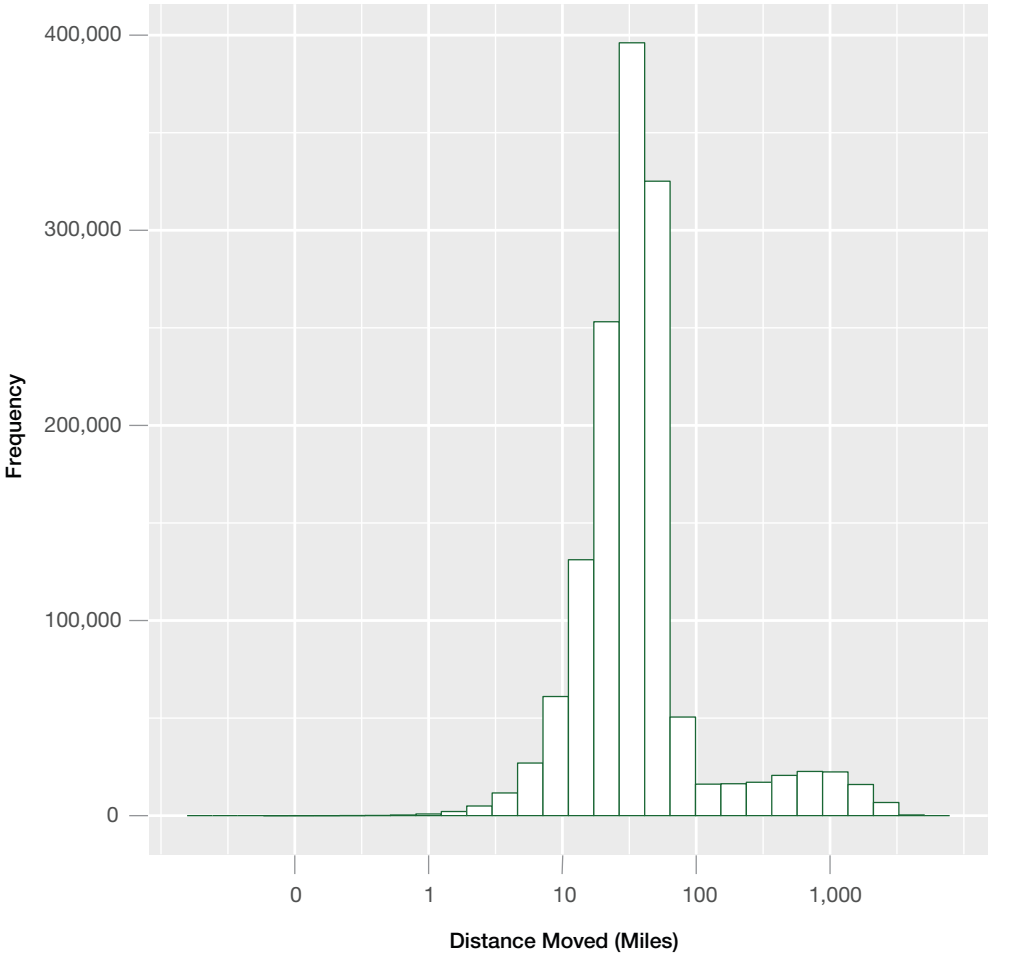


Notes: This plot compares the distribution of the total mortgage amount among the full set of foreclosure records we constructed from the registrar of deeds data and the 34.6 percent of those records for which we were able to identify where the foreclosed household moved afterwards, based on matching the foreclosure records with the RefUSA consumer mail marketing database. If, for instance, more affluent individuals were more likely to be matched, that would be reflected in this plot. We view this Q-Q plot as a sufficient statistic for evaluating comparability of the matched and unmatched samples along the dimension of mortgage amount.

On the basis of those results, which suggest that the data reliability is still preserved given the matching process, we now present several summary results that demonstrate some of the investigations this data set makes possible. First, exhibit 2 presents the distribution of distances moved by people post-foreclosure. The exhibit shows that the large majority moved between 10 and 100 miles to a new location, with smaller numbers moving lesser and greater distances. Exhibit 3 presents specific empirical quantiles for the distribution of distances moved. The median distance moved is 34 miles—comparable, for instance, to a move from Oakland, California to San Jose, California. The 90th percentile for distance moved was 95 miles—comparable to a move from Philadelphia, Pennsylvania to New York, New York. The 95th percentile for distance moved was 548.1 miles—comparable to a move from Memphis, Tennessee to Atlanta, Georgia.

Exhibit 2

Distance Moved After Foreclosures



Notes: This graph depicts the distance (measured in miles) that individuals moved after foreclosure. As evidenced here, the vast majority of people moved within 10 and 100 miles, with smaller numbers moving less than 10 miles and a moderate number moving between 100 and 1,000 (or more) miles to a new location post-foreclosure. See exhibit 3 for specific quantiles of the distribution presented here.

Exhibit 3

Distance Moved After Foreclosure—Percentiles

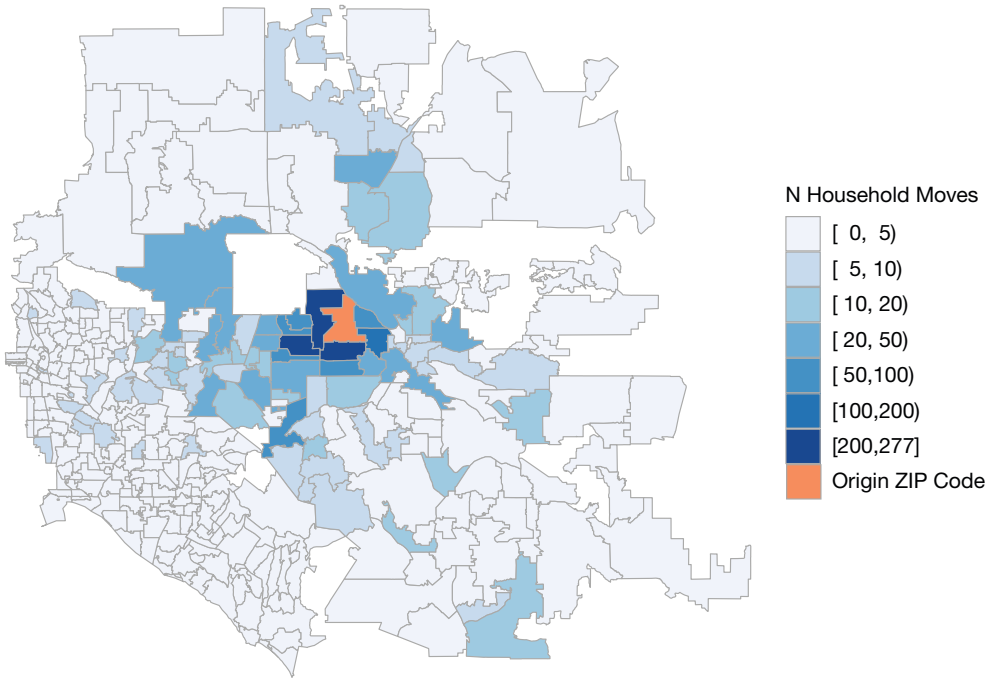
Quantile (%)	1	5	10	25	50	75	90	95	99
Miles Moved	3.8	8.8	12.7	21.5	34.0	49.1	95.0	548.1	1,744.9

Notes: This table presents several of the empirical percentiles for distance moved after foreclosure, corresponding to the distribution plotted in the histogram in exhibit 2.

Next, exhibit 4 focuses on one of the ZIP Codes (92336) in San Bernardino, California that was at the epicenter of the foreclosure crisis. The figure plots the diaspora of people moving from ZIP Code 92336 to a neighboring area.

Exhibit 4

Foreclosure Diaspora from ZIP Code 92336



Notes: This exhibit presents ZIP Code 92336 (depicted here in orange), in San Bernardino, California, which is one of the ZIP Codes with the largest number of foreclosures in the United States. We observe 4,112 foreclosures in this ZIP code from 2006 through 2011 for which we can identify where the household moved after foreclosure. Of that number, 29.3 percent (1,205 households) remained in the given ZIP Code. Of the 2,907 that left the ZIP Code, 2,403 remained within a roughly 25-mile radius of the ZIP Code. This plot depicts how many individuals moved from ZIP Code 92336 to each other ZIP Code within that 25-mile radius.

Matching Deed Mortgage Records to HMDA

To gain additional information on foreclosed households, such as the income reported at the time of taking out the mortgage as well as race and gender information on the borrowers, we also match the mortgage records created from the registrars of deeds data with the HMDA data.[10] Both sets of data are extremely detailed, which makes accurate matching possible, although not without some substantial challenges that must first be overcome. We match on the basis of agreement of census tract, loan amount, year of loan origination, lender name, whether the loan is a purchase or a refinance loan, and whether the loan is a conventional mortgage, FHA backed, or VA backed.

Census tracts are extremely small pieces of land, making them ideal for such a matching operation. The U.S. Census Bureau targets between 2,000 and 8,000 people in each tract and re-adjusts their boundaries after each decennial census to maintain their populations within that range. As supplied by CoreLogic, Inc., the deed data comes with the census tracts as designated by the 2010 census. HMDA data reports 2010 census tracts starting in the year 2013.[11] For loans originated from 2003 to 2012, HMDA data uses the 2000 census tracts, and for loans from 1993 to 2002, HMDA data uses the 1990 census tracts. The boundaries of census tracts can change over time. Thus, for loans in the deed data originated before 2013, it is necessary to first compute the 2000 or 1990 census tract (depending on the loan's origination date) before matching with HMDA data.

We use ArcGIS software and geographic shape files designating the boundaries of the census tracts to map the addresses in the deed records to census tracts from the appropriate census year.[12] The shape files we use in this process are from the National Historical Geographic Information System (NHGIS).

A final challenge is that the names of lenders in the two databases are not precisely the same (for instance, one may contain "Bank of America," whereas the other lists "Bank of America, NA"). We therefore use an algorithm based on the Levenshtein distance between strings to pair lender names that are not exact matches.

In total, we are able to match 68 percent of the mortgages in the deed data to HMDA data, for a total of 65 million mortgage records, including both foreclosed and non-foreclosed mortgages.[13] In Makridis and Ohlrogge (2018), we perform analyses similar to those discussed previously to validate that the distributions of observable characteristics were very similar in the set of foreclosures that we did and did not successfully match to HMDA data.

[10] For 2006, we obtained HMDA data from the Inter-university Consortium for Political and Social Research: https://doi.org/10.3886/ ICPSR24612.v2; for later years, we obtained it directly from the Federal Financial Institutions Examination Council: https://www.ffiec.gov/hmda/hmdaproducts.htm; for earlier years, we obtained CDs from the National Archives.

[11] For details on HMDA data reporting, see, for example, https://www.ffiec.gov/hmda/guide.htm.

[12] In particular, we first found the latitude and longitude coordinates for each address. We then matched those coordinates to shape files for the census tracts.

[13] This success rate closely matches that of other researchers who have performed similar such matching. See, for example, work by Nancy Wallace: https://bfi.uchicago.edu/sites/default/files/file_uploads/3_Wallace_MFM.pdf. In calculating the percentage of deed records matched to HMDA data, we first removed certain deed records that would obviously not be in HMDA data, such as those for loans made to develop large tracks of land for new subdivisions, loans for commercial properties, and so forth.

Alternative Data Sources

The comprehensive country registrar of deeds data that we use in this study was licensed for a fee from CoreLogic, Inc.; however, other companies also sell comparable data, offering researchers the potential to identify the best vendor for their purposes and to potentially negotiate better prices. Attom Data Solutions[14] is one such option. Some U.S. counties may well also provide this information for free in a form accessible to researchers, although we have not specifically investigated that possibility. Similarly, although we use consumer mail marketing data licensed from RefUSA, other firms, such as the Experian credit rating agency, also offer similar products for license.[15]

Acknowledgments

We would like to thank Brian Wang for diligent research assistance and Kris Kasianovitz for assistance in acquiring the data used in this project. Michael thanks the Olin Foundation for Law and Economics and the Stanford Advanced Financial Technologies (AFT) Lab for financial support. Christos thanks the National Science Foundation Graduate Research Fellowship and the Stanford Institute for Economic Policy Research Shultz Fellowship for financial support. All opinions expressed here are strictly our own and not those of any institution with which we are affiliated. Michael Ohlrogge was a Ph.D. candidate in Stanford University's Department of Management Science and Engineering at the time this article was written.

Authors

Michael Ohlrogge is an assistant professor at New York University School of Law.

Christos Makridis currently serves on the White House Council of Economic Advisers, is a Digital Fellow at the MIT Sloan School of Management's Initiative on the Digital Economy, a non-resident fellow at the Harvard Kennedy School of Government's Cyber Project, and Baylor University's Institute of Religious Studies.

References

Brevoort, Kenneth P., and Cheryl R. Cooper. 2013. "Foreclosure's Wake: The Credit Experiences of Individuals Following Foreclosure," *Real Estate Economics* 41 (4): 747–792. DOI: 10.1111/1540-6229.12021.

CoreLogic, Inc. 2017. *United States Residential Foreclosure Crisis: Ten Years Later*. Irvine, CA: CoreLogic, Inc. https://www.corelogic.com/research/foreclosure-report/national-foreclosure-report-10-year.pdf.

[14] The website for Attom Data Solutions is https://www.attomdata.com/data/property-data/recorder-data/.

[15] Experian's website is http://www.experian.com/small-business/mailing-lists.jsp.

Diamond, Rebecca, and Timothy McQuade. 2016. Who Wants Affordable Housing in their Backyard? An Equilibrium Analysis of Low Income Property Development. NBER Working Paper No. w22204. Cambridge, MA: National Bureau of Economic Research.

Federal Housing Finance Agency (FHFA). 2018. *Foreclosure Prevention Report*. Washington, DC: Federal Housing Finance Agency. https://www.fhfa.gov/AboutUs/Reports/ReportDocuments/FPR_1Q2018.pdf.

Ferreira, Fernando, and Joseph Gyourko. 2015. A New Look at the U.S. Foreclosure Crisis: Panel Data Evidence of Prime and Subprime Borrowers from 1997 to 2012. NBER Working Paper No. 21261. Cambridge, MA: National Bureau of Economic Research.

Lee, Donghoon, and Wilbert Van der Klaauw. 2010. *An Introduction to the FRBNY Consumer Credit Panel*, No. 479. New York: Federal Reserve Bank of New York.

Makridis, Christos, and Michael Ohlrogge. 2018. Moving to Opportunity? The Geography of the Foreclosure Crisis & the Importance of Location. Working paper (April 14). Stanford, CA: Stanford University; Cambridge, MA: Massachusetts Institute of Technology. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3162905.

———. 2017. "Validating RefUSA Micro-Data with the Longitudinal Employer-Household Dynamics Data," *Economics Letters* 152: 83–87. DOI: 10.1016/j.econlet.2017.01.001.

Molloy, Raven, and Hui Shan. 2013. "The Post-Foreclosure Experience of U.S. Households," *Real Estate Economics* 41 (2): 225–254. DOI: 10.1111/j.1540-6229.2012.00344x.

Ohlrogge, Michael. 2018. Bank Capital and Risk Taking: A Loan Level Analysis. Working paper. Stanford, CA: Stanford University, Stanford Center for Financial and Risk Analytics. https://papers.ssrn.com/sol3/papers.cfm?abstract_id= 3053240.

Graphic Detail

Geographic Information Systems (GIS) organize and clarify the patterns of human activities on the Earth's surface and their interaction with each other. GIS data, in the form of maps, can quickly and powerfully convey relationships to policymakers and the public. This department of Cityscape includes maps that convey important housing or community development policy issues or solutions. If you have made such a map and are willing to share it in a future issue of Cityscape, please contact john.c.huggins@hud.gov.

Visualizing and Comparing Residential Permit Data Using Lollipop Plots

Alexander Din
U.S. Department of Housing and Urban Development

The views expressed in this article are those of the author and do not represent the official positions or policies of the Office of Policy Development & Research, the U.S. Department of Housing and Urban Development, or the U.S. Government.

Residential permits are a common indicator of housing market activity. Residential permits indicate the demand for new homes, and by categorizing homes into different construction types, it is possible to understand what types of homes are in-demand in the market and the types of homes that the market is producing. In this article, I use a cross between a scatter plot and a bar chart called a lollipop plot to visualize residential permits by year for single-family dwellings (SFDs) and townhomes in Montgomery County, Maryland. These data were obtained from dataMontgomery (2019), the open data portal for the county. These data are for construction permits that were finalized between 2000 and 2018 for SFDs and townhomes, as far back as data were available. Between 2000 and 2018, there were 14,831 and 6,322 permits for SFDs and townhomes, respectively.

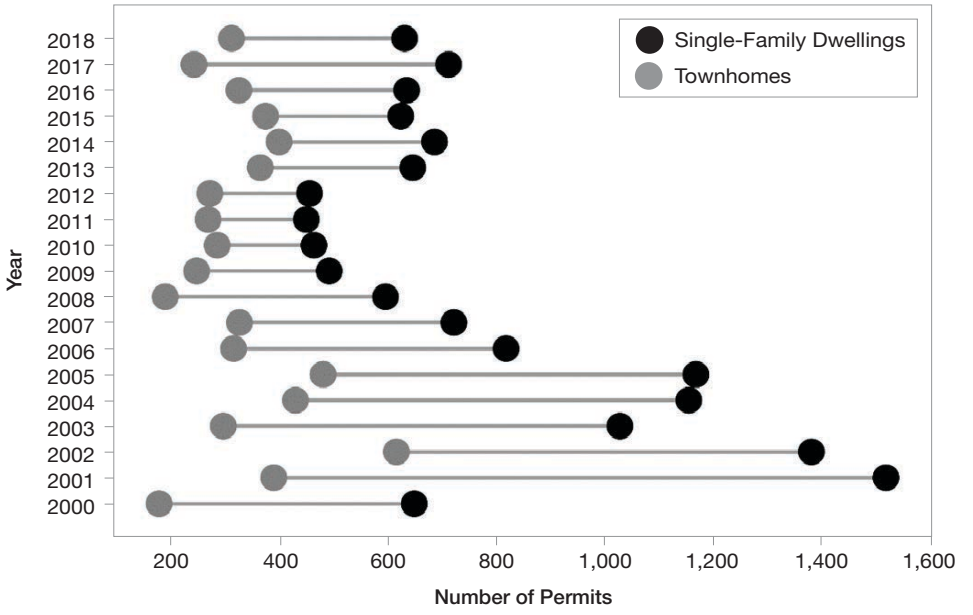
The lollipop plot compares and offers a visualization of a pair of variables in a given record. In this demonstration, each year is a record, and the variables are the number of permits finalized for SFDs and townhomes. A pair of points is plotted on the graph and a line is established between each pair, demonstrating the difference between the two values. For example, as seen in exhibit 1, there were 649 SFD permits and 179 townhome permits in 2000; the lollipop plots a value at each data point

and then connects the point via a line. Values for SFDs and townhomes are color-coded. In addition to being able to understand how many permits of each type were issued each year, the reader can also discern the difference between the numbers of permits of each type issued in a given year.

The first exhibit shows a standard lollipop plot. For each year, there are a pair of plotted points for the SFD and townhome values that are color-coded by category. The line between each pair indicates the difference between the numbers of permits issued by each type. In the years prior to the 2007–2008 financial crisis, SFD permits dwarfed townhome permits issuances. During the recession years, from 2009 to 2013, SFD permits continued to be about twice the number of townhome permits. In all years, permit issuances for SFDs were greater than those for townhomes. Since 2013, permits for townhomes generally increased slightly, while permits for SFDs increased considerably. In addition to showing the number of permits issued each year, the lollipop plot adds value by providing a visualization of the difference between the types of permits issued, allowing patterns to emerge and be more easily recognized.

Exhibit 1

Lollipop Plot for Single-Family Dwellings and Townhomes by Year

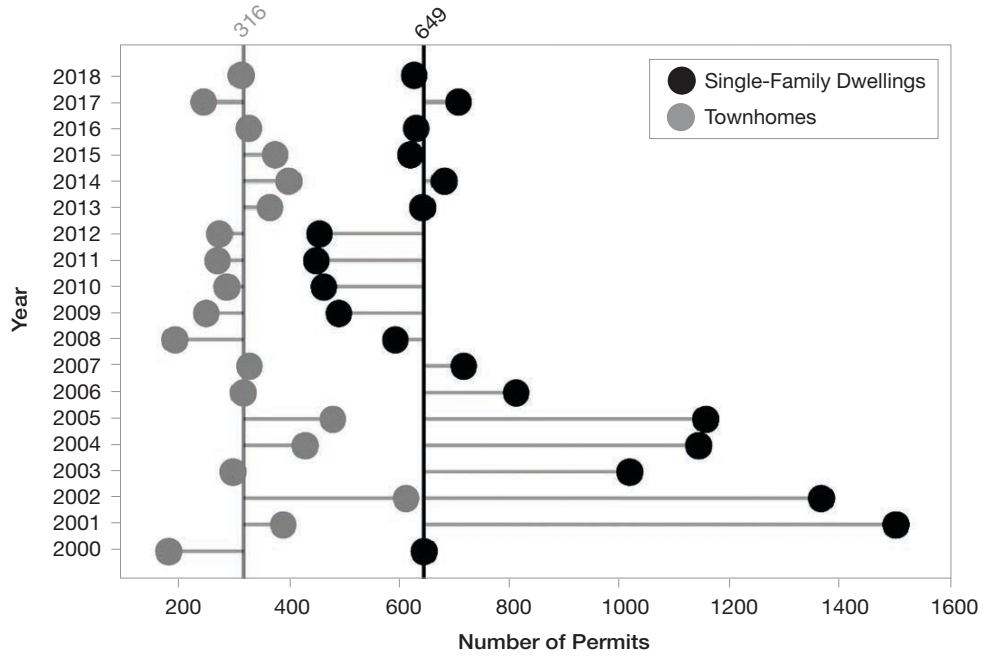


The second exhibit is a modification of the lollipop plot that, instead of comparing the difference between the types of construction, compares the median value for that type of construction over the time period. Between 2000 and 2018, the median values for final SFD and townhome permit issuances per year were 649 and 316, respectively. From 2001 to 2007, SFD permits were generally far and above the median value and then steadily decreased, reaching a low point in 2011 and 2012. Since 2012, permits for SFDs have generally been similar to the median value for the timeframe as a whole. Prior to the market crisis, permits for townhomes were more variable but in

most years were above the median value. Similar to SFD permits, counts for townhome permits fell during the recession years but rose again entering the middle decade and have remained above the median value in most years since 2013. Using the median value as the mark for comparison allows for values between years to be more directly comparable.

Exhibit 2

Lollipop Plot for Single-Family Dwellings and Townhomes by Year Compared to Median Value for Each Permit Type



Acknowledgments

Alexander Din was a housing research and GIS analyst for the Maryland Department of Housing and Community Development at the time this article was written.

Author

Alexander Din is a social science analyst at the U.S. Department of Housing and Urban Development, Office of Policy Development & Research.

Reference

dataMontgomery. 2019. Residential Permits. https://data.montgomerycountymd.gov/Licenses-Permits/Residential-Permit/m88u-pqki/data.

Impact

A regulatory impact analysis must accompany every economically significant federal rule or regulation. The Office of Policy Development and Research performs this analysis for all U.S. Department of Housing and Urban Development rules. An impact analysis is a forecast of the annual benefits and costs accruing to all parties, including the taxpayers, from a given regulation. Modeling these benefits and costs involves use of past research findings, application of economic principles, empirical investigation, and professional judgment

Acceptable Separation Distance Standards for Residential Propane Tanks

Maria Chelo Manlagnit De Venecia
U.S. Department of Housing and Urban Development

Summary of Analysis

Propane is the third most widely used fuel in the United States by number of households, after electricity and natural gas. Residential households primarily use propane for space heating, water heating, and cooking (U.S. Energy Information Administration, 2015). The U.S. Department of Housing and Urban Development's (HUD's) safety regulations establish standards for the acceptable separation distance (ASD) between HUD-assisted projects and hazardous operations that include the storage of substances that have the potential to cause an explosion or fire. Currently, the ASD from an aboveground tank ranges from 125 to 139 feet. The proposed rule would reduce the ASD of aboveground tanks to 10 feet if the capacity of the tank is 250 gallons or less and if the storage tank complies with National Fire Protection Association (NFPA) Code 58 concerning liquefied petroleum gas (LP-gas) (NFPA, 2017).[1]

The benefits of the rule include reduction in costly mitigation measures, increased availability of HUD-assisted projects, and improved administrative efficiencies. The reduction in cost

[1] NFPA is a nonprofit organization that develops consensus codes and standards and provides research, training, and education to eliminate death, injury, property loss, and economic loss due to fire, electrical, and related hazards. NFPA developed Code 58, which provides voluntary consensus standards used by the propane industry for the storage, handling, transportation, and use of LP-gas.

from mitigation measures is estimated to be from $100,000 to $4 million per year and involve approximately 20 projects per year. Given the quality of propane tanks, allowing a more proximate placement of a propane tank to property will present an imperceptible increase of risk to that property. Consequently, the savings from reducing mitigation costs is likely to outweigh the reduced safety benefits.

Costs of the Proposed Rule

The costs of the rule stem from increased danger due to proximity. LP-gas is flammable at a wide range of concentrations. Allowing greater proximity to a potential hazard will expose properties and individuals to greater risk from explosions and fires. The reliability of propane tanks has increased significantly, however, since HUD's rule was originally promulgated more than 30 years ago. HUD's original regulation was based on studies from 1975 and 1982 and so, does not take into consideration increased safety features of new technology in tank designs and updated fire safety codes and standards, including NFPA Code 58.

More recent studies suggest that the evolution of industry safety practices has reduced the probability of propane tank failure (Aherns, 2018, 2017; Flynn, 2010; Hall, 2014). NFPA, for example, asserts that home structure fires in which LP-gas was the type of material first ignited have fallen by 62 percent from 1980 to 2007 (Flynn, 2010), or an average annual decrease of 3.5 percent. Among the reasons for the safety improvement are improved construction techniques and materials and strengthened fire codes (U.S. Fire Administration, 2017).

One study (Bullerdiek, 1986) revealed that particulate contamination; leaks from corroded or damaged piping; nonprofessional installation, inspections, and maintenance; and failure to detect odor or leakage were the major factors leading to appliance failure and subsequent injury. Later studies examined the adequacy of ethyl mercaptan as an odorant to warn residents of potential leaks (Arthur D. Little, Inc., 1987; U.S. Consumer Product Safety Commission, 1987). Many research studies tested several odorants, but ethyl mercaptan remains the odorant of choice.[2] Related to the issue on odorant fading, two new requirements have recently been added to the 2017 edition of NFPA Code 58 to facilitate tracking and recognition of LP-gas. Although the risk of tank failure cannot be totally eliminated through better design, significant progress has been made in safety standards to make HUD's ASD overly cautious.

Benefits of the Proposed Rule

Reduced Costly Mitigation Measures

If a developer is not able to meet the minimum ASD requirement, there are acceptable ways of mitigating the hazard. Those measures—including tank burial, building a blast wall to surround the tank, or building a structure on a HUD-assisted property site to shield a proposed project from the hazard—require an elaborate process involving costly construction procedures. The

[2] This conclusion was based on series of tests and research studies published by the Joint Industry Task Force on Propane Odorization (see Nette, 2017). It is recognized, however, that no odorant will be completely effective as a warning agent in every circumstance. The most recent study is Roscioli et al. (2017).

labor and materials for burying a tank or constructing a blast wall can range from $5,000 to $200,000 per unit.[3] Specifically, the cost of burying a tank ranges from $20,000 to $100,000 per unit, depending on several factors, such as local costs of labor and materials, volumetric capacity of the tank, location of the water table, frost level, and type of soil where the tank will be buried. Additional costs include permit and design fees—especially in an urban environment—corrosion protection, and tank vaulting. Blast walls, which usually are made of concrete block, mortar, or rebar, can range from $5,000 to $200,000 per unit (a wider range than that for tank burial). Aside from the local costs of labor and materials, the type of materials used for the tank, proximity of the wall to the tank (the closer to the wall, the more robust the design of the wall should be), size of the footing of the wall (depending on the size and weight of the wall), and length of the wall are among the other factors that can affect the costs of constructing a blast wall.

The proposed rule would eliminate the need for mitigation measures such as tank burial or a blast wall. Using the available information on the cost per mitigation measure, the construction cost would be $5,000 to $200,000 per unit of mitigation measure. At least 20 mitigation measures per year involve propane tanks of 250 gallons or less in HUD-assisted projects. The aggregate annual reduction in cost would be $100,000 to $4 million (multiply $5,000 to $200,000 by 20). Over a 30-year structure life, the aggregate discounted savings range from $2.0 million to $80.8 million, assuming a 3-percent discount rate, and $1.3 million to $53.1 million, assuming a 7-percent discount rate.

Increased Availability of HUD-Assisted Projects

With the proposed rule, HUD-assisted projects could be closer to propane tanks, increasing the availability and choices for siting projects, especially in areas in need of additional housing units. An estimated 30,890 housing projects with onsite propane tanks could be affected by the proposed rule. Land is usually a significant cost of development and, depending on location, can be the greatest cost. Reducing the amount of land needed for development will reduce a regulatory barrier to development and possibly increase the supply of housing.

Improved Administrative Efficiency

The proposed rule clarifies HUD's regulations on propane containers and aligns them with industry practices. Under NFPA Code 58, the minimum separation distance required for containers between 125 and 500 gallons from a building or line of an adjoining property is 10 feet (NFPA, 2017).[4] Having a regulation more consistent with standard practice makes siting projects easier for HUD grantees.

[3] Estimates include only costs of labor and materials. Other costs such as associated permits, design, and other administrative fees involved in burying a tank or building a blast wall are not included.

[4] See NFPA 58, Table 6.4.1.1. The separation distances are based on the potential hazard of LP-gas; size and type of equipment used to contain it; possibility of leaks (which can ignite); and need for fuel in buildings—not on a worst-case scenario in which the LP-gas container fails catastrophically but as a minimum safe distance for radiant heat exposure to the containers and from the containers. Research has been done to evaluate the effects of radiant heat from fires to LP-gas containers. For example, see Raj (2005).

Conclusion

This proposed rule would reduce the regulatory burden on communities restricted in their ability to site HUD-assisted projects—including low- and moderate-income housing—because of nearby stationary aboveground propane storage tanks. The proposed exception will save HUD grantees the cost of constructing mitigation measures to address residential propane tanks on properties that do not meet the ASD. Based on current technology, developments in industry safety practices, increased awareness to new safety standards, and improved quality of propane tanks, HUD does not expect a noticeable increase in risk.

Acknowledgments

The author thanks Alastair McFarlane for encouraging the development of this article and providing guidance in preparing the regulatory impact analysis on the proposed rule.

Author

Maria Chelo Manlagnit De Venecia is an economist in the Public Finance and Regulatory Analysis Division, Office of Policy Development and Research, U.S. Department of Housing and Urban Development.

References

Ahrens, Marty. 2018. *Home Structure Fires in Which LP-Gas Was the Type of Material First Ignited and the Type of Material First Ignited in All Home Structure Fires*. Quincy, MA: National Fire Protection Association. https://www.nfpa.org/-/media/Files/News-and-Research/Fire-statistics-and-reports/US-Fire-Problem/Fire-causes/osHomeStructureFiresLPGasFirstIgnited.ashx.

———. 2017. *Home Structure Fires Involving Lightning and/or Natural Gas or LP-Gas*. Quincy, MA: National Fire Protection Association.

Arthur D. Little, Inc. 1987. *Characterization of LP-Gas Odorant Fade*. Report prepared for the U.S. Consumer Product Safety Commission. Washington, DC: Government Printing Office.

Bullerdiek, W.A. 1986. *Engineering Hazard Analysis of Residential LP-Gas Fuel Usage*. Frontier Technical Associates Report. Report prepared for the U.S. Consumer Product Safety Commission. Washington, DC: U.S. Government Printing Office.

Flynn, Jennifer. 2010. *Natural Gas and LP-Gas Home Structure Fires*. Quincy, MA: National Fire Protection Association.

Hall, John. 2014. *Fires Starting with Flammable Gas or Flammable or Combustible Liquid*. Quincy, MA: National Fire Protection Association. https://www.nfpa.org/-/media/Files/News-and-Research/Fire-statistics-and-reports/Hazardous-materials/osfiresstartinggas.pdf.

National Fire Protection Association (NFPA). 2017. *NFPA 58, Liquified Petroleum Gas Code*, 2017 ed. Quincy, MA: National Fire Protection Association. https://www.nfpa.org/codes-and-standards/all-codes-and-standards/list-of-codes-and-standards/detail?code=58.

Nette, Eric. 2017. *NFPA 58 LP-Gas Code Handbook*, 11th ed. Quincy, MA: National Fire Protection Association.

Raj, Phani. 2005. "Exposure of a Liquified Gas Container to an External Fire," *Journal of Hazardous Materials* 122 (1–2): 37–49.

Roscioli, Joseph R., Scott Herndon, David D. Nelson, and Tara Yacovitch. 2017. *Methane, Ethane, and Propane Sensor for Real-Time Leak Detection and Diagnostics*. Washington, DC: U.S. Department of Energy. https://www.osti.gov/servlets/purl/1347976.

U.S. Consumer Product Safety Commission. 1987. *Information on LP-Gas (Propane) Odorant*. Memorandum. Washington, DC: U.S. Consumer Product Safety Commission.

U.S. Energy Information Administration. 2015. *Residential Energy Consumption Survey (RECS)*. Washington, DC: U.S. Energy Information Administration. https://www.eia.gov/consumption/residential/data/2015/.

U.S. Fire Administration. 2017. *Fire in the United States 2006–2015*, 19th ed. Emmitsburg, MD: U.S. Fire Administration. https://www.usfa.fema.gov/downloads/pdf/publications/fius19th.pdf.

Additional Reading

Arthur D. Little, Inc. 1975. *Safety Consideration in Siting Housing Projects*. Report prepared for the U.S. Department of Housing and Urban Development, Office of Policy Development and Research. Washington, DC: U.S. Government Printing Office.

Rolf Jensen & Associates, Inc. 1982. *Urban Development Siting with Respect to Hazardous Commercial/Industrial Facilities*. Report prepared for the U.S. Department of Housing and Urban Development, Office of Policy Development and Research. Washington, DC: U.S. Government Printing Office.

U.S. Department of Housing and Urban Development, Office of Community Planning and Development. 2017. *Memorandum for the Approval of Waiver for the State of Vermont of 24 CFR Part 51, Subpart C*. Washington, DC: U.S. Department of Housing and Urban Development.

————. 2011. *Acceptable Separation Distance Guidebook*. Washington, DC: U.S. Department of Housing and Urban Development.

————. 2008. *Memorandum for the Approval of Waiver for the State of Mississippi of 24 CFR Part 51, Subpart C*. Washington, DC: U.S. Department of Housing and Urban Development.

————. 1987. *Siting of HUD-Assisted Projects Near Hazardous Facilities*. Washington, DC: U.S. Department of Housing and Urban Development.

Policy Briefs

The Policy Briefs department summarizes a change or trend in national policy that may have escaped the attention of researchers. The purpose is to stimulate the analysis of policy in the field while the policy is being implemented and thereafter. If you have an idea for future Policy Briefs, please contact david.l.hardiman@hud.gov.

Confirmations, New Insights, and Future Implications for HOPE VI Mixed-Income Redevelopment

Taryn H. Gress
Mark L. Joseph
Seungjong Cho
Jack, Joseph and Morton Mandel School of Applied Social Sciences
Case Western Reserve University

Abstract

As HUD advances its mission to create strong, sustainable, inclusive communities and high-quality, affordable homes for all, the promotion of mixed-income communities has become a core strategy. Across the United States, local governments and private developers are increasingly turning to mixed-income development as an approach to deconcentrate poverty and revitalize urban neighborhoods. With the Choice Neighborhoods Initiative, launched in 2010, the federal government extended its commitment to supporting the mixed-income approach to public housing transformation first implemented through the HOPE VI initiative in the mid-1990s. A comprehensive analysis of mixed-income units produced through the $6 billion HOPE VI program has not yet been undertaken, however. Using HOPE VI grantee quarterly report data from 1993–2014, we analyze the income and tenure mix of housing units that have been produced through the HOPE VI program as follows:

- *Examine variations by factors such as age and size of development, region, and developer and describe the evolution of the program over time.*
- *Examine the factors associated with higher or lower reoccupancy of original residents.*
- *Examine funding leveraged through the HOPE VI and other grants.*
- *Review resident participation in community and supportive services.*
- *Undertake an analysis of the duration of various phases of the initiative, including relocation, demolition, construction, and occupancy.*

Based on this descriptive analysis, we propose implications for future implementation and an evaluation of the Choice Neighborhoods Initiative.

Introduction

Since the mid-1990s, mixed-income development has gained increased attention across the United States as an urban revitalization strategy (Chaskin and Joseph, 2015; Joseph, Chaskin, and Webber, 2007). Mixed-income developments attempt to eliminate the isolation that challenges low-income households living in neighborhoods with high concentrations of poverty (Wilson, 1987). The term "mixed-income development" indicates the intentional financing, design, and construction of a residential property with the goal of attracting individuals from a broad range of income and socioeconomic levels (Brophy and Smith, 1997). As the U.S. Department of Housing and Urban Development (HUD) advances its mission to create strong, sustainable, inclusive communities and quality, affordable homes for all, the promotion of mixed-income communities has become one of the agency's core strategies.

The HOPE VI program was launched in 1992 as a response to recommendations by the National Commission on Severely Distressed Public Housing on the condition of the U.S. public housing stock (Cisneros and Engdahl, 2009; Joice, 2017; Katz 2009; Vale, 2019). After 1995, the focus of the program became the mixed-income transformation of deteriorating and isolated public housing sites. To date, the HOPE VI program has been the largest and longest running federal public housing transformation initiative, with 260 grants totaling more than $6 billion. The Obama Administration replaced the program with the Choice Neighborhoods Initiative in 2010 (Joice, 2017; Pendall et al., 2015; Pendall and Hendey, 2013).

In general, the HOPE VI program had a mixed record of accomplishment. The national HOPE VI Panel Study by the Urban Institute found positive consequences of the program, such as improved safety, better physical environments, and lower poverty rates (Buron, 2004; Comey, 2004; Popkin et al., 2004; Popkin and Cunningham, 2009; Popkin et al., 2012). At the same time, the panel study also reported that most of the original public housing residents did not return to benefit from the new housing developments. The program has also been criticized for reducing the number of public housing units and for delays in constructing new units (Center for Community Change, 2003; National Housing Law Project, 2002).

Contribution of this Paper

Despite an extensive literature of evaluation reports and articles on various aspects of the HOPE VI program, there has been no comprehensive outcomes analysis of the overall program. Most of the research on the HOPE VI program has analyzed a single site or a selected subset of locations (Buron et al., 2002; Chaskin and Joseph, 2015; Graves, 2011; Popkin et al., 2004; Popkin, Levy and Buron, 2009; Tach, 2009; Vale, 2013; 2019). Addressing the gap in extant literature, this paper provides the results of the descriptive analysis of data extracted from quarterly progress reports from all 260 HOPE VI grants from 1993 through September 2014.[1] The primary research question that motivated our analysis was: What is the tenant income and tenure mix of housing units that have been produced through the HOPE VI program?

[1] Among 260 sites, one site did not produce any units. The site had been planned to produce 100 units but had not produced any as of the end of the reporting period. Thus, for all analysis using unit production information, 259 sites were analyzed, excluding that site.

Although not an evaluation of the program, the tenant income and tenure mix analysis here provides insight into the nature of the mixed-income developments produced with HOPE VI funding and, therefore, the potential influence that those developments had on residents and communities through provision of a mix of subsidized and market-rate housing. Using the data available in HOPE VI grantee reports, we analyze "income mix" based on the mix of unit types (public housing replacement units, units subsidized through the Low-Income Housing Tax Credit program, and market-rate units without a subsidy). Our focus is on HOPE VI as a place-based, mixed-income development strategy, and we focus primarily on data available in the revitalization grants. Our extraction and compilation of this data on all HOPE VI projects has enabled us to revisit and reconsider some of the accumulated knowledge and conventional wisdom about the program.[2] Based on our analysis of this comprehensive data on unit production, we frame this paper around three sets of findings. First, based on this new level of detail about HOPE VI grants, what have we been able to *confirm* that was known or surmised previously? Second, what *new* information have we been able to uncover with this data and where have we been able to shed new light? Third, what are the implications of these confirmations and insights for future mixed-income research, policy, and practice?

With the Choice Neighborhoods Initiative (CNI), the federal government extended its commitment to supporting the mixed-income approach to public housing transformation, and now more than 100 CNI planning and implementation grants are underway across the United States (HUD, 2015). Although the Trump Administration has continued funding the Choice Neighborhoods Initiative, the future direction of public housing policy is quite uncertain. This newly compiled dataset provides an opportunity to take stock of what was accomplished and learned from the country's longest running public housing transformation initiative. Based on this analysis of HOPE VI grant production, we identify policy implications that may inform the implementation and evaluation of existing Choice Neighborhoods Initiative projects and other future mixed-income development projects. We first summarize some of the existing knowledge from the larger body of literature on the HOPE VI program and then identify some of the key knowledge gaps that will be addressed in this paper.

Relocation, Demolition, and Construction

We begin our review of existing knowledge on the HOPE VI program with a look at the relocation of public housing residents away from distressed public housing units, the demolition of those public housing units, and the construction of new public housing, affordable units, and market-rate units.

Demolition of public housing. In 1992, the National Commission on Severely Distressed Public Housing estimated that nationwide, approximately 86,000 (of 1.3 million, roughly 6.5 percent) public housing units were "severely distressed" (NCSDPH, 1992). Initially, the HOPE VI program focused on revitalization, rather than demolition, of the severely distressed public housing units (Goetz, 2003). The program soon changed priorities, however, instead offering housing vouchers to original residents for housing off-site and demolishing units without the one-for-one replacement requirement (Katz, 2009). HUD shifted to an approach that promoted the economic

[2] The data extraction and compilation were conducted in partnership with a research team at the Massachusetts Institute of Technology led by Dr. Lawrence Vale and Dr. Shomon Shamsuddin, now at Tufts University.

integration of former public housing sites by rebuilding public housing with units available for higher-income residents to prevent the reconcentration of poverty (Turbov, 2006).

Replacement of public housing units. The one-for-one replacement requirement required housing authorities to construct one new unit of affordable housing for every unit demolished. With limited public funding for new construction, that requirement had been "the largest obstacle to implementing HOPE VI" (Goetz, 2003: 59). In 1995, that requirement was removed, with the intention to replace subsidized housing units with a combination of new construction or rehabilitated public housing units and Section 8 tenant-based vouchers.

Some critics of this approach disagreed with the approach of replacing physical public housing with some proportion of vouchers that facilitate a move to the private market. For instance, the Center for Community Change (2003), in their report, *A Hope Unseen: Voices From the Other Side of HOPE VI*, argued that any loss of affordable housing was problematic and by privatizing public housing, the HOPE VI program was exacerbating a dire social problem. They stated that, by the program's close in 2012, 48,643 public housing units "will be gone, forever unavailable to low-income families, particularly extremely low-income families" (Center for Community Change, 2003: 9). They also made the point that construction of other housing on redeveloped sites, including those with Low-Income Housing Tax Credits (LIHTC), do not provide the same level of affordability unless accompanied by other rental subsidies. The LIHTC program generally provides affordability for households earning between 40 and 60 percent of the area median income, whereas public housing typically serves families at or below 30 percent of area median income due to the ongoing annual operating subsidy. In addition, LIHTC units have time limitations on their affordability status, jeopardizing long-term access to affordable units (Khadduri et al., 2012).

Construction of market-rate units. Any review of HOPE VI has to acknowledge how the program developed and changed over the course of its two decades of operation. Although some of the first HOPE VI grantees renovated all of the old buildings for reoccupancy by public housing families exclusively, by 1995, HUD had more fully committed to a strategy of deconcentration together with mixed-income and mixed-finance redevelopment across all sites. This dimension of the HOPE VI program, which included the construction of market-rate housing units, also fit within a broader move to elevate the role of public-private partnerships in the provision of affordable housing while expanding the role of public housing agencies in the mixed-finance realm. To make that possible, in 1996, the HUD Office of Public Housing Investments published the mixed-finance rule, which laid out the provisions for the formation of public-private partnerships to leverage public-sector funds with private-sector financing to produce mixed-income housing. Two years later, the Quality Housing and Work Responsibility Act of 1998 (QHWRA) added an expanded mixed-finance provision that allowed the use of public funds and operating subsidies for projects to be owned by private developers (Baron, 2009). Together, these steps created the opportunity to leverage private funds to design and implement HOPE VI developments that would be owned and operated by private companies (Wexler, 2001). Mixed-finance development provided a way of generating the capital needed to construct new buildings and renovate existing ones and the operating capital to manage and sustain the brick-and-mortar investments. That move raised critiques among affordable housing advocates that the privatization of public housing was more focused on

benefiting capital interests and higher income city residents than positive outcomes for low-income households (Fraser and Kick, 2007; Katz, 2009; Kleit, 2005, Lees, 2008; Polikoff, 2009).

Timing of relocation, demolition, and construction. Increased uncertainty and variation in construction timeframes was one consequence of the privatization of public housing, which meant that redevelopment projects were more dependent on market conditions. Often, demolition and construction had lengthy timeframes in the redevelopment process (Crowley, 2009). The long periods between relocation, construction, and occupancy meant an extended waiting period for residents who hoped to move back to the new developments (Popkin, Levy, and Buron, 2009). Among other factors, such as resident preferences, selection criteria, eligibility, and physical and mental health needs, long construction timeframes likely had an influence on the low return rates among original residents (Popkin, Cunningham, and Burt, 2005).

Funding

The trajectory of funding for the HOPE VI program changed considerably over time.

Decrease in funding over time. The HOPE VI program was affected by significant funding cuts beginning in 2000. And beginning in 2003, the Bush Administration tried to completely eliminate the HOPE VI program each fiscal year (Polikoff, 2009). Although Congress successfully continued the program, the funding was drastically cut and continued to be reduced through the final award year in 2011. The Obama Administration closed down the program in 2010 and replaced it with the Choice Neighborhoods Initiative, with lower overall levels of federal funding than the original HOPE VI program.

Income and Tenure Mix

The HOPE VI program made it feasible for public housing, affordable and market-rate housing, and homeownership units to be within the same housing development and set a precedent for this type of mixed-income redevelopment.

Income mix types. Although no one accepted definition of mixed-income housing exists, a widely accepted definition focuses on the intentional financing and operations of a development to facilitate a socioeconomic mix of residents (Brophy and Smith, 1997). Income mix in the public housing transformation context is most often categorized into public housing, affordable housing, and market-rate housing (Baron, 2009). Income mix can also be categorized by the percentage of households in a certain income range as it relates to area median income (AMI), with conventional categories being less than 50 percent AMI, 50 to 60 percent AMI, and more than 80 percent AMI. Vale and Shamsuddin (2017: 59) identify four dimensions of income mix that can be considered in planning, implementation, and assessment: "distribution of units by subsidy type (allocation); spatial separation of income mix (proximity); distribution of homeownership versus rental units (tenure); and time limits for subsidies that preserve the income mix (duration)."

Production of public housing homeownership units. The HOPE VI program also mixed rental and homeownership units. Some of those homeownership units were set aside for former public

housing residents, as allowed by Section 32 of the U.S. Housing Act of 1937. With the opportunity to transform public housing sites, some authorities and their development partners returned to this original aspiration and worked to include homeownership for relocated public housing residents into the unit mix (Santiago and Galster, 2004; HUD, 2003). Redeveloped sites could also include unsubsidized homeownership units, including in mixed-income, townhouse-style developments.

Reoccupancy

Although residents were given the opportunity to return to the redeveloped public housing within new mixed-income developments, there were often significant barriers to their return.

Rates of return—challenges to returning. One of the major critiques of the HOPE VI program has been the low rates of return by original residents. Buron and his colleagues (2002) reported that across the country, rates of return to revitalized HOPE VI developments averaged 14 percent (see also Marquis and Ghosh, 2008). The HOPE VI grants included support for relocation services designed to help residents move temporarily and then return to sites once the redevelopment was completed. After redevelopment, however, mixed-income properties often implemented stricter requirements and new screening criteria that included adherence to lease stipulations and other stipulations, such as background checks, credit checks, and drug testing, which hampered reoccupancy efforts (Center for Community Change, 2003; Joseph and Chaskin, 2012). Those requirements, in addition to many other factors—such as limited time and information to make housing decisions, delays of several years in the delivery of replacement units, mismatch of household size to new unit sizes, and apprehensions about the monitoring and social climate in the new developments—contributed to the low return rates (Joseph and Chaskin, 2012; Popkin et al., 2004).

Community and Supportive Services

In addition to the redevelopment of deteriorating public housing, the HOPE VI program sought to provide support for households as they transitioned from high-poverty housing to new developments or alternative housing. To achieve that goal, HUD HOPE VI grants included funding for Community and Supportive Services (CSS) that gave public housing residents the "tools to enable them to improve their life skills and capacities and secure living wage jobs and, when they choose to do so, to relocate to a new neighborhood of their choice" (HUD, 2000: 1). Programming provided through CSS varied greatly across sites; examples of programs included literacy training; job preparation, training, and retention; personal management skills; daycare; youth activities; health services; community policing or security activities; and drug treatment (Naparstek et al., 2000). Research on CSS programs shows that they have not been very effective in achieving measurable outcomes (Center for Community Change, 2003; Holin et al, 2003; Levy and Woolley, 2007; Oakley, Fraser, and Bazuin, 2015; Popkin, Levy, and Buron, 2009).

Data and Methods

Data analyzed in this paper are from HUD's HOPE VI quarterly grant progress reports from 1993 through the third quarter of 2014. These quarterly reports provide the most comprehensive look

at the unit production and financial information of HOPE VI developments available to date. The quarterly project reporting was used by HUD and by the grantees as an administrative tracking tool rather than a research tool, and the projection numbers were updated over time, as new outcome targets were agreed to between HUD and grantees. The data include information on actual and projected unit production, reoccupancy, funding, and timeframes of redevelopment phases. Separate quarterly and national reports on Community Supportive Services were extracted and compiled for analysis.

Descriptive and comparative analyses of the 260 revitalization grantee reports and Community Supportive Services reports were completed using SPSS statistical software.[3] The raw data were checked for accuracy and discrepancies through extensive searches of electronic and hard-copy reports and peer-reviewed articles on the HOPE VI program. Then data were extracted, cleaned, and categorized by individual site level (n=259), state level (n=41), and award year (n=19), for a comprehensive descriptive analysis.

Findings

We have organized this section as follows. We review major elements of the HOPE VI program: relocation, demolition and construction, funding, return and relocation, and community and supportive services. For each element, we present two types of findings: (1) existing knowledge that we have been able to confirm with our analysis of the full grantee report data and (2) new information and insights that we have been able to glean.

Demolition, Production, and Replacement Housing

We begin by confirming existing knowledge about the replacement of distressed public housing units through the demolition and new construction of public housing, affordable units, and market-rate units within HOPE VI developments. We then explore new knowledge and insights about the income and tenure of units that were planned and built as part of the HOPE VI program.

Existing Knowledge Confirmed

We confirmed that there was a net loss of public housing units replaced by public housing units, affordable units, and market-rate units. ***Net loss of total housing assistance***. Our analysis of redeveloped public housing sites includes consideration of the resulting mix of public housing units and large numbers of units with shallower subsidies, mostly funded with Low-Income Housing Tax Credits, with unsubsidized units (often referred to as "market rate" units, although their affordability may vary). This analysis should be placed in the context of what seems to be an overall net loss of rental assistance—available units that are most affordable to families at the lowest part of the income scale. About 155,000 public housing units were demolished through

[3] Significance testing was not conducted because we analyzed the total population of 260 revitalization grants.

the HOPE VI program, including 56,755 units torn down through Demolition grants[4] and 98,592 units demolished through Revitalization grants that included funding for demolition. Public housing units that were demolished were to be replaced with a combination of new public housing units (either rebuilt on site or at other sites), together with Section 8 Housing Choice Vouchers. The program thus aimed to deconcentrate public housing through a combination of strategies.

A precise accounting of the apparent net loss of rental assistance is not possible in this paper based on the data available to the authors. A comparison of the approximately 155,000 total demolished public housing with the total number of replacement units from both public housing and Section 8 vouchers based on the available data, however, suggests a significant shortfall in the net total of rental assistance that was provided as replacement housing assistance. The redeveloped sites included 55,318 public housing units, including both new construction and rehabilitation units, of which the latter would likely not be considered "replacement units" for demolished units. Likewise, a precise number of total Section 8 Housing Choice Vouchers provided as replacement units is not available to the authors. A review of existing research conducted by the Urban Institute in 2004 stated,

> An estimated 63,000 to 70,000 of these supplemental vouchers were allocated to replace emolished public housing units between 1995 and 2003. But it is not known how many were for HOPE VI projects and how many were for other public housing demolition.[5]

Although further research on this topic would be needed for more accurate estimates of the total replacement housing assistance provided, this overall context is important to keep in mind when considering the remainder of this paper's analysis of development, construction, and resulting unit types at the redeveloped sites.

Redevelopment: Public Housing, Low-Income Housing Tax Credits, and Market-Rate Units

According to grantee reports, 97,389 units have been produced through the HOPE VI program (85,934 newly constructed units and 11,455 rehabbed units), including 49,949 replacement

[4] As stated, HOPE VI had two types of grants: Revitalization grants and, later, Demolition grants. There were 285 Demolition grants. Demolition grants were added, in part, to provide funding for new statutorily required demolition requirements for large projects with long-term vacancies. See HUD "List of Demolition Grants" (October 2004) at https://www.hud.gov/sites/documents/DOC_9890.PDF. Projects that received Demolition grants were eligible to apply for Revitalization grants to fund redevelopment. Among the grantees that received Demolition grants, 56 also received Revitalization grants to produce replacement units, and that unit production is thus captured in our analysis here. For statutory demolition requirements, see, for example, the mandatory demolition requirement enacted in 1996 through Section 202 of the Omnibus Consolidated Rescissions and Appropriations Act of 1996 ("to make a further downpayment toward a balanced budget") (Public Law No: 104–134) at https://www.congress.gov/104/plaws/publ134/PLAW-104publ134.pdf.

[5] For estimates on numbers of replacement vouchers, see Popkin et al., 2004, at https://www.urban.org/sites/default/files/alfresco/publication-pdfs/411002-A-Decade-of-HOPE-VI.PDF. Although assessing the reasons for the apparent shortfall of rental assistance units to replace demolished units is beyond the scope of this paper, one contributing factor may be an inconsistent and uneven implementation over time of HUD's policy on whether to provide vouchers to replace demolished public housing that were vacant (as opposed to providing vouchers for families in need of relocation assistance). See, for example, HUD Notice PIH 2002-21, "Applications for Housing Choice Vouchers for Relocation or Replacement Housing Related to Demolition or Disposition," which specifies that replacement vouchers would be provided as replacement for demolished vacant public units (see examples in Section 4 of the Notice). See PIH 2002-21 at https://www.hud.gov/sites/documents/DOC_9093.PDF. Compare with other PIH Notices on the topic in previous and subsequent years.

public housing rental units and 5,369 public housing homeownership units.[6] Of the total units, 55,318 units (56.8 percent) were public housing, 28,979 units (29.8 percent) were LIHTC units (referred to in this paper as "affordable units"), and only 13,092 units (13.4 percent) were unsubsidized, market-rate units. Developments built in the first few years of the HOPE VI program tended to include more rehabilitation of existing public housing units and also tended to replace more of the demolished public housing units with new construction replacement public housing units. As the program progressed, however, more affordable and market-rate units were also built at the redeveloped public housing sites alongside new public housing units. Sites developed later in the HOPE VI program (post-1997) averaged 53.9 percent public housing, 34.9 percent affordable units, and 11.2 percent market-rate units, whereas the sites developed in the earlier years of the program had a higher proportion of public housing units, averaging 67.5 percent public housing, 23.5 percent affordable units, and 9.0 percent market-rate units. Larger sites (≥ 319 units)[7] tended to have a lower proportion of public housing and a higher proportion of market-rate housing than did smaller sites (< 319 units). On average, the larger sites had 56.9 percent public housing units and 12.8 percent market-rate units, whereas the smaller sites had 64.4 percent public housing units and 7.4 percent market-rate units.

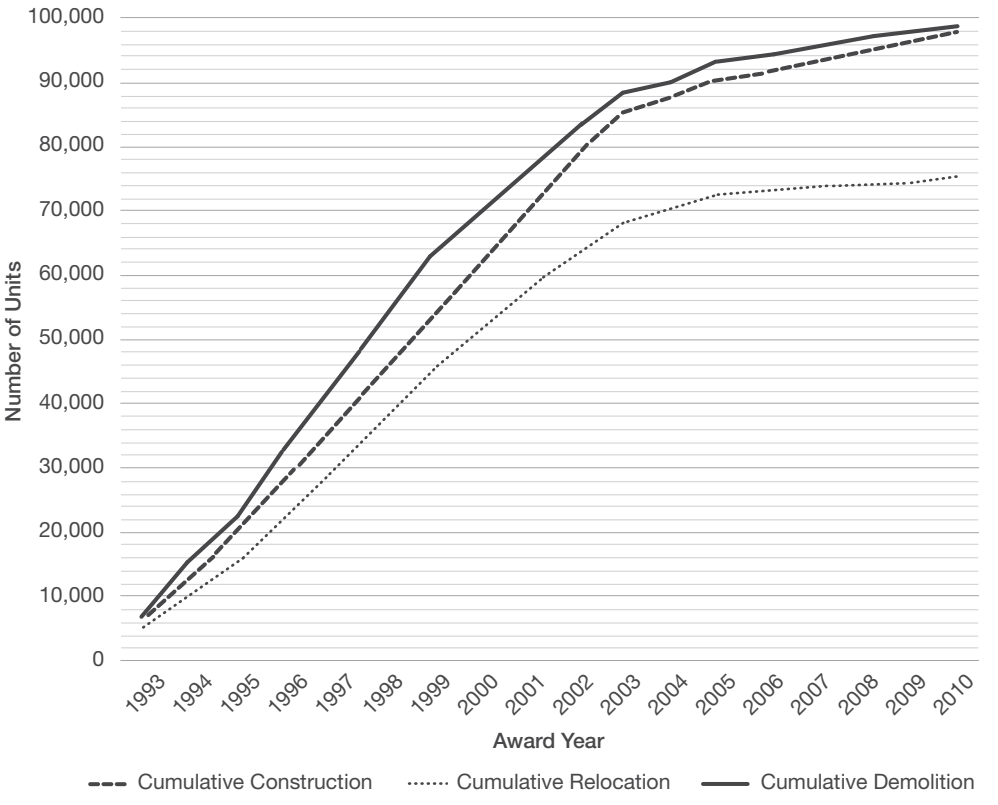
Decrease in production over time. Starting in 2003, appropriations for the HOPE VI program were dramatically reduced, causing a decrease in demolition and construction of units (see exhibit 1). As a result, HOPE VI program grantees demolished and constructed more units in the first 10 years of the program than in the remainder of the program. Likewise, grantee award cohorts after 2003 did not produce as many units for rent or for sale as did those before 2003 (see exhibit 2).

[6] Although new grants through the HOPE VI program were discontinued in 2010, many existing grantees are still completing the remaining construction and occupancy phases of their grants.

[7] The median size of a HOPE VI development is 319 units.

Exhibit 1

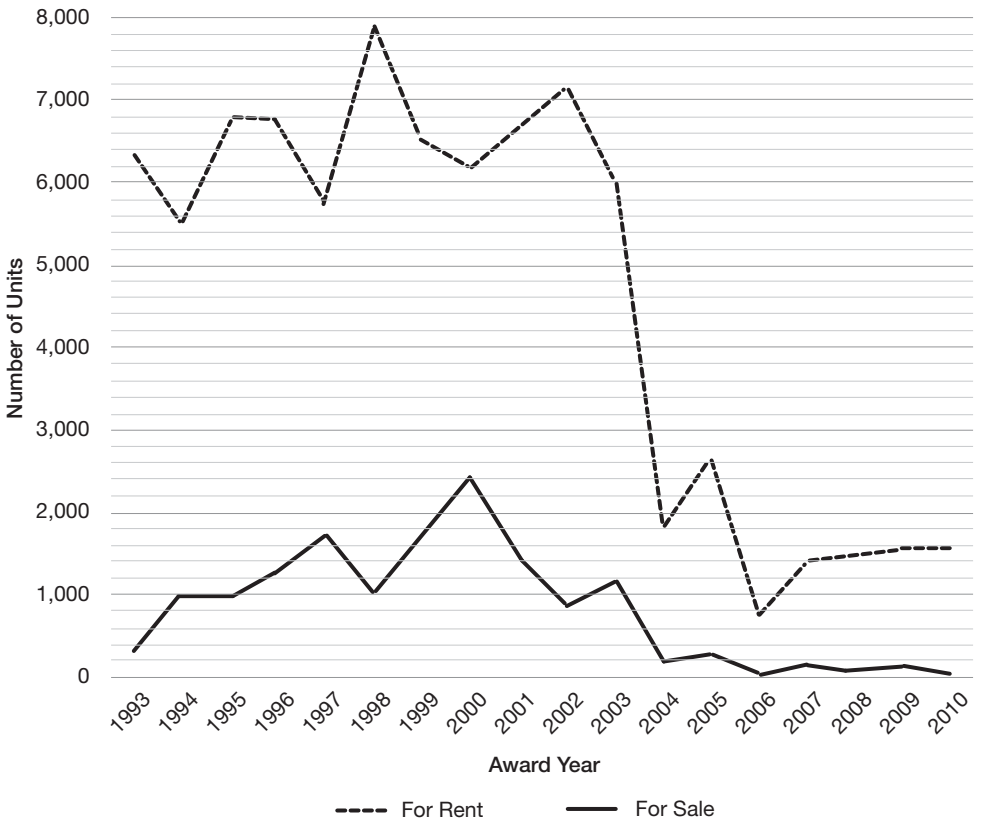
Cumulative Redevelopment Progress by HOPE VI Grant Award Year (N=18)



Note: In the award year information for 2011, only 12 units were constructed; that information was excluded.

Exhibit 2

Rental and Homeownership Unit Production by HOPE VI Grant Award Year (N=18)



New Information and Insights

Our descriptive analysis confirmed much of the knowledge on the HOPE VI program, and we identified some new information and insights about the program that have implications for future mixed-income housing development.

Most units were built as ultimately agreed to with HUD. Sites were generally able to meet the construction goals established in their most recent formal agreements with HUD. Because the quarterly reports to which we had access were overwritten with the most recent projection data, we were not able to analyze original site projections in this study. Overall, 88.3 percent of the units projected in the most recent grantee agreements have been produced. Among the 259 sites we analyzed, 66 (25.5 percent) produced fewer than their most recent projections and 193 (74.5 percent) produced the number of units planned.

Relatively more rental units were produced as planned. As described in exhibit 3, a higher percentage of the planned rental units (91.9 percent) was produced than homeownership units

(72.7 percent). Also, a higher percentage of the projected public housing units (94.7 percent) than affordable units (87.1 percent) and market-rate units (71.8 percent) units was produced.

Exhibit 3

Comparison of Projected and Actual Unit Production by HOPE VI Grantees

		Projected*	Actual	Percent
Rental	Public Housing	53,226	49,949	93.8
	Affordable	26,674	23,899	89.6
	Market	9,786	8,530	87.2
	Total	89,686	82,378	91.9
Homeownership	Public Housing	5,594	5,369	96.0
	Affordable	6,607	5,080	76.9
	Market	8,459	4,562	53.9
	Total	20,660	15,011	72.7
Grand Total	Public Housing	58,820	55,318	94.0
	Affordable	33,281	28,979	87.1
	Market	18,245	13,092	71.8
	Total	110,346	97,389	88.3

*Notes: * "Projected" refers to the most recent projection numbers in the HOPE VI quarterly reports, which have been updated over time as agreements with HUD have been finalized. Figures for public housing units include both new construction and rehabilitation.*

Redevelopment activities included a mix of public housing, affordable (LIHTC) units, and unsubsidized (market-rate) housing. Ultimately, the redevelopment side of the HOPE VI program focused on producing a mix of housing units with varying affordability levels, including replacement public housing, with units subsidized by LIHTC with more moderate affordability, and a relatively limited number of market-rate units. Although HOPE VI production resulted in a decrease of physical public housing units, the redevelopments usually did not include a large proportion of truly market-rate, unsubsidized housing; instead they mainly included subsidized housing, replacing 85.5 percent of the original public housing with units that were affordable for low- and moderate-income residents. An important outstanding question is, how much affordability was provided by the non-public housing "affordable" units, and what was the economic status of residents in those units? The data available in the quarterly reports did not enable us to explore that question, and subsequent research would be needed to assess income data, depending on availability, at redeveloped sites.

The distribution of mix types across the 259 HOPE VI developments demonstrates that affordable housing was a priority across the sites (see exhibit 4). To compare the income mix of HOPE VI projects, we adopted Vale and Shamsuddin's 2014 four-category typology: Narrow Low-Income, Polarized Bimodal, All But The Poorest, and Broad Continuum.[8] We assigned each HOPE VI development project to one of Vale and Shamsuddin's income mix categories based on the following criteria: Narrow Low-Income developments have no more than 15 percent market-rate

[8] Vale and Shamsuddin revised this typology in their 2017 article, "All Mixed Up: Making Sense of Mixed-Income Housing Developments" published in the *Journal of American Planning Association* (Vale and Shamsuddin, 2017). Our results using their new typology are quite similar.

units, Polarized Bimodal developments have no more than 15 percent affordable units and thus mainly have units at the high and low ends of the income mix, All But The Poorest developments have no public housing units, and Broad Continuum developments have at least 15 percent of all three unit types (public housing, affordable, and market rate). Most redevelopments had a Narrow Low-Income mix (*n*=179; 69.1 percent), followed by Broad Continuum (*n*=66; 25.5 percent) and Polarized Bimodal (*n*=13; 5.0 percent). Only one site had an All But The Poorest mix.[9] Exhibit 5 shows the differences in average unit production and their proportion by mix type. On average, the Narrow Low-Income developments had 97.8 percent public housing and affordable units, with only 2.2 percent market-rate units. The average proportion of market-rate units was 26.3 percent in the Broad Continuum developments and 33.7 percent in the Polarized Bimodal developments.

Exhibit 4

Income Type Mix in 259 HOPE VI Developments

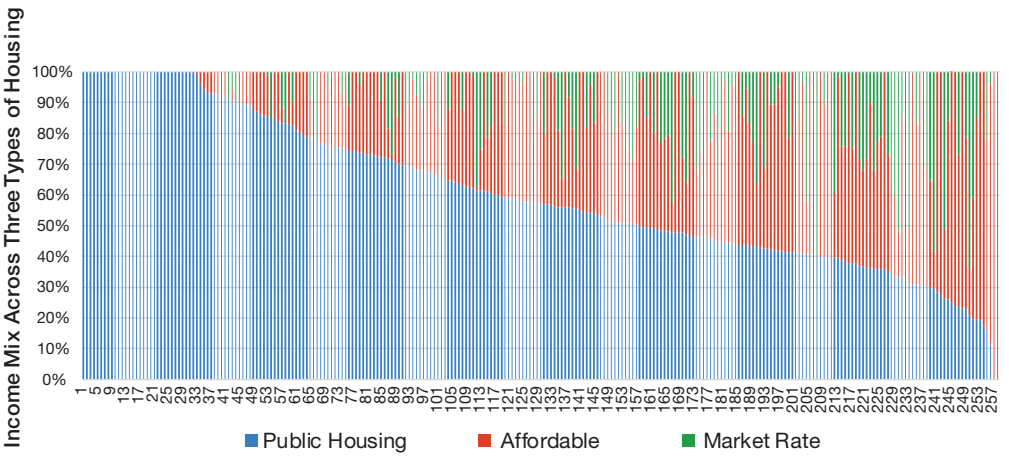


Exhibit 5

Comparison of Housing Types by Mix Type in the HOPE VI Program (N=258)

		Public Housing		Affordable		Market Rate		Total	
	n	M (SD)	%	M (SD)	%	M (SD)	%	M (SD)	%
Narrow Low-Income	179	217 (184)	67.8	101 (110)	30.0	8 (20)	2.2	326 (212)	100.0
Broad Continuum	66	195 (103)	42.1	158 (122)	31.6	140 (112)	26.3	492 (285)	100.0
Polarized Bimodal	13	283 (173)	60.9	37 (54)	5.4	179 (223)	33.7	500 (391)	100.0
Total	258	214 (167)	60.9	112 (115)	29.2	51 (99)	10.0	377 (255)	100.0

SD = standard deviation.

Notes: Only one site had an All But The Poorest mix type and was excluded. In all the tables, n refers to the sample of HOPE VI sites and M refers to the mean number of units at those sites.

Dominance of mixed-tenure developments. Although the HOPE VI program produced more rental units than homeownership units, a majority of sites (*n*=184; 71.0 percent) were mixed tenure, with both rental and homeownership units. Mixed-tenure sites tended to be larger (M=402) than sites

[9] Prospect Plaza in New York City included market-rate rental and affordable for-sale units and no public housing.

without homeownership units (M=312) and included a higher proportion of affordable and market-rate units (see exhibit 6). Overall, 82,378 units (84.6 percent) of the HOPE VI units produced have been for rental housing, and 15,011 units (15.4 percent) have been designated for homeownership.

Exhibit 6

Comparison of Housing Types by Tenure Mix in the HOPE VI Program (N=259)

		Public Housing		Affordable		Market Rate		Total*	
	n	M (SD)	%	M (SD)	%	M (SD)	%	M (SD)	%
Mixed Tenure	184	220 (178)	58.5	122 (118)	30.2	60 (111)	11.4	402 (272)	100.0
Rental Only	75	197 (137)	66.0	87 (104)	27.1	28 (56)	7.0	312 (196)	100.0
Total	259	214 (167)	60.6	112 (115)	29.3	51 (99)	10.1	376 (255)	100.0

SD = standard deviation.

Note: The total percentages in this chart may not add to 100 percent due to rounding.

Public housing homeownership units. The HOPE VI program had a greater focus on public housing homeownership than many might realize. About 10 percent of the replacement public housing units were designated for homeownership (5,369 units). Among the 259 sites analyzed here, a relatively high number (*n*=105; 40.5 percent) produced public housing homeownership units. Among those, 34 sites produced 50 or more public housing homeownership units. It is an untold dimension of the HOPE VI program that warrants closer investigation by researchers.

Variation in timeframes of redevelopment progress. Timeframes for relocation, demolition, construction, and occupancy ranged from weeks to years (see exhibit 7). We found that the relocation phase took the longest amount of time on average (M=694 days), followed by construction (M=667), demolition (M=516), and occupancy of the developments (M=260).[10] Timeframes for redevelopment varied among sites. On average, as would be expected, larger sites reported more days for each redevelopment phase ($M_{relocation}$=829, $M_{demolition}$=611, $M_{construction}$=706, and $M_{occupancy}$=293) than smaller sites ($M_{relocation}$=539, $M_{demolition}$=400, $M_{construction}$=620, and $M_{occupancy}$=218). Among smaller sites, the longest phase was construction, but among larger sites, the longest phase was relocation. Longer timeframes for relocation, construction, and occupancy directly influenced residents who hoped to return to the newly constructed development.

Exhibit 7

Average Duration of HOPE VI Relocation, Demolition, Construction, and Occupancy (days)

	N	M (SD)	Median	Min if >0	Max
Relocation	234	694 (623)	466	33	3,643
Demolition	218	516 (677)	268	23	4.318
Construction	216	667 (336)	580	208	2,949
Occupancy	206	260 (367)	172	13	3,939

SD = standard deviation.

Note: Among the 259 sites, 234 sites provided relocation timeframes, 218 sites provided demolition timeframes, 215 sites provided construction timeframes and 206 sites provided occupancy timeframes.

[10] Redevelopment phases are not necessarily sequential; one can start after another ends, but they can also be concurrent and overlapping.

Funding In our descriptive analysis, we confirmed information about HOPE VI program funding and provided insights on private funding leveraged from the HOPE VI program.

Existing Knowledge Confirmed

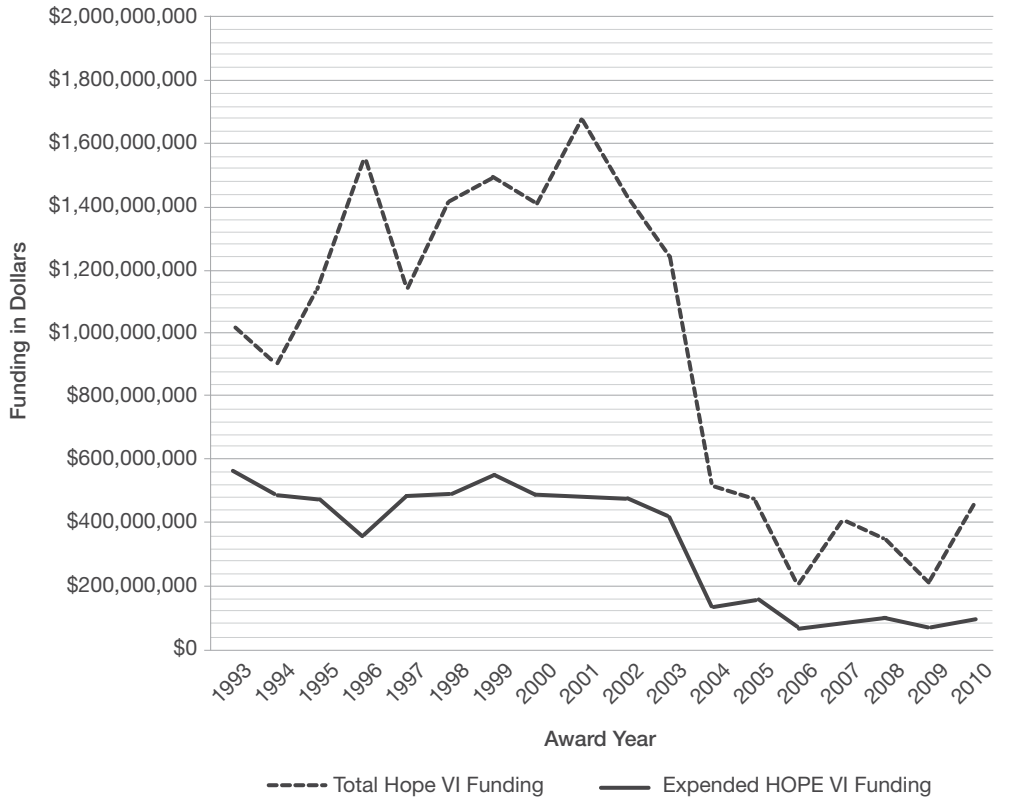
Funding allocated for the HOPE VI program was successfully expended and decreased over time.

Most budgeted funding was expended. Most (95.3 percent) of the budgeted HOPE VI grant funding was spent (M $_{HOPE\ VI\ budgeted}$=$24.1M and M $_{HOPE\ VI\ expended}$=$22.9M). HOPE VI grantees spent 85 percent of total budgeted funds for redevelopment, which includes both federal grants (HOPE VI and other federal agency funds) and money from private sources (M $_{total\ budgeted}$=$76.9M and M $_{total\ expended}$=$65.3M).

Decrease in funding over time. The annual expended HOPE VI funding ranged from $66.2M (in 2009) to $566.1M (in 1993), with an average of $332.7M and a median of $445.9M. Funding expended in the HOPE VI program decreased over time (see exhibit 8). The total funding expended was highest in the 2001 award year and dramatically decreased between 2003 and 2004. The lowest funding was expended in 2009 and remained relatively low through 2010.

Exhibit 8

HOPE VI Total and Expended Funding by Award Year



New Information and Insights

Below, we explore the private funding that was leveraged by the HOPE VI program.

Leveraging redevelopment funding. HOPE VI funds were used to leverage significant amounts of other public and private funds. The approximately $6 billion of HOPE VI funds leveraged approximately $11 billion of non-HOPE VI funds to complete redevelopment for a total of approximately $17 billion expended. For every dollar of HOPE VI funds expended, about $1.80 was leveraged for the redevelopment projects.

Sites targeting an economic mix of residents consistently leveraged more funding (see exhibit 9). Among the four types of income mix, the Narrow Low-Income developments (M *total funding*=$57.0M), which have only public housing and affordable units, generally had less funding and were not able to leverage as much in additional funds (Polarized Bimodal: M *total funding*=$78.6M; Broad Continuum: M *total funding*=$86.9M). Sites with a market-rate component were able to leverage more funds, and the Broad Continuum developments were able to leverage the most funds because those developments consistently incorporated market-rate units.

Exhibit 9

Leverage Ratio of Funding by Mix Type in the HOPE VI Program (N=258)

		Unit Production	HOPE VI Funding	Total Funding	Leverage Ratio
	n	M (SD)	M (SD)	M (SD)	
Narrow Low-Income	179	326 (212)	$22.5M (11.4M)	$57.0M (36.5M)	1:1.5
Broad Continuum	66	492 (285)	$23.9M (94.8M)	$86.9M (50.3M)	1:2.6
Polarized Bimodal	13	500 (391)	$27.7M (15.9M)	$78.6M (43.8M)	1:1.8
Total	258	377 (255)	$23.1M (11.2M)	$65.7M (42.8M)	1:1.8

SD = standard deviation.

The leverage ratio also varied among the regions designated by HUD (see exhibit 10). Compared with other regions, the Rocky Mountain region showed the highest leverage ratio: About $3 was leveraged for every dollar of HOPE VI funding. This region, however, obtained the smallest amount of HOPE VI funding and constructed the smallest number of units. On the other hand, the Southeast-Caribbean region received the largest amounts of HOPE VI funding and produced the largest number of units with leverage ratio of 1:1.7, close to the overall grantee average (1:1.8).

Exhibit 10

Leverage Ratio of HOPE VI Funding by HUD Region

Name	Number of Grants		Unit Production	HOPE VI Funding	Total Funding	Leverage Ratio
	n	%	Sum	Sum	Sum	
Rocky Mountain	5	1.9	1,740	$87.8M	$380.4M	1:3.3
Northwest	12	4.6	6,654	$323.9M	$1.3B	1:3.0
New England	15	5.8	4,060	$346.6M	$1.2B	1:2.5
New York-New Jersey	24	9.2	9,308	$586.7M	$1.9B	1:2.2
Pacific	18	6.9	5,665	$391.5M	$1.2B	1:2.1
Southeast-Caribbean	75	28.8	31,815	$1.6B	$4.3B	1:1.7
Mid-Atlantic	43	16.5	12,712	$903.4M	$2.5B	1:1.7
Midwest	41	15.8	14,389	$1.0B	$2.6B	1:1.6
Southwest	19	7.3	7,767	$488.2M	$1.2B	1:1.4
Great Plains	8	3.1	3,279	$179.1M	$420.8M	1:1.3
Total	260	100.0	97,389	$6.0B	$17.0B	1:1.8

Return and Relocation

We confirmed that many public housing residents did not return to the newly redeveloped housing, and we now look more closely at the trends in return rates across the HOPE VI developments.

Existing Knowledge Confirmed

The rates of return were low across HOPE VI developments.

Low return rate. The grantee report data confirmed what has become conventional knowledge about the HOPE VI program: a relatively low percentage of all occupied units within HOPE VI developments were occupied by the original residents returning to the developments. At least two ways are used to explore how original residents have benefited from the redevelopments. One method looks at redeveloped units as the denominator and the other method looks at the original household population as the denominator. Of the 96,476 units that have been produced *and* occupied, only 19,993 units (20.7 percent) were occupied by the original tenants at each development. That number can be considered the "proportion occupancy by original residents" rate. The more familiar way to consider return rate (the "proportion of original residents that return" rate) is to examine the percentage of original households that have returned to occupy new units available after redevelopment. Of the households originally relocated, across sites an average of only 27.6 percent (with a median of 18.2 percent) returned to the new units.[11] The relatively substantial difference between the average and the median is due to a skew in the distribution toward the lower range. Even among the units *designated for public housing residents*, a low percentage was occupied by the *original* residents. Of the 55,318 units that were designated as replacement units for original tenants, only 36.1 percent was occupied by residents of the original development.

[11] This figure may continue to increase over time as remaining units are built in sites that are not yet complete and *if* more original residents return to those units.

New Information and Insights

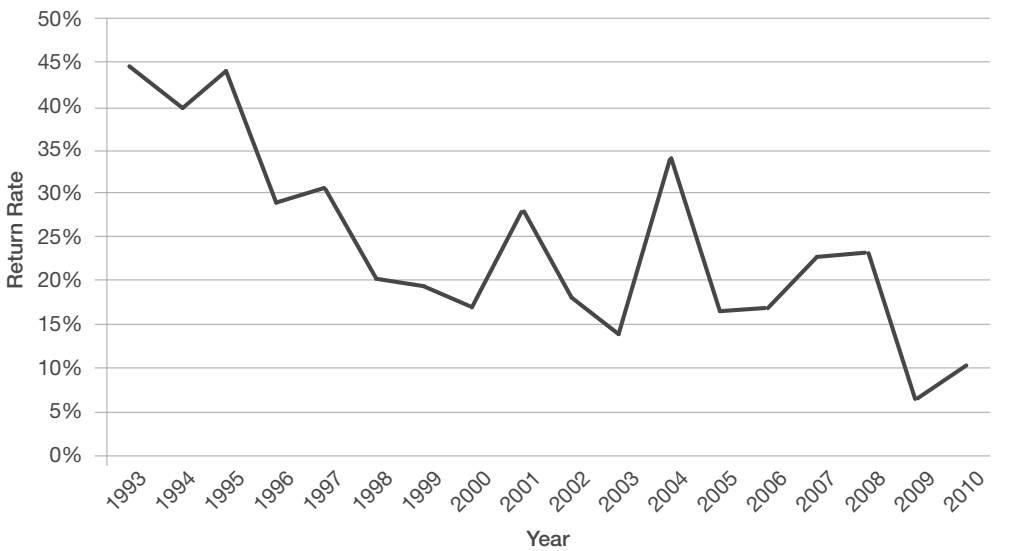
We now discuss how the rates of return were low across HOPE VI developments despite the variance in the income mix types in each development. We also found that these rates remained low over time.

Consistently low return rates. Return rates do not change significantly when the 32 100-percent public housing replacement developments are removed from the analysis (the average becomes 26.8 percent and the median becomes 17.5 percent), suggesting, surprisingly, that return rates were relatively low even when the developments did *not* mix incomes and instead replaced *all* of the public housing on site. Also, the return rates were almost identical for both larger (M=27.6 percent) and smaller (M=27.5 percent) sites, indicating that the return rates were very low regardless of the size of developments.

Declining return rates over time. As indicated in exhibit 11, the average return rates by award year ranged from 6.3 percent (2009) to 44.7 percent (1993). Return rates mostly declined between 1993 and 2000, with spikes in 2001 and 2004 followed by a sharp drop-off in 2005 and another relatively small peak in 2008. Newer sites have, thus far, experienced lower return rates (21.1 percent) than older sites (33.8 percent), which may be a function of time and a lag in construction and occupancy. The return rate at older sites, however, likely indicates the upper end of the average that can ultimately be expected.

Exhibit 11

Return Rates by HOPE VI Award Year



Community and Supportive Services

Some HOPE VI developments also received grants for Community and Support Services to support residents through programs and services that focused on helping residents get jobs, access education, improve their health and wellness, and even cover basic needs. We analyzed grantee reports of CSS caseloads pre- and post-revitalization regarding resident enrollment in and completion of services. We then conducted a comparative analysis of the following eight CSS programs by tenure mix, size, and age: employment preparation/placement/ retention, job skills training programs, high school or equivalent education, English as a Second Language (ESL) courses, childcare, transportation assistance, counseling programs, and substance abuse programs. We also compared broader enrollment and completion goals data for employment, economic development, and homeownership at each development.

New Information and Insights

We share new information and insights on the use of CSS across HOPE VI developments by residents who remained on site at a development during redevelopment and residents who were relocated and then returned once redevelopment was complete. We compared the use of CSS funds by developments with different types of income mixes. We found that most of the basic goals outlined by developments receiving CSS funds were achieved, but broader goals of program enrollment and completion were not.

Overall provision of CSS. CSS served both original public housing residents and new incoming residents between the ages of 19 and 64 who moved to redevelopment sites and original residents who continued to live off site. Of the eight main CSS programs, 249 sites (95.7 percent) provided at least one, and 93 sites (35.8 percent) provided all eight services. Those programs had an average caseload of 1,226 cases and a median of 767 cases.

Low proportion of remaining or returning individuals in CSS caseload. Among former public housing residents served in CSS programs (N=54,453 cases), only 31.8 percent were served while they remained on site (n=5,869; 10.8 percent) or returned after redevelopment (n=11,473; 21.1 percent). The other 68.2 percent who were served were relocated to other public housing (n=13,680; 25.1 percent), to housing not receiving any HUD assistance (n=5,423; 10.0 percent) or to housing using Section 8 vouchers (n=18,008; 33.1 percent).

Comparative analysis of CSS. The number of available services varied by tenure mix and the size and age of the redevelopment. Mixed-tenure sites (M = 1,301) had more residents enrolled in the CSS program than did rental-only sites (M = 1,044). Among the 184 mixed-tenure sites, only 4 sites (2.2 percent) did not provide any of the eight CSS programs (9.3 percent in rental-only sites). Larger sites (M = 1,626) tended to provide more CSS support to residents than did smaller sites (M = 830). Among the 130 larger sites, only 2 sites (1.5 percent) did not provide any of the eight programs (7.0 percent in smaller sites). Finally, on average, older sites (M = 1,415) have provided more CSS support to residents than have newer sites (M = 1,046); however, a slightly higher proportion of older (4.7 percent) than newer (3.8 percent) sites did not provide any of the eight programs.

Most program enrollment goals achieved. Overall, the HOPE VI sites achieved their goals for enrollment in the eight CSS programs (see exhibit 12). On average, employment preparation / placement /retention had the largest number of enrollments, followed by transportation assistance, counseling, job skills training, and childcare. Still, a notable observation is that substance abuse programs, high school or equivalent education, and language courses enrolled the fewest number of residents.

Exhibit 12

Enrollments of Caseloads Receiving HOPE VI CSS Programs

	Enrollment		Goal	
	M*	Median	M	Median
Employment Preparation/ Placement/Retention	333	170	179	100
Transportation Assistance	291	110	136	70
Counseling Programs	239	87	146	444
Job Skills Training Programs	135	77	100	70
Childcare	116	61	97	60
High School or Equivalent Education	71	39	52	32
Substance Abuse Programs	24	6	22	10
English as a Second Language (ESL) Courses	16	0	15	0

*Note: * Mean number of enrollees per HOPE VI site.*

Broader enrollment and completion goals not met. In our comparison of the broader enrollment and completion goals for employment, economic development, and homeownership efforts, the data revealed that some of the goals for those areas of focus were achieved but others were not. In the area of employment, despite exceeding the number of projected total new job placements, which includes all types of employment (full time, part time, seasonal, and temporary), the number of participants currently employed and those employed for 6 months were far below their enrollment goals. In economic development, entrepreneurship training exceeded the enrollment goal but fell short on its completion goal. There were also fewer resident-owned businesses and residents employed by those businesses than projected. For homeownership, the average number enrolled in homeownership counseling exceeded the goal; however, the number of residents who completed counseling was lower than the intended goal, and the number of participants who purchased a home was also less than expected.

Discussion

Through a detailed descriptive analysis of production and financing data from 1993 to 2014, this paper confirms previous expectations and findings regarding the HOPE VI program and provides new information and insights. In addition to exploring the income and tenure mix of the units produced, we answered questions about the nature of HOPE VI reoccupancy, timeframes of production stages, financing, and the CSS programs available to residents of HOPE VI developments. Our findings confirm many longstanding concerns about the program, such as the reduction in the overall public housing stock and the low return rates of original residents. We also

uncovered some compelling insights about the program, such as the overall focus on low-income housing production and the extent of the production of public housing homeownership units.

Overall, the main HOPE VI redevelopment housing product was replacement public housing units, complemented by the inclusion of affordable housing units and, in lower proportions, market-rate units. Almost one-half of the redevelopments included no market-rate housing. The vast majority of production was rental, rather than homeownership, units. Even within the properties developed to include homeownership, a substantial number of units were designated for public housing. Sites with rental-only units tended to be smaller than mixed-tenure sites, and included fewer affordable and market-rate units.

Using Vale and Shamsuddin's (2014) income mix typology, we found that well over two-thirds of the sites have a Narrow Low-Income mix of public and affordable units and about one-fourth have a Broad Continuum mix of public housing, affordable housing, and market-rate units. Our analysis also shows that the majority of HOPE VI projects have focused primarily on producing housing for low-income households, replacing more than one-half of the physical public housing units demolished and building out another one-third of all new units as subsidized affordable units.

Looking at production over time, there was a greater focus on producing a more diverse income mix at developments as the program progressed. The greatest demolition and construction of units occurred in the first 10 years of the HOPE VI program. Starting in 2003, demolition, relocation, and construction leveled off and production of homeownership units dropped sharply.

The data we received included only the most recent agreed-upon projections between the grantees and HUD, thus we were not able to analyze original production projections against actual production. In most cases, grantees met their most recent projected production goals. As would be expected given the timing of the Great Recession and the associated housing market crash, where production goals were not met, it was homeownership units that most often fell short, particularly market-rate homeownership; thus, the intended mix of incomes was not achieved at many sites. This can be interpreted in at least two ways: for those who feared that HOPE VI mixed-income projects would create environments in which public housing residents in particular and low-income households in general were a significant minority, that has not come to pass; on the other hand, to the extent that the success of the mixed-income strategy depended on a broader mix of income and a critical mass of higher income residents, including homeowners, that objective was not widely achieved through the HOPE VI program.

The number of phases and the time it took for each phase of development to begin and end had a major impact on how the HOPE VI program affected residents. The long periods between relocation, construction, and occupancy meant an extended waiting period for those residents who hoped to move back to the new developments and likely contributed to the low return rates.

HOPE VI funds were used to leverage significant amounts of other public and private funding. The $6 billion of expended HOPE VI funds leveraged $11 billion in nonfederal funds—$17 billion expended from all sources. For every dollar of HOPE VI funds, about $1.80 was leveraged. Narrow Low-Income developments, which have only public housing and affordable units, generally had

less funding and were not as able to leverage additional funds as were projects that included market-rate or homeownership units.

The CSS programs provided services to original residents of the public housing developments and to families who later moved into the revitalized site. We analyzed caseload output data, which included goals and outputs for each service. The services with the highest number of enrollments were employment preparation/placement/retention, transportation assistance, counseling, job skills training, and childcare. Overall, CSS programs met or exceeded most enrollment goals, although, on average, the sites fell short of completion goals for high school or equivalent education. The CSS programs also fell short on employment retention and entrepreneurship support goals.

Study Limitations

This study had some important limitations. The findings are drawn from HOPE VI administrative project reports from 1993 to 2014 and include quantitative data on production, financing, and the CSS program and thus provide a numerical documentation of the program. Given the limited information available in the reports, however, we were constrained to basic descriptive and comparative statistical analyses. A text field for qualitative comments was included in the CSS reports, but those data were of limited depth, consistency, and utility. Also, although the reports provide "projected" estimates regarding financing and unit production, those estimates were updated over the course of the program and thus do not provide accurate baseline projections for comparison.

Implications for Research, Policy, and Practice

HOPE VI produced predominantly public housing and subsidized rental housing, with less production of market-rate rental and homeownership units than had been intended. Although it produced higher quality replacement public housing in more mixed-income environments, the HOPE VI program substantially decreased the stock of physical public housing units. Furthermore, extended construction and occupancy timeframes and stringent reoccupancy requirements may have prevented many residents from successfully returning to the revitalized housing. HOPE VI funds successfully leveraged substantial levels of private funds for development. Due to the decrease in federal funding over time, however, unit production levels were not sustained over the life of the program. CSS programs generally achieved many of the stated enrollment goals but were limited in the number of residents they could serve and fell short of the stated goals in several important areas, such as high school equivalency achievement. These findings suggest a number of implications for policy and further research.

Implications for Research

Given the continued national and local investment in mixed-income redevelopment, a number of research topics could be pursued, either from further data collection and analysis of HOPE VI projects or from research on Choice Neighborhoods grantees and other emerging mixed-income redevelopments.

1. To fully understand the projected versus actual unit production through the HOPE VI program—and thus inform future development efforts and negotiations—an analysis using additional information about *original* projected goals may be useful.

2. A deeper understanding of the income levels of residents in each band of subsidy would provide far better information about which households are benefiting from the housing and what the income mix continuum looks like in various contexts.

3. Stronger performance measurement and reporting of resident outcomes, such as those that follow, would provide a critical knowledge base on the upward mobility outcomes of residents in diverse mixed-income contexts and on those who did not return:

    a. The results of various levels and combinations of supportive services.

    b. How the outcomes for residents compare across different levels of site income mix, for example, Narrow Low-Income developments compared with Broad Continuum developments.

    c. How the outcomes for residents compare by level of physical integration among different income groups within a development.

    d. How outcomes for original residents who return compare with those of residents who do not return.

    e. How outcomes compare among original residents who return, newcomers to the public housing replacement units, residents of the affordable units, and residents of the market-rate units.

    f. How resident outcomes vary with factors such as the size of the development and the tenure mix at the development.

4. The nature, outcomes, and effect of public housing homeownership programs is an underexamined area, including which residents qualified and participated and their outcomes.

5. The focus here, given the data available, was on the developments and not on the broader neighborhood context. The surrounding neighborhoods, however, are an important dimension to be included in future research both as a context that shapes the strategies and outcomes in specific developments and as a unit of change brought about by the redevelopment. Analyzing how site income mixes vary relative to the characteristics of the neighborhood context would also be interesting.

Implications for Policy and Practice

Stepping back, we can draw from our findings to suggest some key areas for continued policy focus and improvement.

1. **Balancing dual priorities: ending segregation and concentrated poverty and increasing affordable housing for the poor.** This is an enduring strategic tension in this arena of housing policy. There are legitimately conflicting policy imperatives here. It would be ideal to be able to produce more high-quality public housing *and* facilitate access to more vibrant, socioeconomically diverse neighborhoods. No housing policy has yet fully resolved this dual challenge and the major policy approaches—mixed-income redevelopment, housing choice vouchers, and the Rental Assistance Demonstration program—each have operational advantages and important downsides. Policymakers must continue to be as intentional and comprehensive as possible in their efforts to use existing public resources to maximize the provision of housing for the poor while leveraging private-sector resources to generate investment in housing developments and their surrounding neighborhoods. Clarity about the intended balance of these conflicting goals and vigilant accountability will be key.

2. **Managing the market risks of privatization.** The mixed-finance approach to HOPE VI and the intentions of including market-rate rental and homeownership made the redevelopment efforts extremely vulnerable to market conditions. The result—evident in the early phases of the program and then exacerbated by the housing market crash—was extended delays in unit production and, ultimately, substantial shortfalls in the production of market-rate units. Policymakers should consider to what extent the market-rate shortfalls affected program "success" in various local contexts and examine ways to offset the market risk; for example, might smaller redevelopment phases, increased capacity building, and engagement of nonprofit entities and community-based organizations be beneficial?

3. **Increasing return rates.** Clearly a major program shortcoming was the limited proportion of original residents who benefited from the new, higher quality housing and living environments. Although some proportion of the nonreturners possibly used the relocation opportunity for upward mobility to a low-poverty neighborhood, the literature suggests mixed results for those who did not return and for voucher holders in general. The policy implications include increased resources, attention to the relocation support process, and strategies to make return easier, such as smaller redevelopment phases and phased relocation on site or in close proximity.

4. **Program enrollment is not sufficient: Providing the requisite support services for positive outcomes.** The CSS data analyzed here provide a severely limited purview into the details of service provision and results. The program reports clearly show, however, that although the HOPE VI grantees were able to exceed their enrollment goals in many cases—indicating that they had achieved the objectives of linking with local partners and programs that were offering the forms of support needed—evidently those enrollments generally did not turn into sustained engagement or meaningful outcomes for participants. What is the scope, quality, and duration of support needed to help residents affected by a mixed-income redevelopment move toward self-sufficiency?

Acknowledgments

We thank Jenni Bartholomew for her editorial assistance. We thank our colleagues Dr. Claudia Coulton and Dr. David Crampton for their advisory roles on this research study. We thank research assistants Biwen Liu, Siyang Sun, and Michael Salwiesz for their data compilation work. We thank our partners at the Resilient Cities and Housing Initiative at the Massachusetts Institute of Technology, led by Dr. Lawrence Vale and Dr. Shomon Shamsuddin, for their collaboration and insights in extracting, compiling, and analyzing this data set. We thank HUD for the Research Partnership grant that made this study possible and in particular Ron Ashford and Lars Gnissen for their help with our orientation to the data set, Madlyn Wohlman-Rodriguez for her encouragement and responsiveness in supporting this research, and David Hardiman for his helpful comments on an earlier version of the article. The work that provided the basis for this paper was supported by funding under a grant with the U.S. Department of Housing and Urban Development, Office of Policy Development and Research (Instrument Number RP-14-OH-005). The substance and findings of the work are dedicated to the public. The authors are solely responsible for the accuracy of the statements and interpretations contained in this paper. Such interpretations do not necessarily reflect the views of the government.

References

Baron, Richard. 2009. "The Evolution of HOPE VI as a Development Program." In *From Despair to Hope: HOPE VI and the New Promise of Public Housing in America's Cities*, edited by Henry G. Cisneros and Lora Engdahl. New York: Brookings Institution: 31–47.

Brophy, Paul. C., and Rhonda N. Smith. 1997. "Mixed-Income Housing: Factors for Success," *Cityscape* 3 (2): 3–31.

Buron, Larry. 2004. *An Improved Living Environment? Neighborhood Outcomes for HOPE VI Relocatees*. Metropolitan Housing and Communities: A Roof Over Their Heads Policy Series, Brief No. 3. Washington, DC: Urban Institute.

Buron, Larry, Susan J. Popkin, Diane K. Levy, Laura G. Harris, and Jill Khadduri. 2002. *The HOPE VI Resident Tracking Study: A Snapshot of the Current Living Situation of Original Residents From Eight Sites*. Washington, DC: Urban Institute.

Center for Community Change. 2003. *A Hope Unseen: Voices From the Other Side of HOPE VI*. Washington, DC: ENPHRONT.

Chaskin, J. and Mark L. Joseph. 2015. *Integrating the Inner City: The Promise and Perils of Mixed-Income Public Housing Transformation*. Chicago, IL: University of Chicago Press.

Cisneros, Henry G., and Lora Engdahl, eds. 2009. *From Despair to Hope: HOPE VI and the New Promise of Public Housing in America's Cities*. New York: Brookings Institution.

Comey, Jennifer. 2004. *An Improved Living Environment? Housing Quality Outcomes for HOPE VI Relocatees*. Metropolitan Housing and Communities: A Roof Over Their Heads Series, Policy Brief No. 2. Washington, DC: Urban Institute.

Crowley, Sheila. 2009. "HOPE VI: What Went Wrong." In *From Despair to Hope: HOPE VI and the New Promise of Public Housing in America's Cities*, edited by Henry G. Cisneros and Lora Engdahl. New York: Brookings Institution: 229–247.

Fraser, James C., and Edward L. Kick. 2007. "The Role of Public, Private, Non-Profit and Community Sectors in Shaping Mixed-Income Housing Outcomes in the U.S.," *Urban Studies* 44 (12): 2357–2377.

Goetz, Edward. G. 2003. *Clearing the Way: Deconcentrating the Poor in Urban America*. Washington, DC: Urban Institute Press.

Holin, Mary Joe, Larry Buron, Gretchen Locke, and Alvaro Cortes. 2003. I*nterim Assessment of the HOPE VI Program: Cross-Site Report*. Washington, DC: U.S. Department of Housing and Urban Development.

Joice, Paul. 2017. "HOPE and Choice for HUD-Assisted Households," *Cityscape* 19 (3): 449–473.

Joseph, Mark L., and Robert J. Chaskin. 2012. "Mixed-Income Developments and Low Rates of Return: Insights From Relocated Public Housing Residents in Chicago," *Housing Policy Debate* 22 (3): 377–406.

Joseph, Mark L., Robert J. Chaskin, and Henry S. Webber. 2007. "The Theoretical Basis for Addressing Poverty Through Mixed-Income Development," *Urban Affairs Review* 42 (3): 369–409.

Katz, Bruce. 2009. The Origins of HOPE VI. In *From Despair to Hope: HOPE VI and the New Promise of Public Housing in America's Cities*, edited by H.G. Cisneros and L. Engdahl. New York: Brookings Institution: 15–29.

Khadduri, Jill, Carissa Climaco, Kimberly Burnett, Laurie Gould, and Louise Elving. 2012. *What Happens to Low-Income Housing Tax Credit Properties at Year 15 and Beyond?* Bethesda, MD: Abt Associates, Inc.

Kleit, Rachel G. 2005. "HOPE VI New Communities: Neighborhood Relationships in Mixed-Income Housing," *Environment and Planning A: Economy and Space* 37 (8): 1413–1441.

Lees, Loretta. 2008. "Gentrification and Social Mixing: Towards an Inclusive Urban Renaissance?" *Urban Studies* 45 (12): 2449–2470.

Levy, Diane K., and Mark Woolley. 2007. *Relocation is Not Enough: Employment Barriers Among HOPE VI Families*. HOPE VI: Where Do We Go from Here? Brief 6. Washington, DC: The Urban Institute.

Marquis, Gerald P., and Soumen Ghosh. 2008. "Housing Opportunities for People Everywhere (HOPE VI): Who Gets Back In?" *The Social Science Journal* 45 (3): 401–418.

Naparstek, Arthur J., Susan R. Freis, G. Thomas Kingsley, Dennis Dooley, and Howard E. Lewis. 2000. *HOPE VI: Community Building Makes a Difference*. Washington, DC: U.S. Department of Housing and Urban Development.

National Commission on Severely Distressed Public Housing (NCSDPH). 1992. Final Report to the Congress and the Secretary of Housing and Urban Development. Washington, DC: U.S. Government Printing Office.

National Housing Law Project. 2002. *False Hope: A Critical Assessment of the HOPE VI Public Housing Redevelopment Program*. Oakland, CA: National Housing Law Project.

Oakley, Deirdre, James Fraser, and Joshua Bazuin. 2015. "The Imagined Self-Sufficient Communities of HOPE VI: Examining the Community and Social Support Component," *Urban Affairs Review* 51 (5): 726–746.

Pendall, Rolf, Martin D. Abravanel, Audra Brecher, Alex Curley, Elizabeth Davies, Megan Gallagher, Monica Getahun, David Greenberg, Taryn Gress, Chantal Hailey, Leah Hendey, Reed Jordan, Mark Joseph, Amy Khare, G. Thomas Kingsley, Nancy Latham, Diane K. Levy, Brianna Losoya, Jen McGraw, Hannah Melnicoe, Jamie Pfluecke, Kathryn L.S. Pettit, Susan J. Popkin, Aesha Rashid, Hortencia Rodriguez, and Rachel Scheu. 2015. *Developing Choice Neighborhoods: An Early Look at Implementation in Five Sites*. Washington, DC: Urban Institute: 1–267.

Pendall, Rolf, and Leah Hendey. 2013. *A Brief Look at the Early Implementation of Choice Neighborhoods*. Washington, DC: Urban Institute: 1–14.

Polikoff, Alexander 2009. "HOPE VI and the Deconcentration of Poverty," In *From Despair to Hope: HOPE VI and the New Promise of Public Housing in America's Cities*, edited by H.G. Cisneros and L. Engdahl. New York: Brookings Institution: 65–81.

Popkin, Susan J., and Mary K. Cunningham. 2009. "Has HOPE VI Transformed Residents' Lives?" In *From Despair to Hope: HOPE VI and the New Promise of Public Housing in America's Cities*, edited by H.G. Cisneros and L. Engdahl. New York: Brookings Institution: 191–203.

Popkin, Susan J., Mary K. Cunningham, and M. Burt. 2005. "Public Housing Transformation and the Hard-to-House," Housing Policy Debate 16 (1): 1–24.

Popkin, Susan J., Bruce Katz, Mary K. Cunningham, Karen D. Brown, Jeremy Gustafson, and Margery A. Turner. 2004. *A Decade of HOPE VI: Research Findings and Policy Challenges*. Washington, DC: Urban Institute.

Popkin, Susan J., Diane K. Levy, and Larry Buron. 2009. "Has HOPE VI Transformed Residents' Lives? New Evidence From the HOPE VI Panel Study," *Housing Studies* 24 (4): 477–502.

Popkin, Susan J., Michael J. Rich, Leah Hendey, Chris Hayes, Joe Parilla, and G. Galster. 2012. "Public Housing Transformation and Crime: Making the Case for Responsible Relocation," *Cityscape* 14 (3): 137–160.

Santiago, Anna Maria, and George C. Galster. 2004. "Moving From Public Housing to Homeownership: Perceived Barriers to Program Participation and Success," *Journal of Urban Affairs* 26 (3): 297–324.

Tach, Laura. 2009. "More Than Bricks and Mortar: Neighborhood Frames, Social Processes and the Mixed-Income Redevelopment of a Public Housing Project," *City and Community* 8 (3): 273–303.

Turbov, Mindy. 2006. "Public Housing Redevelopment as a Tool for Revitalizing Neighborhoods: How and Why Did It Happen and What Have We Learned?" *Northwestern Journal of Law and Social Policy* 1 (1): 167–201.

U.S. Department of Housing and Urban Development (HUD). 2015. Choice Neighborhoods 2015 Grantee Report. Washington, DC: HUD. https://www.hud.gov/sites/documents/CNGRANTEEREPORT2015.PDF.

———. 2003. Guidance for PHAs Developing a Section 32 Homeownership Plan. Washington, DC: HUD. https://www.hud.gov/sites/documents/DOC_8104.PDF.

———. 2000. Community and Supportive Services for Original Residents. General Guidance for the HOPE VI Program (Draft). https://portal.hud.gov/hudportal/documents/huddoc?id=DOC_9840.pdf.

Vale, Lawrence J. 2019. *After the Projects: Public Housing Redevelopment and the Governance of the Poorest Americans*. New York: Oxford University Press.

———. 2013. *Purging the Poorest: Public Housing and the Design Politics of Twice-Cleared Communities*. Chicago, IL: University of Chicago Press.

Vale, Lawrence J., and Shomon Shamsuddin. 2017. "All Mixed Up: Making Sense of Mixed-Income Housing Developments," *Journal of the American Planning Association* 83 (1): 56-67.

———. 2014. All Mixed Up: Defining Mixed-Income in Public Housing Redevelopment. Working paper. Cambridge, MA: Massachusetts Institute of Technology.

Wexler, Harry J. 2001. "HOPE VI: Market Means/Public Ends—The Goals, Strategies, and Midterm Lessons of HUD's Urban Revitalization Demonstration Program," *Journal of Affordable Housing and Community Development Law* 10 (3): 195–233.

Wilson, William J. 1987. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago, IL: University of Chicago Press.

SpAM

SpAM (Spatial Analysis and Methods) presents short articles on the use of spatial statistical techniques for housing or urban development research. Through this department of Cityscape, the Office of Policy Development and Research introduces readers to the use of emerging spatial data analysis methods or techniques for measuring geographic relationships in research data. Researchers increasingly use these new techniques to enhance their understanding of urban patterns but often do not have access to short demonstration articles for applied guidance. If you have an idea for an article of no more than 3,000 words presenting an applied spatial data analysis method or technique, please send a one-paragraph abstract to rwilson@umbc.edu for review.

Where is the City's Center? Five Measures of Central Location

Matthew J. Holian
San Jose State University

Abstract

In this article, the author describes and evaluates five databases that provide different measures of city center location. He identifies research that has used each database, calculates average distances between the locations, presents case studies for two cities, and provides suggestions for analysts searching for appropriate measures of city center location. Among the findings are that the location of a city's city hall is a better proxy for the location of a city's central business district (CBD) than are other measures in current use.

Introduction

Researchers across many disciplines need city center location data to identify such places as central business districts (CBDs). The notion of a "city center"—as distinct from the "central city" of a metropolitan statistical area (MSA)—has different meanings depending on the context. One may envision the peak of a city skyline (often in financial centers); a historic transportation hub, such as Grand Central Station in New York City (Schelling, 1960); or a location defined on objective criteria, such as population or employment density (Taubenböck et al., 2013).[1] Other meanings of the term *city center* could be possible. The concept of a CBD, however, is widely used in research.

In this article, I present five measures used in urban economics and related research to identify city center locations. I describe strengths and weaknesses of the measures and present an analysis comparing the five databases, four of which cover U.S. MSA centers and a fifth that includes all census geographies. In keeping with the applied focus of the SpAM department of this journal, I also describe my methods so that others can adopt the techniques. I conclude the article with suggestions for finding appropriate measures of the socioeconomic center of a city.

Five Measures of Central City Points

A point—that is, a location—is represented on a map by geographic coordinates, which give the latitude and longitude of the "geocoded" point. At least five different measures have been used by researchers to identify city center locations. In this section, I describe those measures and outline their differences.

The first measure is from *Central Business Districts: 1982 Census of Retail Trade* (hereafter, 1982 Census), which was the last time the U.S. Census Bureau attempted to identify CBDs, and it did so for 455 cities[2] defined by existing census tract boundaries. Kneebone (2013: 3) notes that "Though dated, the 1982 CBDs represent the last systematic identification of business districts at the national scale. Furthermore, the 1982 CBDs continue to exhibit significant overlap with the densest job centers in the nation's major metro areas." Despite the strong conceptual appeal of the 1982 Census measure, one limitation for a researcher looking for geographic points is that those tract-based definitions are actually areas.

Researchers working at the Federal Reserve have converted those areas to points and made the data available on the Internet.[3] Although I refer to this measure as the 1982 Census measure, the conversion of areas to points was performed by Fee and Hartley (2013), who required CBD points for 385 MSAs. They describe their methodology on pages 47 to 48 of their study:

[1] A new and promising approach uses machine learning techniques along with both subjective assessment and objective criteria to determine CBD points (Brown et al., 2017).
[2] This is my manual count from the PDF file available at the Internet archive: https://web.archive.org/web/20070221191519/http://www.census.gov:80/geo/tiger/cbdct.pdf. This file lists census tracts identified by the 1982 study as a "central business district." See Brown et al. (2017: 6) for more details on the 1982 Census delineation of CBDs.
[3] These data can be accessed at http://www.danielaaronhartley.com/msas_with_central_city_cbds.csv.

> We identify the latitude and longitude of the CBD by taking the collection of census tracts listed in the *1982 Census of Retail Trade* for the central city of the MSA (the city in the MSA with the largest population) and finding the centroid of that cluster of census tracts. We identify the CBD latitude and longitude for 268 MSAs in this manner.

Note that because the 1982 Census measure is available for only 455 cities, Fee and Hartley (2013) were able to merge these data to only an N=268 subset of their sample of contemporary MSAs. Their solution to this problem, described by the authors on pages 47 to 48, takes us to a second measure of central city location:

> For the remaining 117 MSAs, whose central city was not listed in the *1982 Census of Retail Trade*, we use the latitude and longitude found by geocoding the MSA's central city found using the ArcGIS 10.0 North American Geocoding Service. ArcGIS returns points that are, on average, very close to the *CBDs from the Census of Retail Trade*; for the 268 cities for which we have both, the mean distance between the two is 0.39 miles.

One area of interest for Fee and Hartley (2013) was measuring the effect of so-called agglomeration economies, which refers to greater invention, innovation, and enhanced productivity because workers in the same industry live, work, or otherwise interact in close proximity. Fee and Hartley (2013) found an average distance of 0.39 miles between the 1982 Census and the ArcGIS measures.[4] The authors concluded that the measurement error was likely small for most industries because Elvery and Sveikauskas (2010) suggest in their work that distances of between 5 and 25 miles may be relevant for many industries. They discuss the possibility, however, that relevant distances for agglomeration effects may be shorter in industries with more highly skilled workers.[5]

A third measure, similar to the ArcGIS measure, was developed as part of the study by Holian and Kahn (2015),[6] who used Google Earth to geocode the principal city of each of 366 MSAs by recording the latitude and longitude returned—the location where the map view centered—from a city name search. One of the applications in which Holian and Kahn (2015) were interested was whether households living near the city center drive less on average than those who live farther away. Given that the average metropolitan U.S. household lived about 20 miles from their city's downtown area, measurement error—if on the scale of that reported by Fee and Hartley (2013)—would not seem to present a major threat to the internal validity of the analysis. Whether the Google Earth measure is close to the 1982 Census measure is unknown, but this topic is addressed it the next section.[7]

---

[4] ArcGIS is a prominent Geographic Information System (GIS) software package by Environmental Systems Research Institute (ESRI). ESRI provides GIS data files for use in ArcGIS. Fee and Hartley (2013) used these data to create CBD locations, but the authors do not describe how they determined the point location.

[5] This possibility was raised by Fee and Hartley (2013) in endnote 2: 262.

[6] These data are available to download on my blog: http://mattholian.blogspot.com/2013/05/central-business-district-geocodes.html.

[7] In their initial study, Holian and Kahn (2015) recognized the possibility of measurement error but noted that errors of up to a mile would likely not present a major challenge given the nature of the analysis, and that all points in a sample of points they visually inspected seemed to be reasonable approximations of CBD location.

Like the ArcGIS measure, the Google Earth measure is also something of a "black box" in the sense that the variables in the Google algorithm that determined the locations are unknown. Nevertheless, this measure has been used or discussed in articles, including Anenberg and Kung (2018) and Gardner and Hendrickson (2018); Ph.D. dissertations, including Molnar (2014), Resseger (2014), and Su (2018); Federal Reserve Bank publications (Rappaport, 2017, 2016, and 2014); and working papers (Couture et al., 2018; Couture and Handbury, 2017; Hamilton and Dourado, 2018).

A fourth measure was developed by Wilson and colleagues (2012), who provide geocodes based on the address of each city's city hall. These data are from a worksheet in the Chapter 1 spreadsheet file associated with their study.[8] A principal advantage of these measures is that they are conceptually clear, similar to the centers from the 1982 Census. Although the concept of a political center is distinct from that of a business center, one can reasonably expect these political centers to be near the CBDs. This city hall measure provides geocodes for the principal city in each of 368 MSAs.[9] Examples of studies that have used or discussed location information from this source include Dascher (2018); Hall, Palsson, and Price (2018); and Liu (2013).

Finally, a fifth measure of city center used in research is the U.S. Census Gazetteer.[10] These center points are based on the interpolated latitude and longitude associated with the Census Bureau's TIGER/Line Shapefiles. The Gazetteer files provide population, land area, and "representative latitude and longitude coordinates" for *all* census geographies, including tracts, counties, and cities.[11] Documentation explaining how these points were selected or created is unavailable.[12] This measure is similar to Google Earth and ArcGIS with regard to being a black box in creating the central point; however, the Gazetteer is by far the most comprehensive measure I discuss here, which is its principal virtue and can be readily accessed for all census geographies. Holian and Kahn (2015); James and Aadland (2011); and Nelson, Uwasu, and Polasky (2007) provide examples of research use location information from this source.

Analysis

In the preceding section, I described conceptual strengths and weaknesses but posted several questions about their accuracy in representing a city center. This section provides a quantitative assessment of the measures and offers some answers as to their accuracy.

To calculate the distance between two points, I use a formula called the spherical law of cosines. This formula takes into account the spherical shape of the Earth and is accurate over distances

[8] See https://www.census.gov/library/publications/2012/dec/c2010sr-01.html. The points can be found by clicking the link to the Excel file for Chapter 1.
[9] The three measures, Wilson et al. (2012), Fee and Hartley (2013), and Holian and Kahn (2015), all provide geocodes for the principal city of all MSAs in the United States. The reason for the seeming discrepancy in sample sizes is that the MSA definitions change from year to year, and researchers naturally choose different year definitions. An additional complication of using MSAs as units of analysis is that the Census Bureau changes the definition of an MSA's principal city in some years: the principal city is the most populous city in the MSA, so that designation will change when city populations grow or shrink in a way that changes the ranking of cities in the same MSA.
[10] These data can be accessed at https://www.census.gov/geo/maps-data/data/gazetteer2010.html.
[11] This quote is from "2017 U.S. Gazetteer Files" on the Census Bureau website: https://www.census.gov/geo/maps-data/data/gazetteer2017.html.
[12] These centers may be centroids (the geographic center of the city's area, which may or may not be close to the economic center).

typical in intra-urban research. This formula yields what can be thought of as the distance between two points "as the crow flies." I have created a spreadsheet file that implements this formula.[13]

I first calculated the average distance between each of the five measures, for a total of 10 measurements, shown in exhibit 1.

Exhibit 1

Matrix of Average (Mean) Distances (in km)

	Arc GIS	Google Earth	City Hall	Gazetteer
1982 Census of Retail Trade	0.64	0.97	0.52	3.68
ArcGIS		0.98	0.48	3.36
Google Earth			0.94	3.50
City Hall				3.44

The results in exhibit 1 show the City Hall measure to be closest to the 1982 Census measure, at 0.52 kilometers (km), and the Gazetteer measure is farthest from it, at 3.68 km. The mean distance between the ArcGIS and 1982 Census measures used by Fee and Hartley (2013) is 0.64 kilometers, which is extremely close to the mean distance of 0.63 km (or 0.39 miles) that they reported. Of note, I find nearly the same result as Fee and Hartley, even though my sample is slightly different, with N=256 versus N=268. Sample sizes for each mean distance measure in exhibit 1 are shown in exhibit 2.

I now offer a more detailed analysis of the measures. Exhibit 2 shows summary statistics for each of the 10 distances. In addition to average distance, exhibit 2 reports the standard deviation of the distances. These data are useful to report because a measure of location that is farther from the CBD on average but that has a smaller standard deviation may be more desirable for some purposes than a measure that is closer on average but for which the possibility of large errors exists. In addition, exhibit 2 reports the number of cities (N) and the median, minimum, and maximum distances to provide fuller insight into the characteristics of the distances between each CBD measure.[14]

---

[13] This file and accompanying tutorial analysis can be downloaded from http://mattholian.blogspot.com/2017/03/using-spreadsheet-to-calculate-distance.html. For more information on the formula, see http://www.movable-type.co.uk/scripts/latlong.html. I have also written a blog article (Holian, 2017) that contains instructions for using that file, sample data, and a tutorial analysis. My "distance calculator" spreadsheet also contains a worksheet with additional instructions. Although I provide a lot of instructions, the spreadsheet is actually fairly easy to use. There are a few helpful practices that users would learn themselves after using it a few times, but I have tried to spell out all of them in the blog article so that others will not have to reinvent the wheel.

[14] I did not include an estimate of the standard error (SE) of the means in repeated sampling, but that value can be readily calculated based on the formula $SE=SD/\sqrt{N}$. The test statistic for a test for which the mean distance is zero is then the mean divided by the SE. All of the mean distances are statistically significant at zero at conventional levels.

Exhibit 2

Summary Statistics for 10 Distance Measures

	N	Mean	Median	SD	Min	Max
1982Cen to ArcGIS	256	0.64	0.42	0.78	0.01	6.85
1982Cen to Google Earth	256	0.97	0.49	1.53	0.02	12.62
1982Cen to City Hall	256	0.52	0.36	0.77	0.02	6.93
1982Cen to Gazetteer	256	3.68	2.51	4.65	0.06	55.26
ArcGIS to Google Earth	361	0.98	0.50	1.60	0.00	13.07
ArcGIS to City Hall	361	0.48	0.30	0.65	0.01	5.72
ArcGIS to Gazetteer	361	3.36	2.37	4.39	0.12	54.14
City Hall to Google Earth	365	0.94	0.45	1.65	0.01	13.40
City Hall to Gazetteer	365	3.44	2.40	4.46	0.06	54.27
Google Earth to Gazetteer	366	3.50	2.35	4.95	0.07	54.14

1982Cen = 1982 Census of Retail Trade. SD = standard deviation.

The ArcGIS measure is 0.64 km from the 1982 Census measure on average, whereas the City Hall measure is only 0.52 km away on average. That difference is a relatively small 0.12 km, but it suggests, for example, that because the City Hall measure is closer to the 1982 Census measure on average, studies such as Fee and Hartley (2013) and Holian and Kahn (2015) would have been better off using the City Hall measure from Wilson et al. (2012) instead of the ArcGIS or Google Earth measures. Is the difference large enough to be statistically significant? The value of the difference between means test statistic is 1.71 (p=0.087), meaning that the difference is significant at the 10-percent level but not at the more conventional 5-percent level.[15] The standard deviation and maximum distances are very close for both the 1982 Census-to-City Hall and the 1982 Census-to-ArcGIS distances. Finally, in addition to the Gazetteer measure having the highest average distances to all other central location measures, the maximum distance is much higher in the Gazetteer measure and in one case is more than 54 km away from the City Hall measure, which is in the San Francisco MSA and is discussed in the next section.

[15] See McDonald (2014) for a discussion about this test; the author of that article has also created a spreadsheet tool that implements this test, which can be downloaded from that page. This test is conceptually different from what is probably the more familiar difference-in-means hypothesis test.

Application

This section presents qualitative analysis for two MSAs as case studies, starting with New Orleans. The Google Earth measure is not the spot that most people would identify as the center of the city, but it is close to typical locations that people might cite. For example, the Google Earth measure locates the city center at slightly less than 1 mile north of Jackson Square, a central park in the city's historic French Quarter. That difference may or may not be a problem for some analyses, such as housing price change away from the city center, but that distance could be a large enough error to be a problem for other applications, such as in measuring the diffusion of ideas. For example, if workers on their lunch break are not able to walk much more than a mile from their office, measurement error on this scale could severely limit opportunities for informal exchange of ideas.

The New Orleans City Hall location is within the area that is known locally as the CBD, but it is toward the outskirts of that district, near the Superdome and Interstate 10. The ArcGIS measure is one large city block from the City Hall measure. Although these three may be close for many applications, the 1982 Census measure is the best of the five because it is firmly centered in the neighborhood identified by local residents as the CBD and is only a regular city block from the French Quarter. Meanwhile, the Gazetteer measure is the worst of the five; it locates the center in a wildlife refuge more than 10 miles away from the CBD and the French Quarter. Exhibit 3 shows geocodes for all five measures, exhibit 4 presents all five points, and exhibit 5 shows the four nearby points presented in greater detail.

Exhibit 3

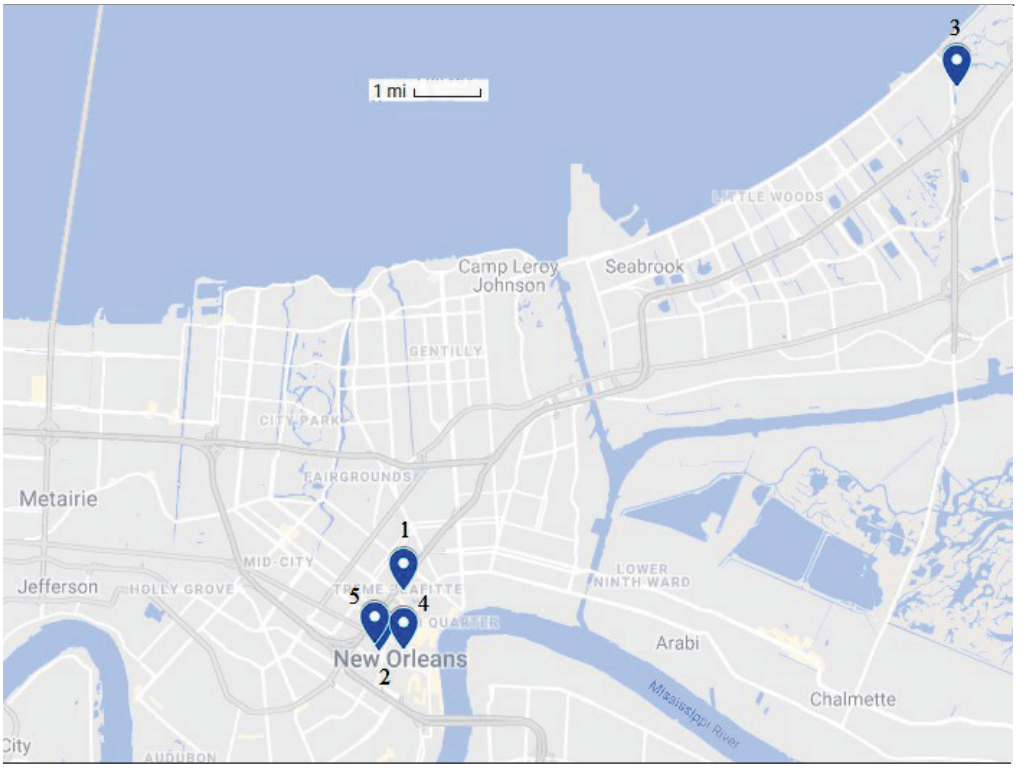
Five Measures of Central Points for Two Cities

Measure		New Orleans		San Francisco	
		Latitude	Longitude	Latitude	Longitude
1	Google Earth	29.964722	– 90.070556	37.774930	– 122.419416
2	City Hall	29.952403	– 90.076487	37.778503	– 122.418307
3	Gazetteer	30.068636	– 89.939007	37.727239	– 123.032229
4	Census 1982	29.952499	– 90.070801	37.785702	– 122.408000
5	ArcGIS	29.953701	– 90.077751	37.777122	– 122.419639

Sources: 1: http://mattholian.blogspot.com/2013/05/central-business-district-geocodes.html; 2: https://www.census.gov/library/publications/2012/dec/c2010sr-01.html; 3: https://www.census.gov/geo/maps-data/data/gazetteer2010.html; 4 and 5: http://www.danielaaronhartley.com/msas_with_central_city_cbds.csv

Exhibit 4

Five Measures of Central Points for New Orleans

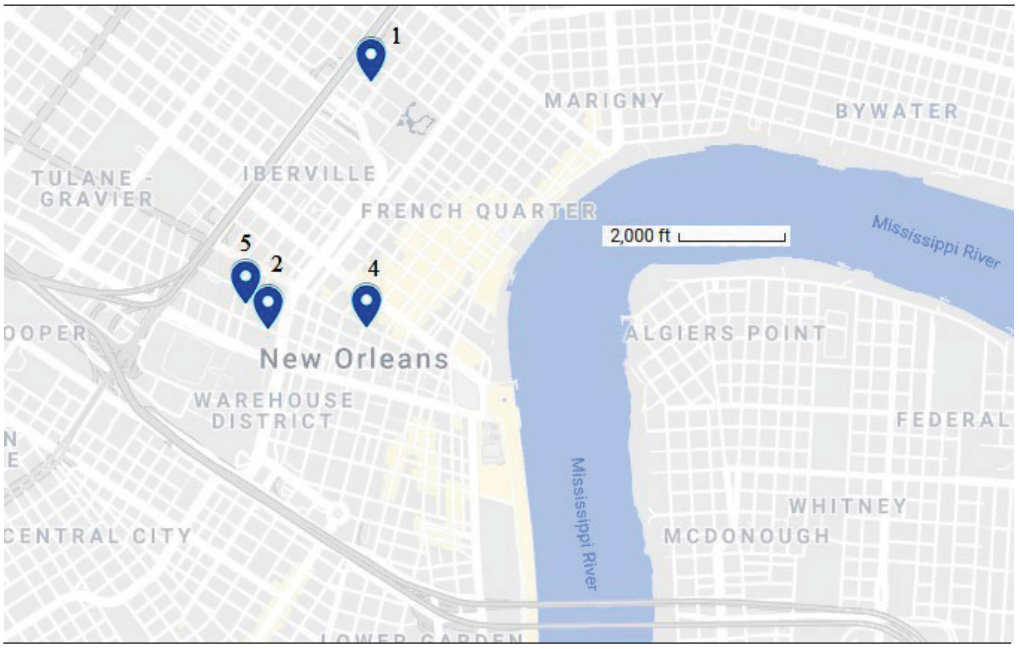


Notes: 1. Google Earth. 2. City Hall. 3. Gazetteer. 4. Census 1982. 5. ArcGIS. Plotted points are shown in Exhibit 3.

As another illustrative example with these measures, consider San Francisco, where the Gazetteer point is on the Farallon Islands, which are in the Pacific Ocean and are visible from the mainland only on the clearest of days. The City Hall, Google Earth, and ArcGIS measures, however, are within one to four blocks of each other. Despite the near consensus reached by those three measures, other contender definitions for the center of San Francisco—the historic transportation center (the Ferry building), financial district (FiDi), or shopping district (Union Square)—are several kilometers away from city hall. The City Hall and Google Earth measures are only a few kilometers away from each of those points, whereas the Gazetteer measure is more than 50 km away. Finally, and as in the case of New Orleans, the 1982 Census measure seems most appropriate for the majority of applications requiring a socioeconomic measure of centrality: it is 2 blocks from the Union Square shopping district, which is between the Financial District, the Ferry Building, and city hall.

Exhibit 5

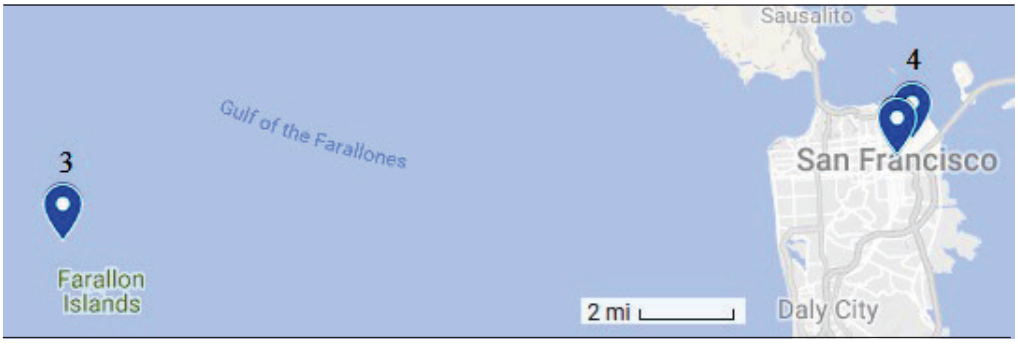
Four Measures of Central Location in New Orleans, Detail Area



Notes: 1. Google Earth. 2. City Hall. 4. Census 1982. 5. ArcGIS. Plotted points are shown in Exhibit 3.

Exhibit 6

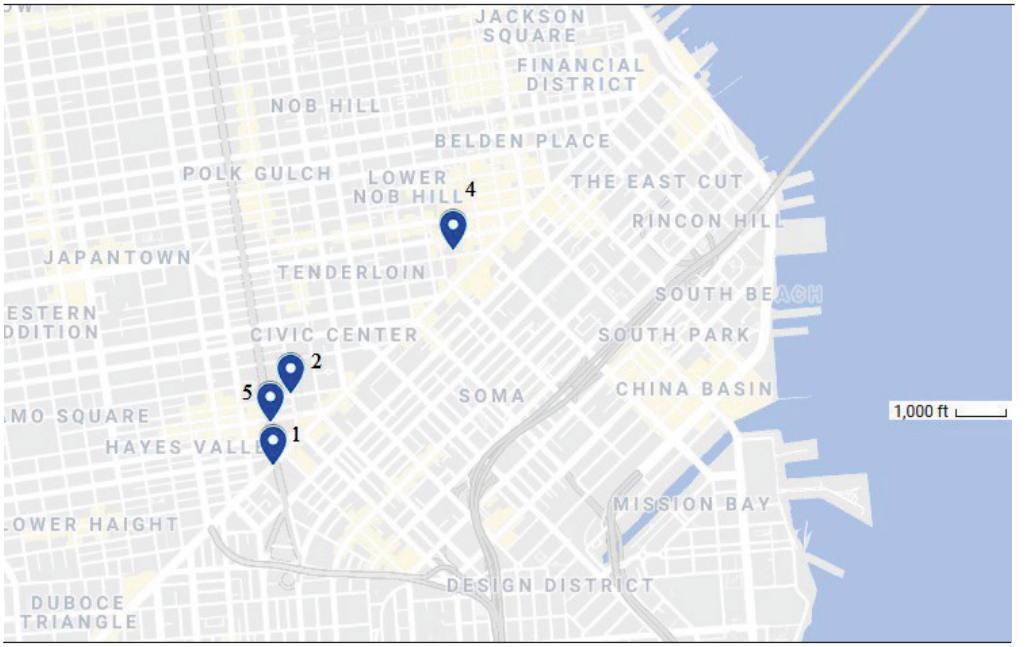
Two Measures of Central Location in San Francisco



Notes: 3. Gazetteer. 4. Census 1982. Plotted points are shown in Exhibit 3.

Exhibit 7

Four Measures of Central Location in San Francisco, Detail Area



Notes: 1. Google Earth. 2. City Hall. 4. Census 1982. 5. ArcGIS. Plotted points are shown in Exhibit 3.

Conclusion

Researchers who need access to a measure of city center locations have at least five databases at their disposal. Depending on the nature of the researcher's study, one may be more appropriate than the others; however, in some situations, the choice between them would likely not matter much, given the documented similarities across the measures. I conclude with the following suggestions for using these measures:

- The 1982 Census points are probably the best measures of the CBD concept; however, Fee and Hartley (2013) produced points for only 268 cities, and even the original survey delineated CBDs for only 455 cities. This source alone will not cover all principal cities in contemporary MSA samples.
- If the 1982 Census does not include the city being studied, use the City Hall database; these points come closest on average to the 1982 Census points in this analysis. These points are available for the principal cities of most of the currently defined MSAs.
- If a city is not available in either the 1982 Census or City Hall databases, the location of the city hall could be geocoded using Google Earth or ArcGIS.

- If time constraints prohibit geocoding City Hall points, the Google Earth or ArcGIS geocoded city points could be used; the ArcGIS points were closer to the 1982 Census points in this analysis, suggesting that they are superior to the Google Earth measures for CBD purposes.
- The Gazetteer should be used only in cases in which large measurement errors are not a problem, such as approximating airline travel distances between cities, or when they are used to measure the center of small census geographies, such as tracts in metropolitan areas, as in Holian and Kahn (2015). The Gazetteer measure is also the most readily accessible and comprehensive in geographic scope of the five measures.

Time permitting, researchers can always complete their tasks using multiple measures and then compare and contrast the results as a sensitivity analysis. I hope that continued research on city center location provides researchers with more and better data, and to that end I share the data and analysis file I used in writing this article.

Acknowledgments

The author thanks Rich Taketa and an anonymous reviewer, who offered helpful comments during the drafting of this manuscript. Any errors are the author's.

Author

Matthew J. Holian is a professor of economics at San Jose State University and can be reached by email at matthew.holian@sjsu.edu.

References

Anenberg, Elliot, and Edward Kung. 2018. "Can More Housing Supply Solve the Affordability Crisis? Evidence From a Neighborhood Choice Model," *Regional Science and Urban Economics*. https://doi.org/10.1016/j.regsciurbeco.2018.04.012.

Brown, Jason, Maeve Maloney, Jordan Rappaport, and Aaron Smalter Hall. 2017. *How Centralized is U.S. Metropolitan Employment?* Working paper RWP 17-16 (November). Prepared for Kansas City, MO: Federal Reserve Bank of Kansas City. https://dx.doi.org/10.18651/RWP2017-16.

Couture, Victor, Cecile Gaubert, Jessie Handbury, and Erik Hurst. 2018. "Income Growth and the Distributional Effects of Urban Spatial Sorting," Unpublished manuscript. https://economics.stanford.edu/sites/g/files/sbiybj9386/f/welfare-implications-urban-current.pdf.

Couture, Victor, and Jessie Handbury. 2017. *Urban Revival in America, 2000 to 2010*. NBER Working Paper No. 24084. Prepared for Cambridge, MA: National Bureau of Economic Research. https://www.nber.org/papers/w24084.

Dascher, Kristof. 2018. "Function Follows Form," *Journal of Housing Economics*. https://doi.org/10.1016/j.jhe.2018.08.003.

Elvery, Joel A., and Leo Sveikauskas. 2010. "How Far Do Agglomeration Effects Reach?" Unpublished paper. Cleveland State University. https://www.researchgate.net/publication/228428788_How_far_do_agglomeration_effects_reach

Fee, Kyle, and Daniel Hartley. 2013. "The Relationship Between City Center Density and Urban Growth or Decline." In *Revitalizing American Cities*, edited by Susan M. Wachter and Kimberly A. Zeuli. Philadelphia: University of Pennsylvania Press: 45–64.

Gardner, John, and Joshua R. Hendrickson. 2018. "If I Leave Here Tomorrow: An Option View of Migration When Labor Market Quality Declines," *Southern Economic Journal* 84 (3): 786–814. https://doi.org/10.1002/soej.12249

Hall, Jonathan D., Craig Palsson, and Joseph Price. 2018. "Is Uber a Substitute or Complement for Public Transit?" *Journal of Urban Economics* 108: 36–50. https://doi.org/10.1016/j.jue.2018.09.003

Hamilton, Emily, and Eli Dourado. 2018. The Premium for Walkable Development Under Land Use Regulations. Mercatus research paper C2010SR-01. Prepared for Arlington, VA: Mercatus Center at George Mason University. https://www.mercatus.org/publications/walkable-development-land-use-regulations.

Holian, Matthew J. 2017. "Using a spreadsheet to calculate the distance between two points." *Matt Holian, Economist blog*, March 4. http://mattholian.blogspot.com/2017/03/using-spreadsheet-to-calculate-distance.html.

Holian, Matthew J., and Matthew E. Kahn. 2015. "Household Carbon Emissions From Driving and Center City Quality of Life," *Ecological Economics* 116: 362–368. https://doi.org/10.1016/j.ecolecon.2015.05.012

James, Alex, and David Aadland. 2011. "The Curse of Natural Resources: An Empirical Investigation of U.S. Counties," *Resource and Energy Economics* 33 (2): 440–453. https://doi.org/10.1016/j.reseneeco.2010.05.006

Kneebone, Elizabeth. 2013. *Job Sprawl Stalls: The Great Recession and Metropolitan Employment Location*. Prepared for Washington, DC: Brookings Institution. https://www.brookings.edu/wp-content/uploads/2016/06/Srvy_JobSprawl.pdf.

Liu, Haofei. 2013. *What Affects the Number of Non-Work Stops Made During Commute Tours? A Study Based on the 2009 U.S. National Household Travel Survey in Large Metropolitan Areas*. Ph.D. diss., University of California, Los Angeles.

McDonald, John. 2014. Handbook of Biological Statistics, 3rd ed. Baltimore, MD: Sparky House Publishing: 126–130. http://www.biostathandbook.com/twosamplettest.html.

Molnar, Gabor. 2014. *Three Empirical Essays on United States Broadband Internet Markets: Quality, Competition, and Welfare*. Ph.D. diss., University of Colorado, Boulder.

Nelson, Erik, Michinori Uwasu, and Stephen Polasky. 2007. "Voting on Open Space: What Explains the Appearance and Support of Municipal-Level Open Space Conservation Referenda in the United States?" *Ecological Economics* 62 (3–4): 580–593. https://doi.org/10.1016/j.ecolecon.2006.07.027

Rappaport, Jordan. 2017. "Crowdedness, Centralized Employment, and Multifamily Home Construction," *Economic Review—Federal Reserve Bank of Kansas City* Q (I): 41–83. https://www.kansascityfed.org/~/media/files/publicat/econrev/econrevarchive/2017/1q17rappaport.pdf.

———. 2016. *Productivity, Congested Commuting, and Metro Size*. Research working paper RWP 16-03 (January). Prepared for Kansas City, MO: Federal Reserve Bank of Kansas City. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2727441.

———. 2014. Monocentric City Redux. Research working paper RWP 14-09 (November). Prepared for Kansas City, MO: Federal Reserve Bank of Kansas City. https://www.kansascityfed.org/~/media/files/publicat/reswkpap/pdf/rwp14-09.pdf.

Resseger, Matthew George. 2014. *Essays in Urban Economics*. Ph.D. diss., Harvard University.

Schelling, Thomas C. 1960. *The Strategy of Conflict*. Cambridge: Harvard University Press.

Su, Yichen. 2018. *The Rising Value of Time and the Origin of Urban Gentrification*. Ph.D. diss., Stanford University.

Taubenböck, Hannes, Martin Klotz, Michael Wurm, J. Schmieder, B. Wagner, M. Wooster, Thomas Esch, and Stefan Dech. 2013. "Delineation of Central Business Districts in Mega City Regions Using Remotely Sensed Data," *Remote Sensing of Environment* 136: 386–401. https://doi.org/10.1016/j.rse.2013.05.019

Wilson, Steven G., David A. Plane, Paul J. Mackun, Thomas R. Fischetti, and Justyna Goworowska (with Darryl T. Cohen, Marc J. Perry, and Geoffrey W. Hatchard). 2012. *Patterns of Metropolitan and Micropolitan Population Change: 2000 to 2010*. Prepared for the U.S. Department of Commerce, Economics and Statistics Administration, U.S. Census Bureau. Washington, DC: U.S. Census Bureau. https://www.census.gov/library/publications/2012/dec/c2010sr-01.html.

Racially Concentrated Areas of Affluence: A Preliminary Investigation

Edward G. Goetz
Anthony Damiano
Rashad A. Williams

Correction

The volume 21, number 1 issue of *Cityscape* contained errors in exhibit 5 and on page 109. The corrected discussion follows.

Income Concentration

Our analysis shows that 36 percent of metro areas had higher rates of concentrated affluence compared with concentrated poverty. Overall, we find 16 percent more areas of concentrated affluence compared to areas of concentrated poverty (2,297 tracts and 1,983 tracts respectively). However, the distribution of concentrated affluence is skewed toward fewer metro areas leading to higher average amounts of concentrated poverty compared to concentrated affluence in our sample. Places with the highest rates of concentrated affluence include wealthy metro areas in the West and Northeast such as San Jose, Washington, D.C., and San Francisco.

Exhibit 5

Average Metro Area Income

		Share Tracts w/ Concentrated Affluence		Share Tracts w/ Concentrated Poverty	
Region	**N**	mean	sd	mean	sd
Midwest	10	2.71	(1.41)	9.2	(4.14)
Northeast	7	4.95	(4.46)	5.13	(2.33)
South	22	4.27	(5.00)	5.33	(4.23)
West	11	7.00	(8.11)	2.27	(1.6)
Total	**50**	**4.66**	**(5.39)**	**5.4**	**(4.14)**

Referees 2018–19

The Office of Policy Development and Research gratefully acknowledges the contributions of the following referees and their assistance in making Cityscape *worth reading.*

Alex Abramovich
Laura Abrams
Arthur Acolin
Tim Aubry
Irina Barakova
Richard P. Barth
Tara Batista
William Clark
Anna Cororaton
Susanna Curry
Richard Dembo
Matthew Desmond
Anne Farrell
Phil Garboden
Maurice N. Gattis
Andrew Hanson
Judith Regina Havlicek
Saahoon Hong
B.K. Elizabeth Kim
Michelle Munson
Nathanael Okpych
Rachel Ong
Vincent Schiraldi
Jonathan Spader
Scott Susin
Halil Toros
Greg Townley
Mary Clair Turner
Kimberly Tyler
Jamie Van Leeuwen
Alex Wagaman
John Weicher
Lei Zhang

Symposium

Office of Policy Development and Research

Cityscape

VOLUME 21 NUMBER 2 • 2019