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, PD&R 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.*

The Importance of Using Layered Data To Analyze Housing: The Case of the Subsidized Housing Information Project

Vincent Reina
Michael Williams
Furman Center for Real Estate and Urban Policy,
New York University School of Law and Wagner School of Public Service

Abstract

The Furman Center for Real Estate and Urban Policy recently developed a new database through its Subsidized Housing Information Project (SHIP). The SHIP database combines more than 50 disparate data sets to catalogue every privately owned, publicly subsidized affordable rental property developed in New York City with financing and insurance from the U.S. Department of Housing and Urban Development (HUD), HUD project-based rental assistance, New York City or State Mitchell-Lama financing, or the federal Low-Income Housing Tax Credit Program. The pooling and layering of data, as well as combining the data with other local housing and neighborhood information, in databases like the SHIP allow for a clearer understanding of the existing affordable housing stock and enable practitioners to more effectively target resources toward the preservation of affordable housing.

Introduction

The Subsidized Housing Information Project (SHIP) is a comprehensive, publicly accessible database that New York University's Furman Center for Real Estate and Urban Policy (the Furman Center) developed. The SHIP database catalogues the nearly 235,000 units of privately owned, publicly subsidized affordable rental properties ever developed in New York City with financing and insurance from the U.S. Department of Housing and Urban Development (HUD), HUD project-based rental assistance, New York City or State Mitchell-Lama financing, or the federal Low-Income Housing Tax Credit (LIHTC) Program.[1] This database combines more than 50 government and public data sources to give an overview of the majority of privately owned, publicly subsidized housing in New York City. A database like SHIP can aid in efforts to preserve and manage these properties by providing a clear understanding of the number of units subsidized, an improved capability for assessing challenges to the existing subsidized housing stock, and a sharper view of the potential for properties to leave subsidy programs.

History

In 2007, concerned that the subsidized housing stock was rapidly declining in an overheating housing market, the John D. and Catherine T. MacArthur Foundation funded a Preservation Capacity Assessment for the city. The assessment resulted in a series of recommendations to the five city, state, and federal agencies charged with administering New York City's housing programs, including a suggestion that the agencies create an interagency working group to devise strategies to protect the affordability of subsidized properties. The assessment also highlighted the need for an independent and objective source of information about the subsidized housing stock (Begley et al., 2011). Accordingly, the Inter-agency Working Group (IWG) was formed in 2008 and selected the Furman Center to create a single database of all properties ever subsidized by HUD, the Mitchell-Lama programs, and the LIHTC Program. The Furman Center then applied to the MacArthur Foundation and received funding to develop the database.

The MacArthur Foundation, recognizing the need for local data on the privately owned and publicly funded multifamily rental stock, has supported similar efforts across the country. For example, the foundation funded several government agencies' efforts to develop databases for internal use, including the Los Angeles Housing Department. The foundation also funded other research centers, including DePaul University's Institute for Housing Studies and the University of Florida's Shimberg Center for Housing Studies, to create publicly accessible databases that catalogue the local affordable rental housing stocks.

The SHIP database is part of a larger national effort toward integrating disparate housing data into databases and making that information accessible to the public. HUD has also recognized the need for more comprehensive data and is looking into models for creating a national preservation database.

[1] For a more detailed explanation of these programs, see Furman Center, Institute for Affordable Housing Policy (2011).

Description of the Database

The fundamental value of the SHIP database is its ability to track multiple funding sources associated with a single property. The underlying data sets range from Oracle® databases designed for active asset management, to Microsoft® Excel® spreadsheets containing historical data to paper files of land disposition agreements. Unfortunately, these data sets—almost 50 in total—did not contain a unique identifier to track properties. Within each agency, and across agencies, there were often multiple distinct data sets that had different ways to identify properties. Some used only street addresses; others used Borough, Block, and Lot (BBL) numbers;[2] and others used agency-specific identification numbers.

To link data sets, we first standardized the spelling and punctuation of all of the property addresses and used those standardized addresses to match properties to BBL numbers. We then used the BBL numbers to merge properties across portfolios. Many of the properties that the SHIP database catalogued had multiple buildings or are scattered-site developments, so we often had to identify multiple BBL numbers and ensure they connected to the correct property.

After merging properties across portfolios, the next step was to standardize the data that each agency attached to its subsidies. For example, we found disparities in the way agencies measured indicators such as the number of units in a development, with some agencies counting superintendents' units in the total and others not. In some cases, we compared the data provided to the city's tax roll data to determine which was most reliable.

Advantages of Layering Data

The SHIP database enables users to view the layering of subsidies, which leads to more accurate counts of how many units have been developed and preserved and how many properties will expire from their subsidy program. When we look at all programs together, we see that 29 percent of the properties that received a subsidy from an agency at some point in their history also received a subsidy from another agency (Furman Center, 2011).[3] When we look only at the currently affordable properties, we find that more than 15 percent of these properties are subsidized through multiple programs. Exhibit 1 shows the level of subsidy overlap across the currently affordable properties.

If we ignored this overlap and counted a property each time it appeared in a portfolio, we would conclude that 2,485 properties were currently affordable, containing more than 230,000 units. By taking the overlap into account, however, we find there are currently only 2,129 affordable properties containing about 170,000 units. The fact that the number of units falls much more substantially than the number of properties when the merge is performed highlights that properties with multiple subsidies tend to be larger than properties with a single subsidy source. The average size of a property

[2] New York City uses BBL numbers to uniquely identify a taxable property.

[3] All data in this report come from The Furman Center's Subsidized Housing Information Project Data Search Tool (Furman Center, 2011).

Exhibit 1

Overlap of Financing Sources Across Currently Affordable Properties

LIHTC	Mitchell-Lama	HUD Financing and Insurance	HUD Project-Based Rental Assistance	Number of Properties	Number of Units
✔				1,469	70,373
	✔			36	11,408
		✔		21	4,884
			✔	274	34,912
Total, properties with one subsidy				1,800	121,577
✔		✔		2	82
✔			✔	21	2,898
	✔	✔		8	1,730
	✔		✔	19	10,607
		✔	✔	253	23,611
Total, properties with two subsidies				303	38,928
✔	✔		✔	1	559
✔		✔	✔	11	1,018
	✔	✔	✔	13	9,230
Total, properties with three subsidies				25	10,807
✔	✔	✔	✔	1	146
Total, properties with four subsidies				1	146
Total				**2,129**	**171,458**

HUD = U.S. Department of Housing and Urban Development. LIHTC = Low-Income Housing Tax Credit Program.

with one current subsidy is 68 units. The average size of a property with three current subsidies is 432 units. Counting these properties three times would significantly skew any survey of the affordable housing stock.

The layered data also reveal that some subsidies are more often combined with other forms of subsidy. For example, more than 50 percent of the properties with HUD project-based rental assistance have an additional form of subsidy, whereas only 2 percent of LIHTC properties have more than one form of subsidy. This finding makes sense because many of the older HUD programs coupled project-based rental assistance with some form of mortgage insurance. The LIHTC program is often layered with some form of soft debt or tax-exempt bond financing that does not necessarily mandate an affordability restriction beyond that of the tax credit. If we explore this layering further, however, we also see that every multiple-subsidy property has either HUD project-based rental assistance or HUD financing and insurance that compels affordability. The merged data illustrate that state and local financing programs consistently leverage HUD resources.

Practitioners have generally been unable to know all of the financing layers on any given property. Without taking layering into account, we would believe that all of the 108 properties that no longer receive HUD project-based rental assistance have converted to market-rate rents. The SHIP database, however, shows us that 23 of these properties are still affordable through another subsidy program. Furthermore, more than one-half of the properties that have left a HUD financing or insurance program still receive funding from another subsidy program that restricts rents.

A failure to take subsidy layering into account would also lead us to overstate opt-out risks. Looking at the merged data sets, we find that 62,000 housing units are in developments that no longer receive a subsidy that the SHIP database catalogues. Had we instead simply counted the number of units that had exited one of our subsidy programs, we would have estimated that 108,402 units were in properties that no longer receive a subsidy. In addition, a property often leaves one subsidy program because it refinances through another program that extends its affordability requirements. For example, since 2000, 106 properties containing 24,173 units had an expiring subsidy but were preserved thorough another program. These units amount to almost 15 percent of the units in properties that are currently affordable and would have shown up as opted out in some agency's portfolio. Layered data sets like the SHIP database enable us to know that these properties are still affordable and that the subsidy was used to preserve an existing property as affordable housing rather then develop new housing.

Databases like the SHIP also provide a more accurate count of properties with expiring subsidies. As we see in exhibit 2, if we counted the number of units that will reach the expiration date for each subsidy source, we would believe that more than 480 properties with 81,242 units will no longer have affordability restrictions in the next 5 years. If we layer all of these data sets, however, we see that most of these properties have multiple forms of subsidy and only 226 properties with 38,608 units are actually eligible to leave all affordability restrictions in the next 5 years.

The SHIP database then enables us to look at the list of expiring properties and better assess the effect of subsidy expiration. For example, when we look at properties with project-based Section 8, we see that 364 properties with 45,870 units are set to expire in the next 5 years. These contracts are renewable, however, and in recent years HUD has offered owners only short-term contract renewals. When we layer on the other forms of subsidies, we see that more than 120 of the expiring properties have at least one other form of subsidy that compels affordability beyond the next 5 years. We might therefore assume that, because those properties are still required to be affordable through another program, it is likely that those owners will choose to renew their Section 8 contracts.

Some forms of project-based rental assistance, namely the Rental Assistance Program and the Rental Supplement Program, are nonrenewable. If we look at these properties, we find that 17 properties with nearly 8,000 units are expiring between 2011 and 2015. Only one of these properties will still have some form of affordability restriction that compels it to be affordable after the project-based rental assistance expires. The challenge with these expiring properties is that they have a subsidy that provides direct rental income to the property and ensures tenants do not pay more

Exhibit 2

Subsidies Expiring in the Next 5 Years

Program	Properties	Units
HUD project-based rental assistance	382	47,353
Mitchell-Lama	67	29,188
HUD financing and insurance	29	4,600
LIHTC	2	101
Total—without layering	480	81,242
Total—with layering	226	38,608

HUD = U.S. Department of Housing and Urban Development. LIHTC = Low-Income Housing Tax Credit Program.

than 30 percent of their income on rent. Because this subsidy is not renewable, the property's financial structure will undergo a serious change, which could result in increased rents for low-income tenants or deteriorating property conditions. Using databases like the SHIP helps practitioners know that these subsidies are expiring and that no other subsidies on these properties will extend affordability restrictions.

Advantages of Pooling Data

The SHIP pools data sets, which allows for the analysis of financing trends. As shown in exhibit 3, more affordable housing was developed in the 1970s than in any other decade in the past 50 years. The Mitchell-Lama program and HUD's project-based rental assistance programs supported much of that development. New developments stopped using the Mitchell-Lama program by the 1980s, when project-based rental assistance and HUD financing and insurance became the most common subsidies. Since 1990, the LIHTC program has financed almost all new affordable housing developments.

Exhibit 3

Number of Units Developed by Decade and Program Subsidy Category

	1960s	1970s	1980s	1990s	2000s
HUD financing and insurance	3,079	11,361	14,898	995	173
HUD financing and insurance/project-based rental assistance	467	7,494	22,797	5,326	2,915
Project-based rental assistance	891	4,889	11,803	2,164	1,390
Project-based rental assistance/ Mitchell-Lama	0	22,996	248	0	0
Mitchell-Lama	14,772	19,199	0	0	0
Mitchell-Lama/HUD financing and insurance	0	1,788	0	0	0
HUD financing and insurance/project-based rental assistance/Mitchell-Lama	0	9,029	0	0	0
LIHTC	0	0	2,928	29,697	38,383
LIHTC/HUD financing and insurance	0	0	327	1,358	0
LIHTC/project-based rental assistance	0	0	0	103	0
LIHTC/HUD financing and insurance/ project-based rental assistance	0	0	0	0	82

HUD = U.S. Department of Housing and Urban Development. LIHTC = Low-Income Housing Tax Credit Program.

Advantages of Leveraging Existing City Data

Databases like the SHIP provide a platform for combining housing data with other local data about these properties and the neighborhoods where properties are located. The SHIP database includes more than 360 neighborhood-based indicators. These indicators range from detailed information about the physical and financial condition of properties to changes in local market and neighborhood characteristics.

Assigning each property a BBL creates new opportunities to link housing data with city data about physical or financial distress. Government officials might want to compare how different portfolios have performed when developing preservation priorities. The SHIP database enables us to compare

the property characteristics—such as housing code violations and tax delinquencies—across portfolios. For example, 8.7 percent of properties with HUD insurance in New York City have been delinquent on their taxes by at least $1,000 per unit for more than 1 year, whereas 5.1 percent of properties receiving project-based rental assistance have been delinquent. Mitchell-Lama properties have only a 1.3-percent delinquency rate. By using the SHIP database, we can see that none of the 22 properties that currently receive both HUD insurance and Mitchell-Lama financing are delinquent on their taxes, whereas 2.9 percent of properties that receive both HUD project-based rental assistance and Mitchell-Lama financing are delinquent.

The combination of housing and market data can also assist researchers. For example, the Furman Center is working on a model to predict which properties are likely to leave the Mitchell-Lama housing program. One theory is that the difference between restricted rents and potential market rents can predict an owner's likelihood of opting out of the program. Between 2002 and 2007, 56 Mitchell-Lama properties opted out. Of those properties, 36 were located in community districts where multifamily properties appreciated in price at a rate higher than the city average.

Finally, this information could help local officials target specific neighborhoods for preservation efforts. For example, government officials might want to prioritize preserving properties located in areas with high-performing schools. In New York City, the Upper East Side was one of the community districts with the highest share of students performing at or above grade level in 2009, so policymakers may want to focus resources on the 1,600 units in properties with expiring subsidies in that neighborhood over the next 5 years. The SHIP database also shows that the per-unit median price in 2010 for a multifamily rental building on the Upper East Side was almost twice the citywide median, which would suggest that these properties are relatively more expensive to preserve. Neighborhood-level quality of life and housing cost indicators such as these are critical as government officials, researchers, and advocates analyze properties and design preservation initiatives.

Conclusion

The SHIP database represents a proactive effort from government, advocates, researchers, and funders across the country to better understand the privately owned, publicly subsidized affordable multifamily rental housing stock. Integrated data sets like the SHIP database are important because they enable us to arrive at better estimates of how many properties have been developed, offer richer descriptions of property characteristics, enable more effective tracking of how many properties have left affordability programs, and help identify which properties will be eligible to leave their affordability programs in the near future. This information will help local, state, and federal government officials in their efforts to preserve affordable housing, enabling them to be more proactive. These data also provide a platform for researchers to better understand the intricacies of these programs and ultimately produce policy-relevant research. Integrated data sets like the SHIP database will become all the more crucial going forward, as older subsidies expire and new ones are used to finance the development and preservation of affordable housing.

Acknowledgments

The authors thank Vicki Been, Ingrid Gould Ellen, and Caitlyn Brazil for their insight on this article. They also thank Max Weselcouch for her assistance with compiling and analyzing the data; Stephen Roberts, our database project manager; and all the research assistants at the Furman Center who helped us along the way, including Brice Chaney, Matthew Murphy, Sajan Ravidran, and Rushin Shah. Finally, they thank the MacArthur Foundation for funding the development of the Subsidized Housing Information Project.

Authors

Vincent Reina is a research fellow at the Furman Center for Real Estate and Urban Policy, New York University School of Law and Wagner School of Public Service.

Michael Williams is a data manager and research analyst at the Furman Center for Real Estate and Urban Policy, New York University School of Law and Wagner School of Public Service.

References

Begley, Jaclene, Caitlyn Brazil, Vincent Reina, and Max Weselcouch. 2011. *State of New York City's Subsidized Housing: 2011*. New York: The Furman Center for Real Estate and Urban Policy.

The Furman Center for Real Estate and Urban Policy (Furman Center). 2011. "Subsidized Housing Information Project Data Search Tool." http://datasearch.furmancenter.org (accessed September 28).

Furman Center, Institute for Affordable Housing Policy. 2011. Directory of New York City Affordable Housing Programs (Beta). http://furmancenter.org/institute/directory/ (accessed September 28).

Additional Reading

The John D. and Catherine T. MacArthur Foundation. 2011. "United States Grantmaking: Affordable Housing." http://www.macfound.org/site/c.lkLXJ8MQKrH/b.943421/k.147B/Domestic_Grantmaking__Affordable_Housing__Recent_Grants.htm (accessed September 28).

U.S. Department of Housing and Urban Development. 2011. "Multifamily Housing Preservation." http://portal.hud.gov/hudportal/HUD?src=/program_offices/housing/mfh/presrv (accessed September 28).