

# THE FEASIBILITY OF DEVELOPING A NATIONAL PARCEL DATABASE:

COUNTY DATA RECORDS PROJECT FINAL REPORT



PD&R



Visit PD&R's website

[www.huduser.org](http://www.huduser.org)

to find this report and others sponsored by HUD's Office of Policy Development and Research (PD&R). Other services of HUD USER, PD&R's research information service, include listservs, special interest reports, bimonthly publications (best practices, significant studies from other sources), access to public use databases, and a hotline (800-245-2691) for help accessing the information you need.

# **THE FEASIBILITY OF DEVELOPING A NATIONAL PARCEL DATABASE:**

**COUNTY DATA RECORDS PROJECT FINAL REPORT**

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

Submitted by  
Abt Associates Inc.  
Bethesda, MD

In partnership with  
Fairview Industries  
Pendleton, SC

January 2013

## Acknowledgments

This report was prepared by Abt Associates Inc. (Abt) and Fairview Industries for the U.S. Department of Housing and Urban Development (HUD), Office of Policy Development and Research. The primary authors of this report are Nancy von Meyer, David Cowen, and David Stage of Fairview Industries, and Carlos Martín of Abt.

The authors acknowledge Will Huguenin and Daniel Boada, both from Abt, who provided the final data collection and standardization information from which the feasibility projections were estimated, along with staff from Smart Data Strategies who were primarily responsible for the data collection. Satyendra Patrabansh and Joshua Leopold were also instrumental

to this work's completion. Larry Buron provided advisory reviews, Suzanne Erfurth performed editing, and Melissa Connors provided production assistance.

The authors appreciate the guidance that Ashaki Robinson Johns and Jon Sperling of HUD provided in managing the entire effort. Finally, this study would not have been possible without the cooperation of those county and state public staff who provided data and answered questions about their ongoing parcel data attributes and collections. Their willingness to share information about the collection and quality of their data—so that others might learn from their experiences—was invaluable.

## Disclaimer

The contents of this report are the views of the contractor and do not necessarily reflect the views or policies of the U.S. Department of Housing and Urban Development or the U.S. government.

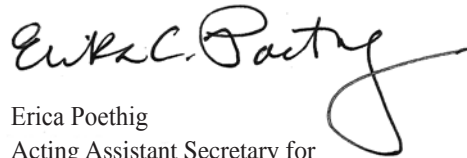
## Foreword

Neighborhood-level information is critical to HUD’s mission to provide quality, affordable homes located in strong, sustainable, inclusive communities. This feasibility study has charted new territory for HUD in an effort to develop a national database of standardized parcel-level (property) data collected directly from the authoritative sources: local government. The challenges are immense, as this study details, but the technical issues are minimal compared with the organizational, cultural, and political challenges in working with more than 4,000 local entities legally responsible for collecting and maintaining parcel data.

Although census data is an excellent resource for understanding and detecting changes in socioeconomic, demographic, and housing characteristics at multiple geographic levels across the country, parcel-level data, particularly if it is enhanced with geographic attributes, can provide a more detailed and nuanced landscape of the changes that occur at the neighborhood level. For example, parcel-level data has proven to be a critical resource in post-disaster recovery efforts and may have application as a source for early warning indicators of financially distressed housing and mortgage markets. However, to capture

the greatest social utility from the more standardized use of this data, it will be critical to inject a stronger national policy framework that promotes and balances the need for more open records with improved data quality and minimal risks of misuse.

This feasibility study offers alternative pathways going forward for achieving the vision of a national parcel database. As documented in this report, HUD has an opportunity to provide the leadership to overcome the institutional challenges of developing national parcel information. HUD looks forward to continuing progress and collaboration in this arena and will remain committed to supporting local, state, and tribal efforts to make this data and information more accessible, standardized, and useful.



Erica Poethig  
Acting Assistant Secretary for  
Policy Development and Research

# Table of Contents

<b>Executive Summary</b> .....	vii	<b>Appendix A. History of Federal Interest in Local Parcel Data</b> .....	36
<b>1. Introduction</b> .....	1	<b>Appendix B. HUD Authority, Data Needs, and Activity</b> ...	44
1.1. Project Scope .....	1	<b>Appendix C. Key Terms and Concepts</b> .....	47
1.2. Report Structure .....	2	<b>Appendix D. Data Collection Status of Study Counties</b> ....	49
<b>2. Data Collection and Standardization</b> .....	3	<b>Appendix E. Original Parcel Attributes</b> .....	52
2.1. Selected Counties and Data Collection Design .....	3	<b>Appendix F. Final Parcel Attributes</b> .....	54
2.2. Data Collection Process.....	5	<b>Appendix G. Parcel Data Comprehensiveness Findings</b> ...	56
2.3. Data Collection Response Rate .....	7	<b>Appendix H. Original Feasibility Questions</b> .....	63
2.4. Feasibility Survey Response Rate .....	12	<b>Appendix I. HUD-Approved Data Sharing Agreement Template</b> .....	65
2.5. Data Standardization .....	13	<b>Appendix J. Feasibility Findings</b> .....	69
2.6. Lessons Learned .....	15	<b>Appendix K. Preliminary Standardized Database Shell and Format</b> .....	73
2.7. Conclusion.....	17	<b>References</b> .....	74
<b>3. Feasibility of Ongoing Data Collection for Select Counties</b> .....	18		
3.1. Feasibility Measures and Cost Estimating.....	18		
3.2. Summary of Resource Requirements .....	21		
<b>4. Feasibility of Nationwide Data Collection</b> .....	23		
4.1. Nationwide Parcel Data Collection Constraints .....	23		
4.2. Nationwide Parcel Data Collection Requirements ..	25		
4.3. Nationwide Parcel Data Collection Alternatives .....	26		
4.4. Conclusion.....	34		



## List of Exhibits

Exhibit 1.	Location of HUD Study Counties .....	3	Exhibit 21.	Summary of 2009 Review of National Parcel Data .....	25
Exhibit 2.	Population Distributions of Study Counties and All U.S. Counties.....	4	Exhibit 22.	National Parcel Data System Option.....	26
Exhibit 3.	Populations of Study Counties and All U.S. Counties.....	4	Exhibit 23.	U.S. County Classifications by Resource Categories.....	27
Exhibit 4.	Number of Calls Made Until Primary Contact ..	5	Exhibit 24.	Multiyear National Data Collection Level of Effort Estimates (Hours) .....	28
Exhibit 5.	Average Calls Made Until Primary Contact, by State.....	5	Exhibit 25.	National Parcel Data With State Aggregation Option.....	32
Exhibit 6.	Data Access: Fees, Sharing Agreements, and Use Restrictions .....	6	Exhibit 26.	Estimates of Costs for Grants To Support State-Level Aggregation Efforts.....	33
Exhibit 7.	Data Collection Response Rate and Requirements.....	7	Exhibit A-1.	FGDC Organization .....	37
Exhibit 8.	Percent of Study Counties Providing Specific Datasets .....	8	Exhibit A-2.	Proposed New Coordination for National Parcel Dataset.....	39
Exhibit 9.	Percent of Study Counties by Total Dataset Availability .....	8	Exhibit B-1.	Proposed FGDC Reorganization.....	45
Exhibit 10.	Expected Availability of Original Data Attributes .....	8	Exhibit D-1.	Status of Collection for Selected Study Counties.....	49
Exhibit 11.	Attributes Collected, Percent of Counties.....	9	Exhibit E-1.	Parcel Attributes, Descriptions, and Comments .....	52
Exhibit 12.	Actual Attribute Availability .....	9	Exhibit F-1.	Final Attributes in Primary Attribute File .....	54
Exhibit 13.	Development Schedule for a Typical Assessment Roll .....	12	Exhibit F-2.	Final Attributes in Sales File.....	55
Exhibit 14.	Assessor Activities by Month (Distribution of Study Counties).....	13	Exhibit F-3.	Attributes Dropped From Study.....	55
Exhibit 15.	County Resource Categories .....	19	Exhibit G-1.	Dataset Types Collected (HUD-Selected Counties Only) .....	56
Exhibit 16.	Study County Classifications by County Resource Categories.....	19	Exhibit G-2.	Parcel Attributes Collected (HUD-Selected Counties Only) .....	59
Exhibit 17.	Definitions of Collection and Standardization Tasks.....	20	Exhibit H-1.	Instrument and Data Source for Feasibility Measures .....	63
Exhibit 18.	Four-Year Level of Effort per Task (Hours).....	20	Exhibit J-1.	Method for Accessing Parcel Data.....	69
Exhibit 19.	Summary Level of Effort per County Category (Hours).....	21	Exhibit J-2.	Prepackaged Parcel Dataset .....	69
Exhibit 20.	Multiyear Data Collection Level of Effort Estimates (Hours).....	21	Exhibit J-3.	Parcel Data Fees, Restrictions, and Waivers ....	69
			Exhibit J-4.	Currency of Data .....	70
			Exhibit J-5.	Tax Roll Profile .....	70

Exhibit J-6. Prepackaged GIS Dataset .....	70	Exhibit J-12. Level of Effort for Initial Contact (Number of Attempts) .....	71
Exhibit J-7. Method for Accessing GIS Data .....	70	Exhibit J-13. Future Plans for New Software .....	72
Exhibit J-8. GIS Fees, Restrictions, and Waivers .....	71	Exhibit J-14. Impending Retirements or Workforce Changes .....	72
Exhibit J-9. Level of Effort for Collection (Hours) .....	71	Exhibit J-15. Month of Tax Bill Mailing .....	72
Exhibit J-10. Future Automation Possibility.....	71	Exhibit K-1. Preliminary Standardized Database Shell and Format .....	73
Exhibit J-11. Data Sharing Agreement Type .....	71		



## Executive Summary

In 2010, the U.S. Department of Housing and Urban Development (HUD) charted new territory in an effort to develop a national database of standardized parcel-level (property) data collected directly from the most authoritative sources: local counties. HUD contracted with Abt Associates Inc. and their subcontractors, Fairview Industries and Smart Data Strategies, to embark on an exploratory project for assembling local assessor data, including key attributes such as property address, assessed value, land use, sales price, and sales history, for 127 targeted counties. The primary tasks of the project included identifying the appropriate data sources in each community, assembling the data and metadata, and standardizing the data in a common format to be accessible for HUD research efforts and for possibly aggregating data to higher levels of geography for public dissemination.

The research team identified authoritative sources and collected data and relevant documentation from 109 of the originally selected 127 counties—an 86-percent response rate—at varying levels of comprehensiveness. Achieving this response rate required extensive and unexpected additional outreach because of both the timing of the data collection phase (during the peak work season in many county assessor’s offices) and extensive negotiations for government-to-government (G2G) data-sharing agreements between the counties and HUD. Data from the remaining 18 counties were not collected for reasons that included not having electronically available data (3 counties); having parcel data in the hands of municipalities within the counties rather than those of the counties themselves (2 counties); fees (3 counties); data-sharing requirements (1 county); and other expressions of reluctance, including lacking the internal resources to process the request (9 counties).

Data standardization proved even more challenging. Key obstacles included—

- Lack of full data documentation from many of the study counties.
- Variations in each dataset’s comprehensiveness, attribute definitions and formats, and accuracy.
- Unclear and very diverse methods for internally validating data in each county.
- Wide variations in nomenclature and definitions for attributes (from land use codes to even basic assessment values).
- Incorrectly identified or duplicate values for similar attributes within datasets.

For these reasons, the study’s standardization task was limited to developing internally consistent datasets (that is, within each county) that use a uniformly defined attribute list, as opposed to a truly standardized dataset in which the parcel data are defined, collected, validated, and reported for each county according to identical protocols.

Despite the challenges, the study estimated the levels of effort for an ongoing collection and standardization for the study counties and for a nationwide collection and standardization effort. To make this estimation, the study classified the counties into four categories, based on their ease of collection and clarity of data documentation, and averaged the levels of effort in each category. These averages were defined as the baseline (“Year 1”) collection effort. Subject-matter experts then estimated the diminishing level of effort required on an annual basis until the minimal level of effort would be needed for ongoing activities (reached in “Year 4”).

For example, counties for which data can easily be accessed at the state level (Category 1) require a total level of effort of only 4 person hours in the first year, and counties for which data can be obtained through public websites (Category 2) require only 12 person hours in the first year. This effort is reduced to less than 1 hour for Category 1 and slightly more than 2 hours for Category 2 by the fourth year. By contrast, Category 3 counties require agreements or fee negotiations, and they take an initial 18.0 person-hour effort that diminishes to 4.5 person hours by the fourth year. Category 4, comprising counties with the least accessible data, require a consistent 25-person-hour level of effort for every year because of the expected repeated difficulty in collecting data. For an ongoing data collection and standardization of the original study counties, the first year’s total level of effort would be 1,585 person hours, assuming a similar response rate of 86 percent. Collection and maintenance of the data would be reduced to 636 person hours by the fourth year. ***The cost of the first 4-year effort for the study counties is estimated at more than \$800,000.***

The 3,221 counties and similar jurisdictions in the United States (including the District of Columbia and Puerto Rico) were also classified according to the four previously described categories based on the perceived representativeness of the counties in the sample and professional knowledge about the total population of U.S. counties. For an ongoing data collection and standardization of all U.S. counties, the first year’s total level of effort would be 45,653 person hours. This

figure would wind down annually to 13,694 person hours by the fourth year. *The cost of the first 4-year effort for all U.S. counties is estimated at nearly \$22 million.*

Both the study counties and the comprehensive national county collection efforts, however, would yield the same limitations of data quality as noted in this report. Reducing cost and improving data uniformity are imperative. Because of the magnitude of the resources and the persistent standardization problems associated with scaling up the methods used in the study, alternative approaches beyond the process of contacting and collecting data from individual counties performed in this study are explored. The study considered mid-term strategies for acquiring the data that are lower cost and long-term strategies for improving the uniformity of data that would ultimately lower collection and standardization costs significantly. These alternatives include—

- A need-based collection in which this project’s methods are applied to select counties only.
- The purchase of data from private vendors, which is currently practiced by most agencies.
- The support and encouragement of statewide efforts to aggregate county-produced data.

The last alternative, in particular, may prove to be the most reasonable—although longer term—approach to meeting HUD’s goals of efficiently collecting standardized, valid parcel data that could be available in the public domain. Simultaneously incentivizing states to support standards, consistent data collection, and state-level aggregation with their counties—combined with need-based data collections in the interim—appears to be the most effective approach. Fortunately, more states are developing statewide data standards or actually collecting the data and providing them in aggregate form; two states are currently using this protocol but seven additional states are in various stages of development. Accelerating this trend would reduce the cost requirements at the national level.

Despite the data collection and standardization challenges noted in this report, the results of this exploratory project suggest that an ongoing national data collection and maintenance effort for a subset of well-defined attributes is technically feasible with additional resources that diminish over the span of 4 years. The standardization effort would be more difficult, however, and would require significant resources. The most effective method to approaching both will depend on the amount of federal resources available and the changing landscape of parcel data maintenance in the United States.

# 1. Introduction

The U.S. Department of Housing and Urban Development (HUD) and other federal and local policymakers rely on detailed, reliable, and updated data on land ownership and property transactions to make informed decisions and address emerging issues. Real-time information about a community's housing stock can help gauge the level of neighborhood distress, identify the underlying causes of the distress, and inform the development of appropriate policy responses. Parcel-level information—that is, geographically referenced information about the ownership, rights, and interests of land parcels—can also be used to evaluate existing programs that are designed to stabilize communities suffering from foreclosures, lack of affordable housing, or natural disasters.

Numerous scholarly and policy studies have recommended investment in a national parcel-level data collection effort and the infrastructure to support the ongoing collection.<sup>1</sup> Because of HUD's interest in local parcel data for general housing and community development analysis and for targeted assistance efforts (such as disaster recovery), HUD has acquired or purchased parcel data in the past.<sup>2</sup> By developing and maintaining a national land parcel database then, HUD could accomplish many goals—for example, readily describing neighborhood changes due to HUD funding, or accelerating the federal response to natural disasters.<sup>3</sup> The need for a centralized parcel data infrastructure has been expressed, but no clear path (or resource estimate) for creating it has been identified.

Collecting parcel-level data from individual states and counties is more complicated than simply contacting each state or county and arranging for a data transfer. Core challenges involved in this effort to assemble parcel data include (1) data availability and completeness, (2) the willingness of local governments to provide data, and (3) the varying format and structure of data among counties.<sup>4</sup>

Some counties simply do not have data available in electronic format or may be in the process of constructing such a database. Moreover, for those counties that have at least some data

in electronic format, identifying which office holds the data, information about the data's collection, and the robustness of the desired attributes are further challenges. In most counties, the main source of parcel data is the county assessor's office. The assessment data, however, usually contain property sales data for only the last two or three transactions, and the source of the sales data may be in other offices, such as the deeds recorder's office or that of the county clerk that originally recorded the sales. Thus, in addition to not having older sales records, the assessment records may not provide the entire set of desired sale attributes. For example, the assessment records almost never contain mortgage information.

In addition, some assessment data are several years out of date or inconsistently updated. Some data elements may not be collected at all by some counties. For example, only some of the counties routinely collect information on easements as part of the assessment data. Similarly, the current condition of the property—particularly regarding its interior characteristics—might not be collected at all. To better gain a sense of these challenges, HUD decided to test the waters and build an initial standardized parcel dataset collected directly from the most authoritative local sources.

## 1.1. Project Scope

HUD contracted Abt Associates Inc. and its subcontractors, Fairview Industries and Smart Data Strategies, in August 2010 to assemble and standardize parcel data, including value, land use, sales, sales history, and address information, for 127 of the most populated counties in the United States. The project goals were divided into two categories.

### Data Collection and Standardization

- Identify the sources for the data.
- Collect the data in the most efficient manner possible, including paying reasonable data purchase fees, reviewing

<sup>1</sup> An extensive summary and review of these and other relevant publications are provided in appendix A.

<sup>2</sup> A review of HUD's current authority to collect parcel data, HUD offices with potential need for these data, and HUD's activities to date in parcel-level data collection and analysis are provided in appendix B.

<sup>3</sup> HUD coordinated an earlier exploratory effort with GIS Consultants and Boundary Solutions, Inc., to acquire parcel data and boundaries from several Gulf Coast counties after Hurricanes Katrina, Rita, and Wilma to support disaster assistance and related projects in the Gulf region.

<sup>4</sup> Key terms and concepts related to assessment data are provided in appendix C.

and signing data-sharing agreements or submitting more complicated ones to HUD, and requesting waivers of fees and agreements whenever possible.

- Determine the quality of data from these sources and standardize the data within each county in the most uniform manner possible for later geocoding and analysis by HUD.

### Feasibility Study

- Document the hours and costs for collecting and standardizing the data.
- Estimate the level of effort (in time and costs) for developing an ongoing database for this county sample and a nationally comprehensive database, based on the documentation.

First, the project focused on acquiring, assembling, and standardizing parcel data from 127 counties that received early Neighborhood Stabilization Program funding and the corresponding states with state-level parcel data repositories. HUD was specifically interested in parcel data related to tax assessment, property sales, easement, lien, land use, and condition. The first major task of the project was to conduct a pilot data collection for 9 counties to test the efficacy of the research design and to refine the approach and design before the main data collection phase. The refined strategy was then used to implement the main data collection phase of the project for the remaining 118 counties.

Second, the project focused on the feasibility of collecting and standardizing parcel data to build and maintain an ongoing parcel database that could be extended to have national coverage and be available to the public in aggregated form, based on the findings and lessons learned from the previously described pilot collection effort. Lessons learned during the data collection and data standardization and the team's previous expertise with parcel data were used to assess the feasibility of an ongoing data collection and standardization for these 127 counties and the nationwide effort.

## 1.2. Report Structure

This report summarizes the data collection and standardization processes, describes the original and standardized datasets, and presents the findings of the feasibility projections. Chapter 2 provides a review of the data collected, a description of the data collection and standardization effort, and lessons learned.<sup>5</sup> This chapter also addresses the data quality findings that will affect ongoing efforts. Chapter 3 analyzes the feasibility of ongoing collection and standardization for the study counties. Chapter 4 assesses the feasibility of extending this effort nationally and describes some key trends that are shaping the nation's parcel data infrastructure. The study concludes by presenting alternative strategies in the mid- and long-term for developing HUD's desired parcel database.

<sup>5</sup> The actual data and metadata are provided separately to HUD by Abt Associates Inc.

## 2. Data Collection and Standardization

This chapter reviews the methods and procedures used for the county parcel data collection and standardization, describes the data, and presents key lessons.<sup>6</sup> The effort included monitoring and gathering information on data content and tracking the resources and level of effort needed to predict future resource requirements. These indicators of resource requirements, called *feasibility measures*, were developed and refined during the pilot project phase as a means to measure the level of effort required for parcel data collection directly from the counties in the future.

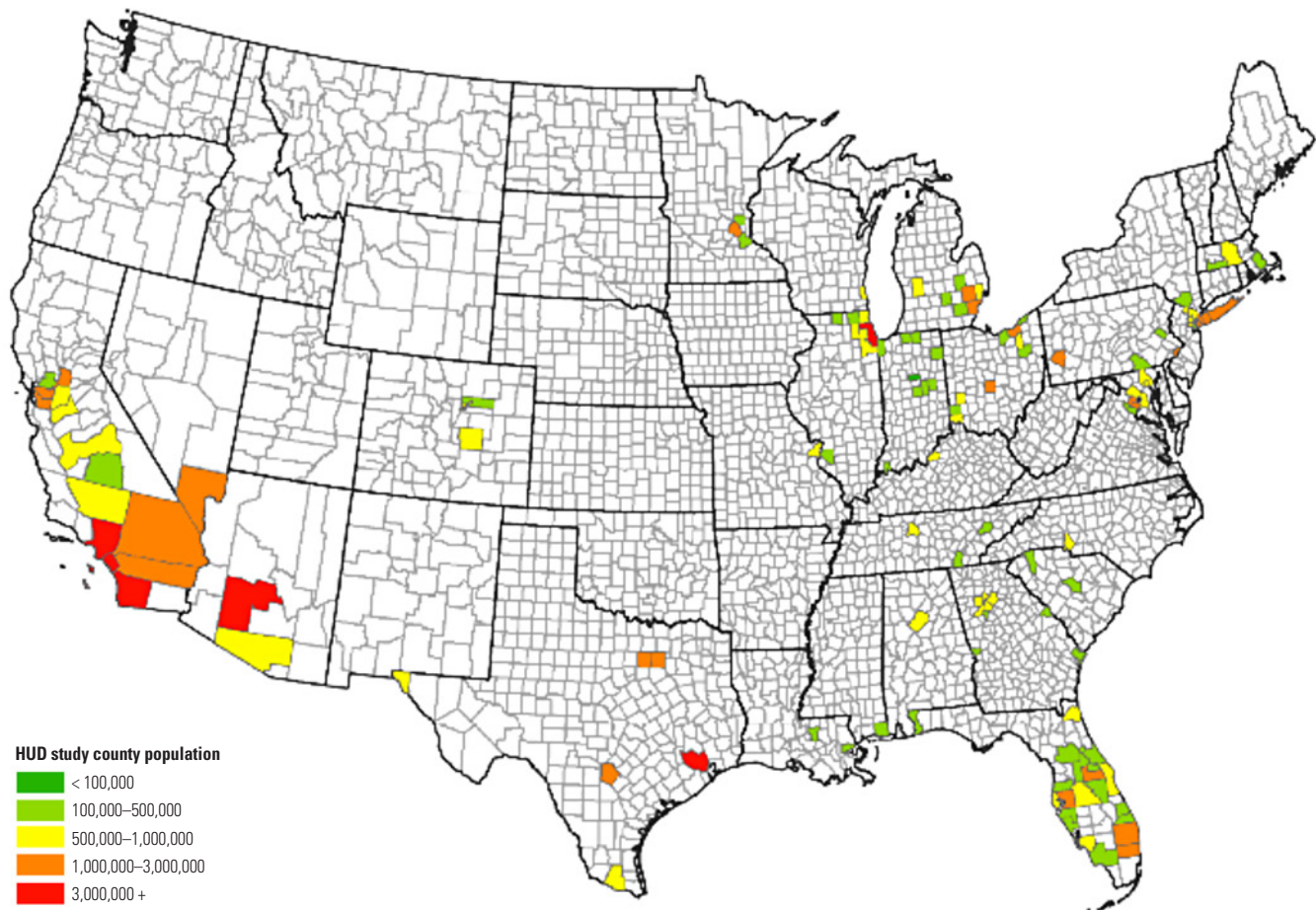
### 2.1. Selected Counties and Data Collection Design

#### 2.1.1. Selected Counties

The data collection included the 127 governmental units listed in appendix D. Their geographic distribution is shown in exhibit 1.

Exhibit 2 compares the population sizes of the 127 project counties with those of all 3,221 U.S. counties, including the

**Exhibit 1. Location of HUD Study Counties**



<sup>6</sup> More accurately, this project included 125 counties and 2 municipalities (New York City and Baltimore). In this study, the term county is used as matter of convenience to be inclusive of all 127 governmental units. In practice, New York City includes 5 county equivalents and the city of Baltimore is 1 county equivalent; the total number of county equivalents in the study, then, is 131. Jurisdictions beyond the 127 in the study that would be classified as counties for the purpose of national projections include Alaska’s boroughs, Louisiana’s parishes, Puerto Rico’s municipios, and the District of Columbia.



**Exhibit 2. Population Distributions of Study Counties and All U.S. Counties**

Population Range	HUD Project Counties	Percent of Total	U.S. Counties	Percent of Total
< 30,000	0	0	1,747	54
30,001–100,000	2	2	891	28
100,001–300,000	19	15	371	12
300,001–500,000	31	24	84	3
500,001–1,000,000	44	35	89	3
1,000,001–3,000,000	25	20	33	1
> 3,000,001	6	5	6	< 1
Total	127	100	3,221	100

Source: 2010 Census

District of Columbia and Puerto Rico’s *municipios*. Exhibit 3 provides various statistical measures of the populations in the project sample and in the universe. The U.S. Department of Housing and Urban Development sample tended to include counties with large populations; the median population of sample counties is approximately 621,000 compared with the national median of about 26,000.

**Exhibit 3. Populations of Study Counties and All U.S. Counties**

Population Statistics	HUD Project Counties (N = 127)	U.S. Counties (N = 3,221)
Total	123,871,446	312,471,327
Average	975,366	97,011
Min	39,370	82
Max	9,818,605	9,818,605
Median	620,961	26,076
Standard deviation	1,298,617	309,299

Although not representative of U.S. counties in general, this selection of counties represented a large proportion of total U.S. parcels. The geographic diversity of these counties (including the variety of parcel infrastructure development stages) and their relatively larger population sizes (associated with more advanced internal parcel data infrastructures) are particularly relevant; these characteristics are used to classify the population of U.S. counties into categories that support the methodology for calculating level of effort in this study.

**2.1.2. Data Collection Design**

The study’s data collection effort included two phases, a pilot project and the full data collection. The pilot provided a preliminary evaluation to determine attributes that the team could reasonably expect to collect from local government assessor’s office. The originally desired set of attributes is listed in

appendix E and the final collected set of attributes is provided in appendix F. The distribution of the attributes actually collected from each study county is provided in appendix G. As a part of this effort, a set of questions for our county data contacts was developed to evaluate the reasonableness of acquiring parcel data directly from local governments. The resulting set of indicators, some of which were collected through interviews and others from the collection process, were used to measure feasibility, as analyzed in chapter 3. The original questions for gauging the feasibility measures are provided in appendix H.

The plan for the primary data collection effort included five core activities to be performed in each county.

- **Contacts.** Local government assessor contacts were acquired for the 127 counties. Introductory letters from HUD were sent to each county in August 2011 with a brief explanation of the purpose of the data request and introducing HUD’s agents (specifically Smart Data Strategies and Abt Associates Inc.). Calls were made to confirm the primary contact information and begin the process of soliciting data and conducting interviews to support the feasibility analysis. Websites were also reviewed to determine if requisite data were publicly available for download.
- **Data collection requirements.** When needed, discussions were initiated to sign memoranda of understanding (MOU) with the counties, principally on agreements regarding limitations on use of their data. A budget was allocated to pay for fees from the counties charging for data purchase.<sup>7</sup> When counties had fees, they were asked if the fees could be waived, given the federal request. HUD also developed a standard MOU to address data-sharing concerns (appendix I). In the event a county required its own MOU, the team provided summaries to HUD of the key concerns in relation to HUD’s terms for approval.

<sup>7</sup> Researchers were given permission to automatically pay counties charging less than \$100, but HUD permission was requested for all counties charging more. The budget was minimally used because most counties either waived or reduced their fees on request.

- Data collection.** Depending on their internal processes and resources, counties that agreed to share data delivered them through various electronic channels (download, internal or external file transfer protocol, compact disk) and provided any existing data documentation available. Subject-matter experts developed crosswalks for each county dataset and the HUD-selected attributes to expedite later standardization.
- Interviews.** As a part of the data collection process, feasibility measures were collected through interviews with appropriate personnel. Informal notes were taken on progress being made, the number of contacts needed, barriers to acquisition, and the date of acquisition. The interviews, informal notes, and acquisition and processing results were documented and used to populate a set of resource measures to evaluate the cost of acquisition and future costs for followup acquisition of the same counties. This information was later used to help predict the resources needed to acquire data from nonproject counties. A categorization schema (presented in chapter 3) was developed that describes four different types of counties to capture their readiness, willingness, and ability to provide parcel data to HUD. This categorization schema was used to profile various local governments to assess the resource requirements for collecting local government parcel data.
- Standardization.** The researchers developed a template for naming the HUD-selected attributes and standardizing each county’s data to those attributes in consultation with subject-matter experts from Fairview Industries and its consultants. The researchers then processed each county’s data into the standard format (noting variations and exceptions within the datasets). SAS, a statistical software product, was used to convert the data into a standard file format.

## 2.2. Data Collection Process

The following discussion reviews some feasibility indicators gathered during the data collection phase, including the processing resources (measured primarily in time) required for collecting data. Key indicators, particularly for contacts and collection requirements, are discussed.

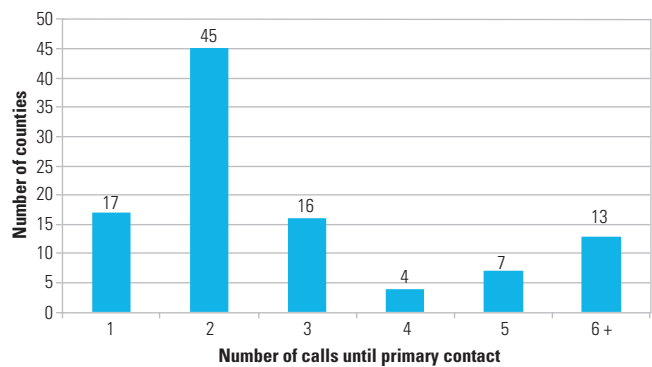
### 2.2.1. Contacts

The directors of assessor’s offices in all 127 counties were sent letters, and then called to acquire a primary contact for the acquisition of parcel data. Although a number of reasons might explain the difficulties in contacting some counties, the number of calls provides a measure of the readiness and willingness of a local government to provide data (exhibit 4). Data from 25 counties were publicly available, so no contact was available or needed. Most counties provided a contact within two or three

calls, but a substantial number required six or more calls. Note that for a multiyear collection effort, this activity is a one-time cost, as the correct contact will be known in future collections.

Exhibit 5, which shows the average time in hours to acquire a contact by state, provides additional perspective on the statewide culture for sharing parcel data, because formal data-sharing responsiveness is affected by state policy and the policies of individual counties. For example, Florida has a strong public access policy, and the state provides the certified roll on its website. Many local governments publish their working roll to their websites simply to reduce the workload associated with data requests. California and Colorado are in the process of developing statewide parcel layers for government agencies. By contrast, in Massachusetts, no county data were collected because the management of parcels takes place at the township

**Exhibit 4. Number of Calls Made Until Primary Contact**



Notes: N = 102 counties. Does not include counties for which data were publicly available for collection.

**Exhibit 5. Average Calls Made Until Primary Contact, by State**

State	Average Time per County (Hours)	Number of Counties
FL	0.63	15
NJ	0.65	5
AZ	0.75	2
NC	0.75	1
TX	0.75	6
OH	0.91	8
MN	1.00	3
CA	1.10	5
VA	1.13	2
MD	1.25	3
TN	1.25	5
WI	1.25	1
PA	1.38	2
GA	1.43	7
MI	1.44	4
CO	1.50	2
IN	1.63	4
NY	1.67	3
IL	1.81	8



level.<sup>8</sup> In general, the more willing and able the county is to provide data, the fewer resources it will take to acquire and process those data in the future. This issue was further explored in the feasibility interviews.

### 2.2.2. Data Collection Requirements

Data-sharing requirements may be shaped by both policy-related and technical issues, and they generally reflect the degree to which the data are already publicly available or the level of restrictions that are put on the data’s subsequent use. Data sharing is further complicated by the fact that different offices may collect and manage different portions of a county’s parcel information, and may impose different requirements within the same county.

Exhibit 6 summarizes the public access requirements that were placed on both the attributes and the parcel map during the primary data collection as part of the feasibility study. The numbers reflect only those counties for which interviews were performed (70 percent of all counties for which data were collected).

A critical concern regarding access to local government parcel data is cost. By contrast with expectation, the fees were not as formidable as budgeted (see appendix D). The total cost from fees for the 109 counties came to \$9,265, with about 91 percent of the fees from 13 counties and the remainder from another 11 counties that charged nominal fees.

These results suggest that future collection efforts may not involve significant financial transactions, but note that five counties charged excessive fees (three for attribute data and two for GIS data) that the project could not pay, and thus these data were not collected for this study. Furthermore, several communities waived or reduced their fees for this specific

collection effort after significant negotiation with the study team—waivers that may not carry over into subsequent data collection efforts.

As was the case with actual fee requirements, a much smaller than expected number of communities required data-sharing agreements with HUD, in the form of either accepting HUD’s template agreement or submitting their own for HUD’s approval. Nine counties required agreements, but two of these counties subsequently became nonresponsive and the data for one county were acquired through other channels without an agreement. Data-sharing agreements generally included similar use restrictions related to citizens’ privacy concerns, the counties’ liability for inaccurate data, or the publication of data at the parcel level.

No county explicitly stated that the publication of aggregated data was a prohibited activity, provided that all appropriate confidentiality measures were taken for individual parcels and parcel owners. Most counties verbally agreed that publication of aggregated data would not constitute a violation of their use agreements, but other use restrictions included—

- Placing any additional parties with access to the data (such as agents used to collect and standardize the data) under similar restrictions to those of the requesting agency (HUD).
- Agreeing to indemnify the counties against later use of the data (largely because of concerns of accuracy within the roll at any one time).
- Prohibiting any resale, lease, or transfer of the parcel-level data.
- Acknowledging the source agency in the county as the provider of the data whenever the data are analyzed or reported.

**Exhibit 6. Data Access: Fees, Sharing Agreements, and Use Restrictions**

Fees, Restrictions, and Waivers	Attributes			GIS		
	Percent Yes	Percent No	Percent NA or Blank	Percent Yes	Percent No	Percent NA or Blank
Fees for data	35	46	19	30	47	23
Fees can be waived	22	15	63	17	15	68
Fees can be waived with agreement*	11	14	75	NA	NA	NA
Fees are recurring	20	54	26	NA	NA	NA
Use restrictions	31	50	20	27	50	23
Data sharing agreement	13	69	1	NA	NA	NA
Template agreement	10	71	19	NA	NA	NA

NA = data not available.

\* The percentage of respondents who reported that the county charges fees for their data is significantly higher than that recorded in the study because counties either waived their fees for this specific study or their data were acquired via an alternative method.

<sup>8</sup> Data were purchased and collected for the city of Worcester, Massachusetts, the largest township within Worcester County.

### 2.3. Data Collection Response Rate

The collection effort resulted in the acquisition of data from 109 of the originally requested 127 counties.<sup>9</sup> Exhibit 7 shows that data from 80 of the 109 counties were either publicly available or readily provided to the researchers. In 21 counties, including counties that reduced standard fees for the purposes of the study, fee purchases of less than \$2,000 were required. In 5 counties, either the county approved the HUD MOU terms, or HUD approved the county’s MOU terms.<sup>10</sup> Finally, 3 counties had both fee and MOU requirements that were successfully transacted.

Data from 18 of the 127 selected counties were not collected for various reasons. Three counties had no electronically standardized data within the county, and the collection and maintenance of the data in 2 counties remain in the hands of municipalities and townships. These constraints exist in Massachusetts and, to a lesser extent, in Michigan, Illinois, and New Jersey.<sup>11</sup> Three counties charged excessive fees (from \$34,000 to \$1,000,000). One county required data sharing terms that exceeded those allowed by the HUD-approved template. Two

other counties formalized MOUs with HUD but did not deliver data. Finally, 7 either chose not to respond to any of the attempts to discuss data collection or were not able to coordinate data collection with the researchers for a variety of reasons.

#### 2.3.1. Dataset Types

Another way of viewing the data collected is in the broader category of dataset types in which it is made available by counties. Specifically, these types include the following datasets—

- **Parcel attribute data.** Data regarding the physical characteristics of parcels, ownership, and assessed values are commonly held by counties, and their collection is typically the charge of county assessors. Often, assessors refer to this data as “CAMA” data—in reference to Computer Assisted Mass Appraisal software that has been in use since the 1970s to maintain assessment data in government agencies. The implementation of this software is significantly tailored to each jurisdiction’s individual attribute terminology, definitions, and data needs.
- **Sales data.** Recent sales (including the names of involved parties and the price and date of the sale) are also commonly documented by counties. Counties in some states do not collect sales data because it is considered private; these states are referred to as nondisclosure states.<sup>12</sup> Most jurisdictions, however, do collect this information in some way, although not necessarily through the assessor’s office.
- **Parcel geographic data.** The geometry and location of parcels are critical identifiers of both the parcel and its key parcel attributes (such as area). These data are commonly kept in GIS software, which is becoming an essential technology and can be found in nearly all communities. The minimum population of a community that can support this technology is estimated at about 30,000. All communities in this project were above this threshold, but only one of those counties where the management of property records occurs at the township level maintained geographic data at the county level (DuPage, Illinois).

**Exhibit 7. Data Collection Response Rate and Requirements**

Category	Number of Counties	Percent of All Counties
Responded readily*	49	39
Additional negotiation (no fees or MOU)	31	24
Fees only	21	17
MOU only (HUD or county)**	5	4
Fees and MOU	3	2
<b>Acquired (original 127)</b>	<b>109</b>	<b>86</b>
Fees rejected by HUD	3	2
MOU rejected by HUD	1	1
Data not available	5	4
Nonresponsive counties**	9	7
<b>Not acquired (original 127)</b>	<b>18</b>	<b>14</b>
<b>Additional counties acquired</b>	<b>278</b>	<b>NA</b>

MOU = memorandum of understanding. NA = data not available.

\* Includes publicly accessible downloads and inaccessible downloads that were readily provided to researchers.

\*\* Two counties signed MOUs but did not provide data during the extended collection timeframe. Both are listed as “nonresponsive.”

<sup>9</sup> Data from an additional 278 counties were collected from statewide sources. These data were also standardized for each county, but their collection effort was not recorded for this study.

<sup>10</sup> Of the 5 counties, 2 successfully underwent the MOU negotiation process with the researchers and HUD but ultimately did not provide their data. These counties are listed in exhibit 3 as “nonresponsive.”

<sup>11</sup> All parcel data in Massachusetts were collected and maintained at the municipal or township level at the time of collection. In Worcester, researchers collected and acquired the parcel data from the primary municipality in the county directly. Several counties in Michigan and Illinois faced similar conditions, but had GIS data at the county level.

<sup>12</sup> The current nondisclosure states (that is, those that do not require collecting sales data) are Idaho, Indiana, Kansas, Mississippi, Missouri, Montana, New Mexico, North Dakota, South Dakota, Texas, Utah, and Wyoming. The prohibition applies only to sales data, not to parcel attributes or geography.

- **Historical records.** States usually mandate at least 3-year retention of tax records because of tax foreclosure time-frames, although the trend is to keep records longer.

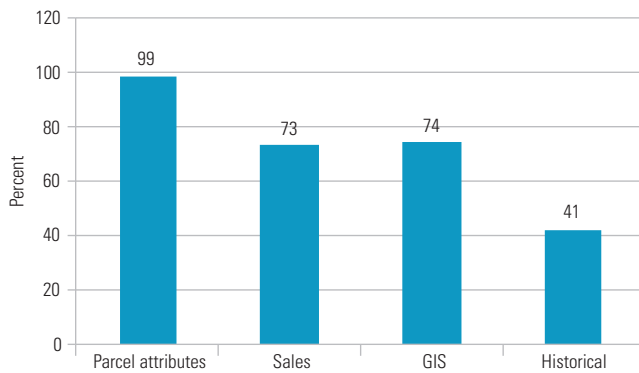
The collection rates for these datasets across all study counties are shown in exhibits 8 and 9, and detailed information for each individual county is in appendix G.

According to the subject-matter experts consulted in the project, the final rate of data collection for the selected counties during this study—86 percent—was comparable with that of similar data collection efforts; response rates from those studies ranged from 80 to 85 percent.<sup>13</sup>

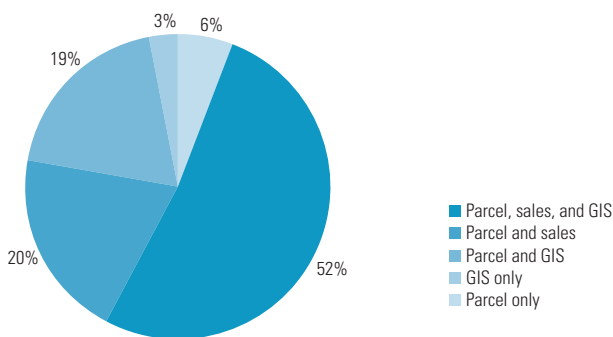
### 2.3.2. Data Attributes

Despite the high response rate, the comprehensiveness of the collected data varied significantly. HUD originally requested a set of attributes based on the expectation that they could be

**Exhibit 8. Percent of Study Counties Providing Specific Datasets**



**Exhibit 9. Percent of Study Counties by Total Dataset Availability**



collected. These attributes were defined by categories that ranged from “highly likely” to “unattainable.” Exhibit 10 shows the original expectations for the availability of desired attributes. The likelihood of any single attribute being available depended on many factors, particularly its importance in the assessment process or whether it had been requested previously and frequently during other data collection efforts.

**Exhibit 10. Expected Availability of Original Data Attributes**

Parcel identifier	Highly likely
Parcel identifier year	
Site address	
Area of lot	
Assessed value	Probable
Neighborhood	
Area units of measure	
Land use code	
Year built	
Structure type	Possible
Assessment date	
Assessment timeframe	
Assessment basis	
Sales price	
Sales date	
Structure area units	Less possible
Area of Structure	
Multifamily unit	
Owner type	Partial
Sales type	
Condition of property	
Condition year	
Easements	Unattainable
Liens	

### 2.3.3. Attribute Availability

Exhibit 11 provides a list of the attributes and the percentage of counties that provided them (including the additional counties beyond the 109 original sample counties). A detailed accounting of each of the collected counties’ attributes is provided in appendix G.

Exhibit 12 shows the same actual attribute response rates for all counties collected as presented in exhibit 11, but using the same color scheme as presented in exhibit 10 (black response rates are new attributes added by the research team using other local resources during the data standardization effort).

With only a few exceptions, most of the desired attributes’ actual availability met expectations. The final collection of specific attributes is presented in the following discussion based on the original expectations for their availability. Definitions for each attribute are also provided.

<sup>13</sup> For example, Stage and von Meyer (2010). Fairview Industries’ data collection experience is based on collecting parcel data in more than 400 counties over a 5-year timeframe.

### Highly Available

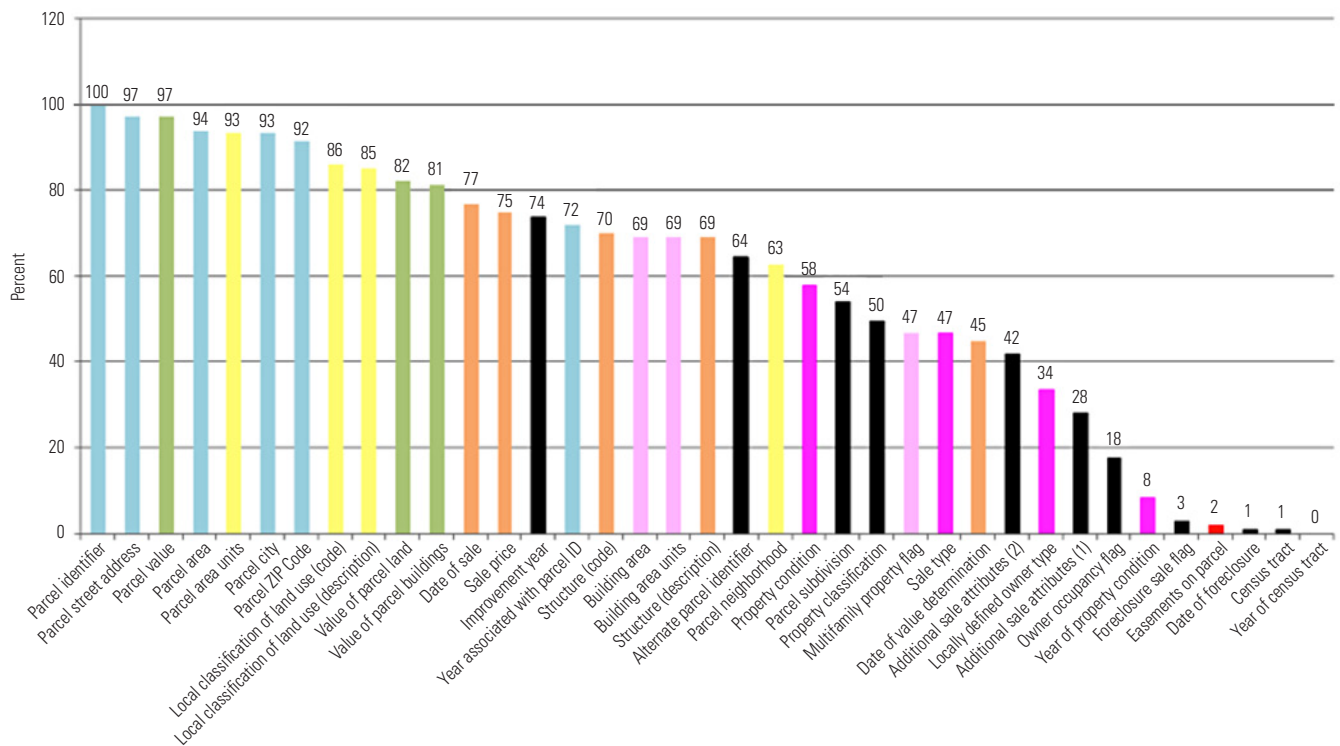
Highly available attributes, not surprisingly, are those that are fundamental to the assessment processes used by counties.

The only attribute that did not conform to the study’s original assumptions (that is, accessible from at least 90 percent of counties) was the parcel identifier year.

**Exhibit 11. Attributes Collected, Percent of Counties**

Attribute	Percent of Counties	Attribute	Percent of Counties
Parcel identifier	100	Structure (description)	69
Parcel street address	97	Alternate parcel identifier	64
Parcel value	97	Parcel neighborhood	63
Parcel area	94	Property condition	58
Parcel area units	93	Parcel subdivision	54
Parcel city	93	Property classification	50
Parcel ZIP Code	92	Multifamily property flag	47
Local classification of land use (code)	86	Sale type	47
Local classification of land use (description)	85	Date of value determination	45
Value of parcel land	82	Additional sale attributes (2)	42
Value of parcel buildings	81	Locally defined owner type	34
Date of sale	77	Additional sale attributes (1)	28
Sale price	75	Owner occupancy flag	18
Improvement year	74	Year of property condition	8
Year associated with parcel ID	72	Foreclosure sale flag	3
Structure (code)	70	Easements on parcel	2
Building area	69	Date of foreclosure	1
Building area units	69	Census tract	1

**Exhibit 12. Actual Attribute Availability**



- **Parcel ID.** The parcel identification number (PIN) serves as the primary index for the data. Because counties need to distinguish parcels regardless of the attributes they track for those parcels, parcel identifiers are essentially always available when the data themselves are available. If a working roll and the GIS files are not synchronized, parcel identifiers may be missing from the attribute file, although this occurrence is rare.
- **Site address.** The availability of the site address, other than its value for spatial location, is an indicator that the data are coming from mature systems. The assessor collects the mailing address for tax bills to provide to the tax collector. Collecting the site address for a parcel is typically one of the first activities that a local government takes on after the parcel database has been completed. Although the site mailing address is not essential for assessment purposes (and therefore certain subattributes such as “city” or “ZIP Code” are occasionally omitted), it has high value for disaster response, fraud detection, data quality, enhanced census operations, and other uses via data matching across local, state, and federal data systems.
- **Lot or parcel area.** This value is an essential component of the assessment process. If it is not provided in parcel attribute data, it can often be calculated from the GIS data.
- **Assessed value.** The various attributes related to the assessed value of land and any buildings on them (both in general and for taxation purposes) is a primary responsibility of the assessor’s office and is rarely omitted if parcel attribute data are provided.
- **Parcel identifier year.** This attribute, which establishes the first year that a parcel is incorporated into the assessment, was generally viewed as common knowledge by assessors because the first year was identified in their historical data. Although inclusion of this attribute would eliminate such cross-checking and improve the efficiency of the dataset, a surprising number of counties did not maintain this field. The limited availability of this attribute was one of the most significant variances from expectations.
- **Land use code.** These attributes are usually similar within a state but can vary considerably among states. Code definitions are needed to make use of the codes because their definitions vary significantly among jurisdictions. After they are acquired, they are fairly stable across time.
- **Neighborhood or subdivision.** These geographic areas are collected as either a part of the legal description or an indication of assessment neighborhood, meaning properties of like characteristics that are assembled into assessment groups. This attribute is not an often-requested field and may be left out of prepackaged datasets. Similar to those of land use codes, the definitions and parameters for this attribute are often particular to the jurisdiction. For county-level datasets, however, this attribute is collected at a lower level than expected.
- **Units of measurement.** The units of measurement are a critical part of a parcel’s attribute information simply because of their relevance to standard reporting methods. They may be considered common knowledge within a local government and not included in the provided datasets, so documentation often must be requested.
- **Year built.** The indicated date that a structure is built depends on how the assessor manages the parcel attribute information. Depending on the county, this year can be the date of the last improvement or it can be the original year of building. The trend in modern systems is to retain both dates. If historical records are kept, only the latest improvement is provided. The trend to include history reflects a mature and well-maintained parcel attribute system and would be more common in urban jurisdictions. Fewer jurisdictions reported this attribute than had been originally expected, however.

### Possibly Available

Some attributes that were originally expected to be possibly available (that is, available for between 70 and 80 percent of responding counties) ultimately were not. Although they likely do exist, many of these attributes, such as structure type, are not often requested. For counties that prepackage data for external requests, then, these attributes are often not included. For a few, the information may not be collected (for example, sales price) or the data may be common knowledge (for example, assessment basis) and maintained in external documentation. The likelihood of collecting these attributes generally matched the assumptions, with a few exceptions. Structure type exceeded expectations, and assessment processing data attributes (date, timeframe, and basis) fell short of expectations.

### Probably Available

These attributes are fundamental components of the assessment process, but considerable variation exists among the counties as to how this information is collected, and these fields often require interpretation and standardization. As predicted, all of the attributes assumed to be probable (that is, having a response rate of between 80 and 90 percent) fell within those ranges.



- **Structure type.** Structure type commonly refers to the physical building type on a parcel (for example, lowrise multifamily building), although it often overlaps with land use values. Although usually collected, this attribute is not frequently requested information and often is not provided in prepackaged datasets. Although related to land use, this attribute’s definitions are less variable. Specific requests for this attribute increased its collection rate significantly.
- **Sales price and date.** These attributes are often perceived as fundamental to the assessment process, and the trend in local government is to retain sales history to assist with the revaluation processes. Counties in nondisclosure states do not collect sales information, although occasionally this information is requested on a voluntary basis. Given these restrictions, these sales attributes were initially believed to be less available. However, since last-sale value is often provided in prepackaged datasets, the study had higher response rates for sales attributes than had been expected.
- **Assessment basis, date, and timeframe.** These three attributes are often viewed as common knowledge within assessor’s offices. In certain circumstances, they are required by mandate to be consistent within a state. In others, they are simply tacit knowledge that is not documented and cannot be acquired other than by personal interviews. Because the number of counties for which we were able to perform interviews was smaller than the number of counties from which we are able to actually collect data, the response rates for these attributes were significantly lower than expected.

### Less Possibly Availability

These data attributes were not expected to be readily available or were not expected to be collected in many jurisdictions. Despite our predictions that high response rates would be difficult, however, rates ranged from 64 to 73 percent of the collected counties.

- **Structure area and units.** These fields are often not a part of prepackaged datasets and must be requested. In more urban environments, this information may be more readily available, explaining the higher than expected collection observations for both attributes.
- **Multifamily unit.** This data element is often embedded in land use codes. Because data definitions for this attribute vary widely, this study was unable to distinguish which land use identifier necessarily signified this type of housing. Furthermore, some jurisdictions defined multifamily parcels as having two or more units, whereas others use the HUD convention of four or more. An additional challenge for

multifamily unit identifiers is that these structures are often classified as commercial property. For these reasons, collection observations for this attribute generally met but did not exceed our expectations.

### Partially Available

Information in this category often must be collected from other available data sources or interpreted from those sources with assessment data. Early on, we assumed that these attributes would be more difficult to acquire. Attributes related to the owner and sales types, however, slightly exceeded expectations and were collected at nearly the same level as sales-related information in the “possibly available” category. Property condition-related data, however, was virtually unavailable.

- **Owner type.** Information on the entity owning a parcel can be derived from the land use code, the tax exempt code or, simply, the owner name. If the land is owned by a nontaxable entity, then the type of entity (federal, state, or local government) might be coded. Otherwise, individual names might be used. In all cases, the classifications for owners vary significantly among jurisdictions and are less likely to be included in available assessment data. The low expected availability of this attribute was confirmed by the actual data collected.
- **Sales type.** The sales type is not always collected, although the trend is to capture this information to assist with the assessment process. The response rate (78 percent) was much higher than expected for this attribute, suggesting that this trend is real.
- **Condition of property and condition year.** The condition of a property is noted in a site visit assessment, although typically site assessments are not conducted on an annual basis. Even in urban areas where reassessments may be done annually, an accompanying annual condition site visit is not common. Property conditions are more likely to be collected in other databases (such as municipal health or property code enforcement offices), but, across all of the cases, the definitions and terms for conditions varied tremendously.

### Unattainable

In general, fields related to specific legal or financial conditions of property owners (such as liens, easements, and foreclosures) were not expected to be a part of or relevant to the assessment process, which was, in fact, the case. This information is often kept by different agencies, such as a sheriff’s office or local courthouse. Although this limitation was expected, it should be considered in the event that these attributes are desired in future data collection efforts.

It is worth noting, however, that there is a growing trend to include parcel identifiers on ownership transfer and mortgage documents that might include these attributes, which would enable these records to be linked to the assessment and GIS data. For this reason, we had some, albeit minimal, collection observations for easements and foreclosures, although none for liens. It is expected that subsequent, ongoing explorations with counties may yield improvements in the rate of collection for these attributes, particularly those attributes that were internally viewed as common knowledge or that might be acquired from other sources.

## 2.4. Feasibility Survey Response Rate

While data were collected from 109 of the original 127 selected study counties, the number of counties responding to the feasibility questions was much lower. Many counties simply provided data publicly (for example, online), and staff was not available to respond to requests for interviews. In other counties, staff chose not to respond to either all or a portion of the feasibility questions. We did not try to interview staff in counties that did not provide data. With response rates for some questions as low as 24 percent, the researchers supplemented this information with subject-matter knowledge regarding the internal processes and data in county assessor’s offices. Appendix J provides the responses to the feasibility questions.

Two key observations are worthy of exploration.

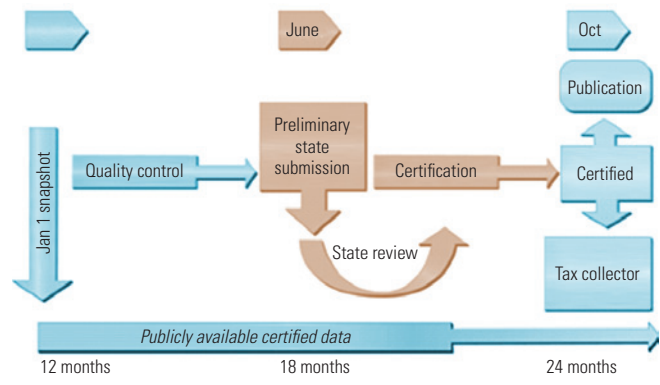
- The assessor’s data collection and certification process and its effect on the data’s timeliness.
- The staffing and resources available in county offices to respond to data collection efforts (particularly at the time of this study’s collection).

### 2.4.1. Data Timeliness and Comprehensiveness

Approximately one-half of the selected counties provided the certified roll, and one-half provided the working roll during data collection based on the time of collection. The working roll is regularly updated, but may not be synchronized with other data regularly throughout the year (exhibit 13). By contrast, the certified roll includes the data that are provided to the tax collector at a snapshot in time (usually January) and then synchronized with other data, verified, and finally published. This process can last from 6 to 10 months—meaning the data can be at least 6 months old by the time of publication. The working roll is the most current data with updated sales and owners but is not verified and may not include assessment data, which would be determined at the time of verification.

Aside from issues related to the age of the data, counties were often prepared to deliver standard prepackaged datasets that did not include all of the attributes HUD needed (although they almost invariably contained the “highly available” and “probably available” attributes described previously). For this reason, negotiating for and receiving the most appropriate and timely datasets from counties depended on the study team’s resources, the individual county’s resources, and the ability to effectively communicate which datasets contained which attributes. Given the desire to determine the most comprehensive data that could be acquired with minimal resources, the preference was to accept prepackaged data. For these counties, this minimal level of effort was used in the later calculations for the feasibility of future data collections, although, conceivably, more comprehensive data could have been collected with significant additional effort.

**Exhibit 13. Development Schedule for a Typical Assessment Roll**



### 2.4.2. County Resources and Availability To Participate

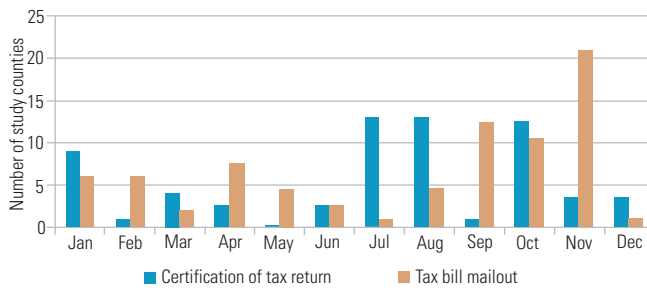
Some counties’ lack of interest and availability in responding to feasibility questions added challenges that influenced the long-term projections. Of the nine counties deemed nonresponsive, no observable characteristics in our target population (for example, size of the county’s population) would enable us to predict which counties are not likely to provide data in the future. Although it can generally be said that all governments are currently strapped for resources, larger jurisdictions typically have more flexibility and are more able to allocate resources to support data sharing. Many of the larger local governments have staff that are specifically tasked with providing access to data, whereas in the smaller communities staff must multitask, which might lead to the expectation that large counties are more likely to provide the data. The larger governments, however,



often had correspondingly large bureaucracies, complicating the number of initial contacts to get connected to the primary contact and adding additional time for internal approvals before data were shared. The size of the counties in our sample was not correlated with whether or not data were provided. As noted in chapter 2.1.1, however, the study counties are not representative of all U.S. counties with regard to population. The reason for nonresponse, then, might simply have been resource limitations that are characteristic of counties of all sizes.

One commonly mentioned reason for not providing data was the timing of the data collection in relation to the workload of assessor’s offices. Although the schedules vary slightly from state to state, the four peak periods for the assessment staff are (1) the acquisition of the snapshot that will be used to create the certified roll, (2) the preliminary reports to the state (if required) based on the certification, (3) notification of proposed tax rates to property owners and associated hearings and appeals from the assessment values, and (4) the production of the certified roll itself. As shown in exhibit 14, most counties approach their major deadlines at the beginning of the year (January), the beginning of summer (June and July) and, most significantly, in the fall (September through November) for notification and roll certification. Because this final, most intense assessment period occurred at the same time as data collection for the project, the response rate may have been higher if data collection had been undertaken at a different time.

**Exhibit 14. Assessor Activities by Month (Distribution of Study Counties)**



## 2.5. Data Standardization

Toward the end of the data collection process, the researchers were charged with standardizing data within each county’s datasets to the set of key attributes requested by HUD. The sample data obtained from the pilot counties provided the basis for constructing a standardized parcel database shell and standardized formats (appendix K). This shell was provided to HUD before standardization and became the basis for exploring each individual dataset from the selected, responsive counties.

Early in the process, however, numerous challenges for standardization became evident, many of which could be only partially overcome. The challenges stemmed from issues with the quality of the counties’ original data, incompleteness or inconsistency across their datasets, and the lack of standard attribute definitions and standards across all of the counties. The following are some of the more specific issues.

- **Internal validity problems.** Information within the county records was sometimes conflicting or implausible. For example, the land usage field might classify a specific parcel as multi-family residential, whereas the number of dwelling units may be less than four (classifying the parcel as not multi-family). Identifying outlier and illogical values is difficult or even impossible for someone unfamiliar with each area’s unique real estate market and assessor data. For example, the county definition of parcel classifications is not always the same as the HUD definition. Correcting for possible internal validation errors was beyond the scope of the project
- **Inadequate or missing date attributes.** Although information regarding the year in which the data were last updated was often collected, often no indication was given of the specific year or date that property valuations or condition inspections took place, degrading the quality of the valuation and property condition data collected.
- **Incomplete or missing data documentation.** County-supplied data documentation—including data attribute definitions, descriptions of variables and database layout, and lookup tables for coded attributes—was often incomplete or missing entirely. Lack of essential documentation made it difficult to understand raw data extracts and thus to select or impute HUD-requested attributes from available raw data. Lack of lookup tables made it difficult or impossible to interpret coded attribute fields. To compensate for this absence, the standardization team had to use other available data collection resources—such as county assessor method documents, publicly accessible assessment descriptions, and, when possible, direct responses from county staff—to obtain necessary lookups for coded fields. For some basic geographic attributes (such as parcel city or ZIP Code), the team had to impute values using GIS software.
- **Presence of similar or interchangeable attributes.** In many cases, especially among single-family residential properties, a single attribute or group of attributes provides sufficient measures for HUD-requested attributes. For example, structure description, land use, and property class often simultaneously measure land activity—in this case, a single-family residence. For those instances in which values were clear, the appropriate attribute values were incorporated.

- **Duplicate records.** Duplicate parcel identifiers (attribute data) or parcel identifiers combined with transfer identifiers (sales data) in raw data files made it difficult to link files. Such duplication could result from—
- **Subparcel level of observation.** Many counties’ data included files in which attributes were measured by structure rather than by parcel. If multiple structures exist on one parcel, then the parcel identifier is duplicated. In response, our standardization team linked all parcel-level data (that is, datasets unique by the parcel identifier), then linked the merged datasets to structure-level data in a many-to-one merge to gain single parcel-level values.
- **Multiple observations for the same parcel.** On occasion, attributes such as property valuation or acreage could be further split depending on land use, structure, or valuation year. For example, total valuation could be split into commercial value, residential value, and total value based on a parcel’s specific use. In this case, each land type would have a separate valuation record, creating parcel identifier duplication. The standardization team only included records measuring total valuation in the standardized primary attributes file. If a valuation was split based on valuation year, the most recent record was used for the primary attribute data file and the others for the historical attribute files. If records were split into land- and improvement-level valuations, one record for land value, one for improvement value, and the sum for total value were used.
- **Data entry error.** On occasion, parcel duplicates appeared for no obvious reason. This situation occurred most often in sales data for which parcel-transfer records appeared multiple times, presumably because data entry staff in the county corrected previously entered data by entering a new line of data for the same parcel and sale. The standardization team included all duplicates in the sales data file and the last occurrence of a duplicate in the primary attribute file (assuming that the record entered last was the most up to date), along with a flag denoting duplicate sales information.
- **Substructure level of observations.** Rarely, HUD-requested attributes were collected on a substructure level. In these cases, different elements of one structure (for example, the main area, garage, and enclosed parts may be elements of a single-family home) were each given their own record, creating many duplicate parcel identifiers. In the event attribute values were numeric, the team summed up the substructure measurements to the structure level. For character attributes, the attribute associated with the main area of the structure was included.
- **Inadequate match among datasets.** When linking multiple datasets, some level of mismatch often emerged (that is, some parcel records are included in one of the linked datasets and not in the other or others). Often, mismatch occurs when linking non-GIS tax roll or sales data to GIS files. In a few cases, mismatch was severe; in Pima County, Arizona, parcel records between datasets mismatched in more than 20 percent of the records. To address this challenge, all attributes were included in the primary attributes file when two datasets of equal importance were merged (that is, two datasets from which multiple important attributes were drawn, such as basic parcel and valuation datasets). However, when merging two datasets of unequal importance (that is, one dataset from which many attributes were drawn and one from which only a few attributes were drawn), the team excluded the records in the less important dataset that did not link with records in the more important dataset.
- **Parcel fracture and consolidation.** Datasets created at different times had different parcel universes. Parcels are not immutable; they are joined, split, redrawn, and relabeled with sales and other legal transactions. In several cases, a county-supplied attributes dataset and a separate sales dataset would not merge well because one dataset was outdated relative to the other. In certain cases, counties kept records of previous parcel identifiers to help navigate these cases. In other cases, however, no appropriate linking identifier was provided.
- **Presence of multiple valuation attributes.** Some counties provided many valuation attributes, such as “appraised value,” “assessed value,” “market value,” “cost value,” “income value,” and “special assessed value.” Combined, again, with the lack of documentation from all counties, the presence of multiple valuation measures made selecting an appropriate assessed value measure difficult even for assessment data experts. In the standardization phase of the project, the two most commonly provided attributes (“appraised value” and “assessed value”) were included in the primary attributes file with flags noting the presence of additional valuation measures in the county-supplied data.
- **Conflicting terminology.** For several attributes, little one-to-one correspondence (bijection) took place between HUD-requested attributes and county-supplied attributes. This challenge was also compounded by different measures or scales across the county-supplied attributes. For example, HUD requested “condition of property,” but a county may instead supply “property grade” (a variable with a measure on an undefined scale from A to F), “construction quality” (a variable from 1 to 5), or “improvement condition.” Even

in the minority of cases in which data documentation was fully available, it was unclear both which of these county-supplied attributes should be mapped to the “property condition” field in the standardized dataset and how to combine them so as not to lose information. In these instances, the standardization team retained the original classifications as provided in their data. These disparate classifications are one example of how possible comparisons between counties are severely limited.

- **Inadequate ability to compare data across counties.** Although beyond the project’s scope, the standardization team made an effort to see if data standardized within each county to the same data format could then be compared across counties. Even when counties collect the same or similar information, however, the data are often incomparable across different counties. For example, counties may use different valuation methods or provide multiple valuation measures for different internal uses, making it difficult to compare even fundamental property values between counties. Another common difference is in the way counties judge property conditions and record land usage, making these attributes essentially incomparable across counties. This situation not only limits the potential for analysis across counties using these data, but also poses a critical challenge to analysis efforts on key attributes in future data collection efforts.

The number of attributes with definitions that varied across the study counties—and the extent of those variations—was remarkable. For this study, these limitations mean that any future analysis with the collected data can only be county-specific. Even if resources could be expended to acquire the full data documentation from each county and perform additional analysis comparing them, the counties will generally continue to apply their own definitions and data collection processes. No consistency exists even across the handful of states that have required consistency in terms and processes of their counties. Until a full standardization process is operationalized nationally, an aggregation effort can only identify values within a county’s dataset that match a prescribed list of attributes, refer to those values as that county’s attribute, and attempt to format the attributes in a similar appearance. Any analysis or reporting of these data assuming them to be fully standardized would be exploratory at best and erroneous at worst.

## 2.6. Lessons Learned

The study demonstrated that the collection of data across a sample of counties (and possibly the entire population of U.S. counties) is feasible, albeit with some effort. With some considerable effort, each county’s data also can be made internally

uniform to a set of general attribute definitions. The resulting data, however, cannot and should not be used in cross-county analysis because of the challenges noted previously.

In addition to the lessons learned with regard to the comprehensiveness and accessibility of specific content within the datasets, the process of collecting the data provided numerous lessons for future data collection efforts. These lessons are useful both for practical implementation of future data collections and in helping predict the level of effort for future collections.

- **Developing relationships.** This first iteration of data collection involved almost entirely cold calls to counties to identify the primary contact. Such calling without introduction results in less productive responses, oftentimes fees (or higher fees), and an extended data collection timeframe. In general, data collections repeated over years would reduce the abruptness of such calls and lead to longer term relationships between the county and the requesting entity. These relationships could lead to a qualitative difference in the response from the counties, and a quantitative difference in the time spent negotiating and soliciting data, when the county official is familiar with the caller from a previous project, a personal connection, or some past relationship. This relationship could be personal in certain circumstances (for example, in smaller, more rural counties). Longer term institutional relationships would ideally be all that is required for most counties with professional staff to gain this collection efficiency. Initially, however, personal relationships are likely to be required.
- **Scope.** The responsiveness in the selected sample of counties varied significantly. If the data collection had been restricted to a more limited timeframe or cutoff time per county, the collection might have been more efficient (although efficiency was not the purpose of this study). For future collections, the selection of a limited number of counties known to be more responsive could lead to a more efficiently collected, albeit less comprehensive, database, should one be needed.
- **Subject-matter expertise.** Preliminary discussions with counties, particularly when discussing the content and processes related to data collection, often required expertise in understanding the nuances of assessment data, including timeliness, collection cycles, accuracy, fitness for use, and lineage. This tacit information about the assessment process tends to get lost when data are aggregated and compiled or when data are obtained from shadow sources. This situation can be remedied either if datasets contain information describing the last assessed date and data explicitly or if questions similar to the study’s feasibility questions can be asked periodically during future data collection efforts.

- **Contact.** The methods used to reach counties and identify the data collection effort and agents were noted as a deterrent in certain counties. Improvements could include—
  - Using a HUD envelope when sending out the background information and introductory letters. Counties recognize the HUD return label and official envelope and may be more likely to respond in a timely manner.
  - E-mailing the questions for the inventory database. Counties that requested the feasibility questions in advance were prepared for the answers.
  - Identifying all agents, including contractors, that may have contact with the county and its data helped improve credibility and reassure counties.
  - Sending thank-you letters to the counties that provided data, especially those that waived fees, was beneficial. Letters from HUD in particular may be helpful in future collections.
  - Avoiding contacting counties during their high-workload months will increase responsiveness and, likely, improve the chances of collecting in later years.
- **Standardization.** Lessons were learned from the standardization process, particularly when counties did not or could not provide any supporting data documentation. Limited to no opportunities arose to ask additional questions of the county staff during the study; the answers to many of these questions may have helped overcome key standardization challenges later. Improvements in this process could include—
  - Verifying the type of roll (working roll or certified roll) before processing, if possible. This information should be a priority if the opportunity to ask additional questions of the counties is limited.
  - Checking for repeated or duplicate parcel numbers, multiple or duplicated sales, and consistency of coded values and code descriptions in the raw data early on. With additional questions, duplication might be explained readily, beyond possible differences in rolls or for other reasons that may appear to be common knowledge to the county assessor in question.
  - Verifying that all codes in the coded value fields have an associated explanation or code description. Often, counties did not provide full descriptions for codes. Although past experience with these data may have led some counties to assume that these codes were common knowledge, variation in many of these codes across counties would suggest the need for better documentation.
  - Checking that the record counts match in all files, particularly across rolls and GIS data and between parcel data and sales and address tables. Larger discrepancies between counts could signify clear differences in the data held in the various datasets.
  - Verifying that the format and structure of date fields are consistent in all records. Even the method of formatting dates in data fields was inconsistent within some datasets and certainly was inconsistent across all of them. Confirming consistency of these data records early in the process could ensure subsequent programming to compensate.
  - Checking comment and note fields for content that might be missing from required standardized data fields. Often, counties may have maintained additional comment or note fields in which the necessary attribute was either elaborated on or, in some cases, maintained.

Perhaps most importantly, the standardization efforts in this study confirmed the fact that, regardless of the quantity of the collected data, the data cannot be standardized in the common analytical definition of the term, that is, manipulated in ways that include but are not limited to methods that ensure that—

  - The same definition for each attribute holds across counties and was applied consistently during each county’s respective data collection.
  - The data collection and entry processes are the same across counties.
  - The data in each county are internally valid (or the process to confirm the validity of the data is similar to that of other counties).
  - The data were collected within the same timeframe.
  - The data are documented in the same manner, using the same definitions and nomenclature to ensure the common meaning.
- **Timing of data collection.** Finally, information collection protocols should be coordinated and timed with the workload of assessor’s offices to ensure ideal data collection timeframes. For this study, clearance requirements added significant delay to the project startup and forced a data collection that coincided with peak periods of activity for local assessors, which may have reduced the response rate and quality of response.



## 2.7. Conclusion

As subsequent collection is done in these counties, the percentage of data that are successfully collected should increase, and the overall process should improve. The challenges in this initial collection aside, counties were for the most part willing and able to provide data to HUD. Several factors that contributed to this overall willingness should be mentioned in addition to the barriers for participation described previously.

- **Resources.** Although all organizations are challenged for resources, in general, the larger the population being served, the greater the resources that are available. The 127 counties in this project are for the most part larger communities, so it is assumed that, relatively, these counties have the resources needed to support data sharing and mature systems with GIS technology, as well as the staffing resources to coordinate with HUD agents.
- **Culture of sharing.** The difference in the willingness to share data can be set by policy—as is the case in Florida, with its strong public records law, or in Maryland, with its access-for-fee policy. States without a specified policy have much greater variability among their counties. On the whole, as demonstrated by the acquisition of parcel data

from 86 percent of the counties, local governments are willing to share data, and most (84 percent in our study) will share with minimal or no fees. These observations are for government-to-government sharing, and may not reflect the fees for other forms of sharing.

- **Infrastructure for sharing data.** Data sharing requests to many organizations, particularly when they are asked for elements that are not normally requested, are an additional workload for staff members who are already busy. Timing requests for information during periods of low activity improves the likelihood of a more positive response from the source organization. Some organizations are prepared for these data requests, and procedures have been established to reduce the effect on the office workload regardless of the timing of the request.
- **Mutual benefit.** Benefits to the county of sharing parcel data are typically indirect. Clearly articulating real benefits of data sharing to the data producer improves the responsiveness to data requests. As indicated by the results of this project, counties are more willing to share their data if they can justify the additional workload and are informed of the purpose and use of the data.

## 3. Feasibility of Ongoing Data Collection for Select Counties

This section reviews the projections for ongoing data collection from the 127 selected counties to set the stage for an expanded and possibly nationwide collection. These projections are based on the feasibility measures from the county interviews and the levels of effort in this study's data collection.

### 3.1. Feasibility Measures and Cost Estimating

The development of an assessment strategy for the current and future cost of parcel data acquisition has two components.

- **Classifying data sources.** Local government data sources are categorized by their availability and willingness to provide data. These factors affect the resources necessary to collect data based on the feasibility measures (if available) and county responsiveness. This results in a categorization of counties.
- **Resource requirements.** The recurring and nonrecurring resource requirements for parcel data collection used to predict levels of effort are based on the level of effort (in hours) for collecting and standardizing data in this project. These costs are determined by breaking the collection effort into its component parts and compiling information on the time required for each task. This aggregate measure results in a level of effort for each category of counties.

#### 3.1.1. Classifying Data Sources

Evaluating the status of local governments' readiness to provide data can be divided into two additional components, (1) the availability of the data in a format that can be used by the U.S. Department of Housing and Urban Development and (2) the willingness of local officials to provide data. Classifying the degree to which a county is willing to provide data is a useful means for tracking the current status of parcel data sharing and for determining where to allocate available resources. County population is used as a preliminary indicator of ability; that is, data from counties with larger populations are assumed to be more available.

The use of county population as a gauge for parcel data availability follows a rule of thumb used by subject-matter experts involved in the GIS area. Communities with populations of at least 30,000 are likely to maintain and use GIS. In counties with populations of 100,000 or greater, it is highly probable that GIS is available; in counties with populations of more than 300,000, it is nearly a certainty.

This assumption is borne out by the 127 counties in question. Only 2 percent of the counties had populations of less than 100,000, 73 percent had populations of between 100,000 and 1 million, and 25 percent had populations of greater than 1 million. These counties' populations were much greater than the general threshold for having the resources to support GIS technology and, presumably, for having sophisticated staff and systems to provide data. Currently, approximately 82 percent of the parcel data in the United States is estimated to have been spatially enabled; only 3 study counties did not have GIS-useable data.<sup>14</sup>

Regardless, a county's true willingness to share its data can be determined only after the county has been contacted. Again, the proportion of counties (86 percent, or 109 counties) that were willing to share data conforms to the experience of other data collection efforts. Of the remaining 18 counties, many may be more successfully targeted in the future as well. For example, the 9 study counties that had electronic data at the county level but simply could not or did not respond to the data request may conceivably be approached at different times of the year. Likewise, the 4 counties that had excessive fees or were unwilling to share data may also reduce their requirements either through continued negotiation or their own choosing.

Study counties were categorized based on the types of county availability and willingness to provide data. For clarity, four categories of classification were developed based on the parcel data availability and the willingness of the county to provide the data. The categories and the classification of counties are summarized in exhibit 15 and depicted in exhibit 16.<sup>15</sup>

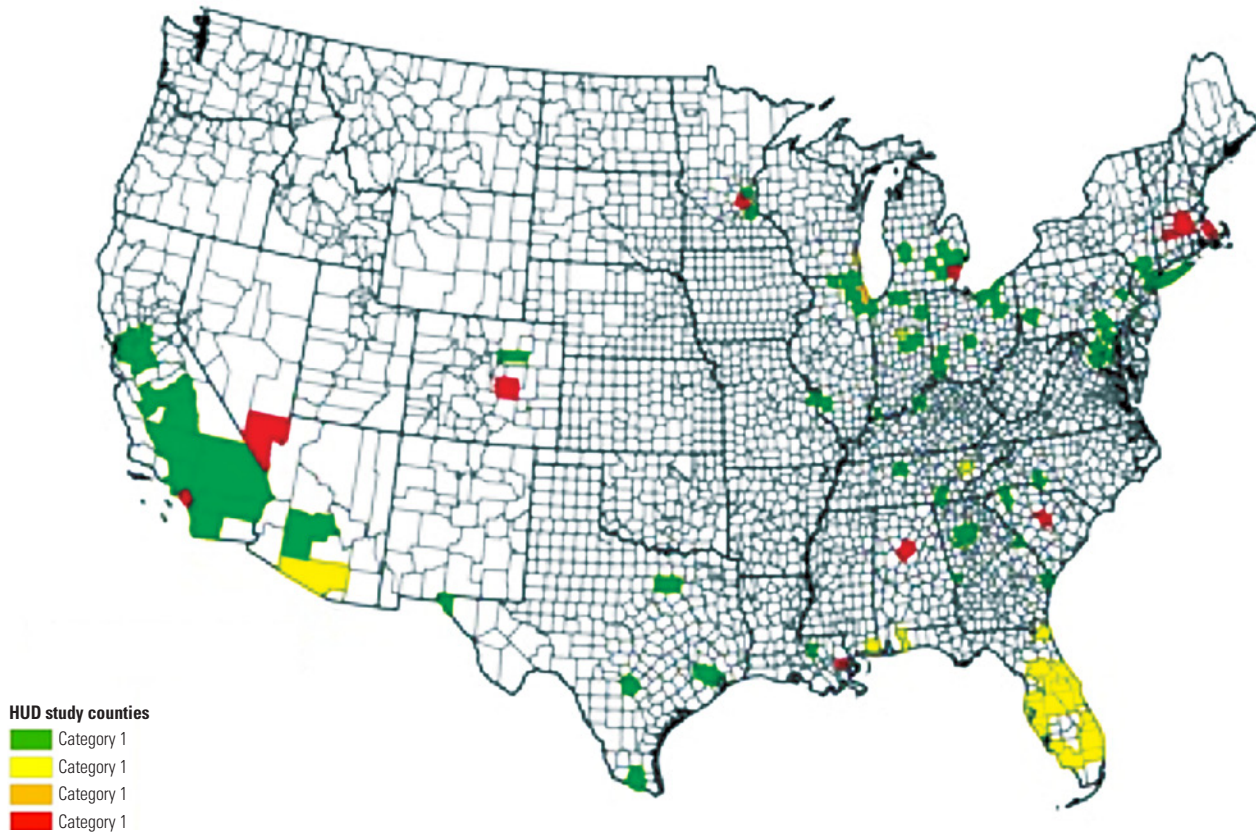
<sup>14</sup> Stage and von Meyer (2010).

<sup>15</sup> This system is a hybrid of the Federal Geographic Data Committee Cadastral Subcommittee's Level of Stewardship developed for the states, as described in appendix A.

**Exhibit 15. County Resource Categories**

Category	Study Counties	Characterization
Category 1—state access or local government web delivery	40	Parcel data is readily available in a prepackaged format from the web. Two types of data access are in this category: (1) Data that has been aggregated to the state level, which may be available from the state either through a single contact or through web access. (2) County data provided through the web that can be accessed without contacting the county. An agreement may be required at the state level, too.
Category 2—local government access only, but freely available	41	Individual counties must be contacted, but the parcel data are freely and readily available, possibly in a prepackaged format, or the county may need to produce a customized report. If they do not have a prepackaged dataset or a publication dataset, the data request may have to be created. Typically after a few repeat requests the county develops a standard report for the requester. The state of Maryland is included in the category even though they have a fee; going to a single source for the data has value enough to include them in this category.
Category 3—local government available with agreements or fees	31	Local governments have data and will share with an agreement. After an agreement is signed and/or nominal fees paid this usually suffices for follow-on requests. A Category 3 county may move to a Category 2 if fees or agreements are waived for follow on requests.
Category 4—unavailable	15	Counties either do not have the data or are unwilling to share it without excessive fees or untenable agreements. The number of counties that do not have data continues to diminish as the cost of technology continues to fall and parcel automation continues to rise. The number of counties without data in the larger counties of HUD interest is very small.
Total	127	

**Exhibit 16. Study County Classifications by County Resource Categories**





### 3.1.2. Resource Requirements

With the counties categorized qualitatively, the specific level of effort for a multiyear data collection of these counties can be determined based on this study’s data collection measures. The level of effort was determined by modeling a Category 2 (median ability and willingness) county using the results of this project’s data collection efforts.

The number of hours required to collect and standardize the data for this typical, or median, county was further broken down into institutional and technical tasks. Institutional tasks include contact, negotiations and approval seeking, and ongoing relationship building. Technical tasks include all activities associated with transfer, documentation, processing, and database maintenance. Exhibit 17 describes each of these individual tasks.

Exhibit 18 lists the level-of-effort projections for all county categories for subsequent years based on the estimated level of effort per county from the first year’s data collection (that is, this study’s average institutional and technical data collection level of effort). This calculation assumes a county-by-county data collection similar to that performed in this study.

The level of effort for each task decreases significantly from the first to the second year, and it gradually decreases to a regular maintenance level by the fourth year. Reasons for the reduction in effort over time include—

- **Institutional resource needs.** Most resource requirements in the institutional section are nonrecurring costs.<sup>16</sup> After the first collection cycle, the organizations are familiar with the data request and requesting agency, and the agreements usually carry into subsequent collections. There is always

**Exhibit 17. Definitions of Collection and Standardization Tasks**

Tasks	Description
<b>Institutional</b>	The tasks in this section focus on the establishment of the initial contact and agreements for sharing data.
Target source agency	Identification of the source agency authority and the technical staff that provide the data.
Acquire approval	Establishment of the terms of data sharing, development of agreements, and agreements on fees.
Institutional maintenance	Maintenance of contact with the organization and the adjustments that occasionally occur because of changes of policy, personnel, and other factors.
<b>Technical</b>	The data transfer, processing and standardization tasks.
Data transfer	Establishment of procedures for transferring the data, reminders, and followup with the counties if errors emerge in the data.
Data documentation	Documentation of the data source and acquisition of ancillary files needed to standardize the data.
Data processing	Conversion of the provided files into a standardized format.
Maintenance	Alterations in processing because of changes in the source data.

**Exhibit 18. Four-Year Level of Effort per Task (Hours)**

	Average Level of Effort															
	Category 1				Category 2				Category 3				Category 4			
	Y1	Y2	Y3	Y4+	Y1	Y2	Y3	Y4+	Y1	Y2	Y3	Y4+	Y1	Y2	Y3	Y4+
<b>Institutional tasks</b>																
Targeting source agency	2.00	0.50	0.50	0.00	2.00	0.50	0.00	0.00	3.00	1.50	0.00	0.00	5.00	5.00	5.00	5.00
Acquire approval	0.00	0.00	0.00	0.00	2.00	0.50	0.00	0.00	3.00	1.50	1.00	0.50	7.00	7.00	7.00	7.00
Institutional maintenance	0.00	0.50	0.00	0.25	0.00	1.00	1.00	0.75	0.00	1.00	1.00	0.75	0.00	0.00	0.00	0.00
<b>Technical Tasks</b>																
Data transfer	0.50	0.50	0.25	0.25	1.00	1.00	0.25	0.25	2.00	2.00	1.00	1.00	3.00	3.00	3.00	3.00
Data documentation	0.50	0.00	0.00	0.00	1.00	0.00	0.00	0.00	2.00	1.00	0.50	0.25	2.00	2.00	2.00	2.00
Data processing	1.00	0.50	0.25	0.25	6.00	3.00	1.00	0.75	8.00	5.00	2.00	1.50	8.00	8.00	8.00	8.00
Maintenance	0.00	0.00	0.00	0.00	0.00	1.00	1.00	0.50	0.00	1.00	1.00	0.50	0.00	0.00	0.00	0.00
<b>Total Hours</b>	<b>4.00</b>	<b>2.00</b>	<b>1.00</b>	<b>0.75</b>	<b>12.00</b>	<b>7.00</b>	<b>3.25</b>	<b>2.25</b>	<b>18.00</b>	<b>13.00</b>	<b>6.50</b>	<b>4.50</b>	<b>25.00</b>	<b>25.00</b>	<b>25.00</b>	<b>25.00</b>

Y = Year.

<sup>16</sup> *Nonrecurring*, in this instance, means “taking place in 2 to 3 years.” Maintenance activities will handle the occasional adjustments that need to be made.

some continuing work with the data sources to refine the procedures and to keep the contacts current. However, beyond the second acquisition the refinement or adjustments to procedures are minimal.

- Data acquisition and processing.** Data acquisition has ongoing costs because updates will need to be processed. The cost of documentation is significantly less after the first acquisition because changes are primarily edits to the original documentation and not wholesale redocumentation. Data processing remains at 50-percent levels in the second year, as the source agency may provide additional data that it did not provide in the first round or different files as the source agencies establish more efficient procedures to provide the data. The formats begin to stabilize after the second cycle until the total resource needs for processing become approximately 20 percent of the initial cost.
- Special request.** Ancillary benefits of establishing a data collection relationship with a county are the reduction of institutional barriers and expedited processing that might occur if additional information is needed. Such needs range from general updates, which often occur during an emergency response, or a data request for a special project. Time is saved by having not only the procedures in place to process the data, but also the critical path of acquiring permission to use the data. Knowing the technical staff that provided the data also greatly expedites the acquisition process. The level of effort for these requests is not provided in the following analysis, but can be assumed to be less than the initial-year request if there is an ongoing collection effort.

Exhibit 19 shows the summary county multipliers based on the previously detailed county calculations.

**Exhibit 19. Summary Level of Effort per County Category (Hours)**

County Category	Year 1	Year 2	Year 3	Year 4
Category 1	4.00	2.00	1.00	0.75
Category 2	12.00	7.00	3.25	2.25
Category 3	18.00	13.00	6.50	4.50
Category 4	25.00	25.00	25.00	25.00

**Exhibit 20. Multiyear Data Collection Level of Effort Estimates (Hours)**

County Category	Number of Counties	Year 1	Year 2	Year 3	Year 4
Category 1	40	160	80	40	30
Category 2	41	492	287	133.25	92.25
Category 3	31	558	403	201.5	139.5
Category 4	15	375	375	375	375
<b>Total hours</b>	<b>127</b>	<b>1,585</b>	<b>1,145</b>	<b>749.75</b>	<b>636.75</b>
<b>Estimated annual cost*</b>		<b>\$317,000</b>	<b>\$229,000</b>	<b>\$149,950</b>	<b>\$127,350</b>

N = 127 counties.

\* Assumes an average \$200-per-hour level-of-effort cost, including labor only.

Although the average level of effort for the acquisition of data from a county in the first year is 12 hours, this figure can vary by more than 100 percent if complications arise with the data and getting permission to use them. For Category 4 counties, this average level of effort was determined to be 25 hours based on the study level-of-effort measures. This preliminary level of effort can also be negligible if the data are provided on a public website and in a standard format, as occurred for the counties in Category 1 for this project (estimated 4 hours of effort).

Finally, these county projections for yearly level of effort can be multiplied by the number of study counties in each category to determine the total yearly level of effort (exhibit 20).

The total cost of the 4-year effort to develop a database that would require subsequent maintenance, then, is estimated at \$823,300, including only labor and using the same research design, attribute template, and collection protocol as used in the study. Nonlabor costs such as software and data storage capacity are not included.

### 3.2. Summary of Resource Requirements

Fortunately, the likely situation is not as resource heavy as the math implies. A percentage of these counties will eventually fall into categories 1 or 2 because of either technology system upgrades or likely changes in data fee and data-sharing requirements. Also, the growing trend among states is to collect and aggregate parcel data themselves. The resource needs for state-wide parcel data collection are significantly lower, not only because the number of states (50 states plus Puerto Rico and the District of Columbia) is significantly lower than the number of counties (3,221). Rather, state-level collection will already have some level of standardization that will likely be based on nationally accepted standards. Hence, an additional standardization savings will be realized.

Florida can serve as a best case scenario. Statewide data, both attributes and the parcel maps, were provided on a simple hard drive. The data were also provided in a standardized format. The collection time was approximately 5 hours. The effort required to perform the basic tasks that would apply to all

counties, including the time it took to complete documentation, create a crosswalk for standardization, and write conversion programs, was about 20 hours in total. The actual standardization and cleaning of each county's data required an additional 25 hours for both the study and the supplementary counties in Florida. In the end, the total time to collect and process the data was about 50 hours (or slightly less than 1 hour per county),

which contrasts with the estimated 804 hours that would have been required to contact and collect data from each of the 67 counties on an individual basis. This difference represents an approximate 94-percent savings. In all cases, the processes and findings from this study of 127 counties provide a framework for understanding how a national parcel database could be developed and alternatives for its development.

## 4. Feasibility of Nationwide Data Collection

This chapter looks beyond this study’s data collection approach by examining the institutional framework that enables or hinders a nationwide data collection effort, and then proposes alternative approaches for meeting the U.S. Department of Housing and Urban Development’s objectives in the long term. The framework and alternatives are considered in light of HUD’s practical needs for parcel data in the short-term and the broader objective of a national parcel data infrastructure that HUD is exploring.

In its 2007 landmark report, *National Land Parcel Data: A Vision for the Future*, the National Research Council (NRC) examined the need and approaches for nationwide parcel data from a generic, multiagency perspective.<sup>17</sup> The report emphasized the need for a coordinated and cooperative approach among all federal agencies for a nationwide, consistent, and regularly updated national parcel dataset. The assessment of nationwide parcel data collection discussed in this report builds on the premise of the NRC report but focuses on the parcel information needed by HUD—that is, the specific set of attributes related to parcel value and sales that is beyond the information identified in the NRC report. HUD’s interest in tying parcel attributes to census demographic and housing information is also considered.

### 4.1. Nationwide Parcel Data Collection Constraints

Three essential components are necessary to the development of a national parcel data system.

1. The data, including the standards and formats.
2. The institutions that create, manage, and provide access to the data.
3. The technology that supports the automation and storage of the data.

Each of these components affects the quality of the results and the approach to a national parcel data collection strategy.

#### 4.1.1. Parcel Data

According to the NRC, approximately 150 million privately owned parcels and approximately 10 million publicly managed parcels exist in the United States. The essential parcel attributes are ownership, value, and use.

- Ownership includes the current owner(s) with ties to the records that define land conveyance, rights, and chain of title, including limitations on ownership (for example, easements). The ownership information describes the full range of rights and interest in land. A system that contains all of the titled rights and interests and any encumbrances on the title is often termed a cadastre or a cadastral system. In most tax parcel systems, the ownership information is represented by the name(s) of the current surface rights owner and perhaps a link to the deed or other document of the most recent sale or conveyance.
- Value includes appraisal and assessment on land and improvements, sales information, and real estate tax rates, levies, and liens. Value includes an inventory or listing of all of the real estate features on the property, including structures and structure characteristics such as construction dates, size, and condition. Mobile homes that are not attached to the land are often excluded from property value and may be tracked as personal property or a vehicle. The valuation is based on the methods used for establishing value, such as market value, use value, or functional value. The sales information can include only the most recent conveyance and the type of sale, or it may include the history of sales.
- Use includes actual utilization as well as a categorization of the property. Most real estate tax systems establish a property use classification to support a valuation. These categories can affect allowable real estate tax credits, such as homestead or commercial tax credits. In some systems, zoning is included as a use, but zoning is more properly defined as a planned or allowed use rather than an actual use. Properties can have multiple uses, and this attribute is one of the most varied related to parcels.

<sup>17</sup> NRC (2007a).

Currently, the naming and defining of these parcel attributes exhibit little uniformity, and the effect of the variability and volume of information on a nationwide collection is formidable. Although this condition is rapidly changing due to the advent of national policy and academic efforts toward generating uniform standards, these standards are not necessarily operationalized in all counties. This state of affairs will pose a constraint to any national data collection, particularly one that seeks to compare counties. This situation is mitigated in states where state-imposed standards lead to a degree of internal consistency across counties, although this condition currently applies to only a handful of states.

#### 4.1.2. Parcel Institutions

There are well over 4,000 entities in the United States that are responsible for collecting and maintaining parcel data. These entities include townships in the New England states and counties or their equivalents in most other states. In some instances, most notably in Montana and Tennessee, parcel data may be maintained at the state level or by regional authorities. By and large, the local government entities that manage parcel data include—

- The clerk of courts or a register of deeds often manages the ownership information with an emphasis on document preservation, perhaps through scanning, and access to documents through indexing. An increasing trend is the use of scanning systems for these records, but the degree to which the content of the documents—such as grantor-grantee, rights conveyed, and value of conveyance—is automated in a retrievable database varies greatly.
- Assessor’s and treasurer’s offices often manage the value aspect of real estate. The assessment and tax management systems may be the same or separate. Typically, these systems are automated in databases; the systems, however, are commonly in a proprietary data structure, and it can be challenging to retrieve data for use outside the local system.
- Both assessors and local planning and zoning agencies often manage the characteristics and use information of land. In the assessor’s database, this information is compiled to support valuation and establish comparable properties. For the planning and zoning agencies, this information is compiled to support code enforcement and master planning. Extracting

specific property conditions from these datasets is not highly likely, but information on planned use is often automated in a GIS.

- Other legal and law enforcement agencies (such as the county sheriff’s department or county courts) are generally responsible for contractual, financial, and criminal records, such as foreclosure information. Due to the recent recession, there has been increased interest in sharing this data with other local agencies interested in parcel and property data.

The number of organizations with a hand in the creation, assembly, and maintenance of parcel data is daunting. In the original formation of the United States, the power to manage and publish landownership information was, by design, relegated to the lowest level of government, closest to the people. If the sheer number of organizations is not overwhelming enough, most local jurisdictions often have an elected official that oversees at least the land transaction information (the land records, notably the deeds), if not the assessment information as well. In many states, the assessment function is even more localized than the land transaction data, with local township-based assessment and county-based tax management.

In many states, a state department of revenue or state taxation management department provides guidance on the content, frequency, and procedures for assessments, sometimes even providing licenses to assessors. Some states equalize values across the state, adjusting local valuation to achieve equitable and balanced valuation across the state. These departments also receive copies of the local assessment information and even copies of automated tax maps. Typically, this information is provided on an annual basis, and some states have a specified format and content for these data. Some states (such as Florida) compile these data into standardized formats and provide either the entire datasets or summary statistics from the data. Although not authoritative, these states are trusted sources.<sup>18</sup>

One additional source for parcel data are those entities that provide shadow data—the term applied to parcel data that have been acquired from the county, and then modified or value added in some way that they can be mistaken for current data. It is not uncommon for questions to arise from, or even legal action to be taken against, the authoritative source based on information in shadow data. Typically, shadow data are datasets accumulated by private vendors, standardized and

<sup>18</sup> Authoritative data come directly from the creator or primary data source, such as the county assessor’s office. Trusted data are published data that are available from a trusted source. The authoritative source has an agreement with the trusted source for publication. For example, a state (trusted source) might aggregate and publish data from many authoritative sources (counties). See appendix C for full descriptions.

augmented with data from other sources, and then provided for sale. In some cases, the augmentation and standardization make the shadow data easier to obtain and use, even though they are further removed from the authoritative sources and the standardization may alter the original datasets.

If HUD were to pursue the collection of national parcel data, the constraints within the local sources may lead to decisions regarding the prioritizing of attributes (to minimize multiple data sources). More importantly, HUD might seek alternative sources, particularly states, but these sources are not uniformly available. HUD has relied on shadow sources for data in the past and may continue to do so at a cost. If HUD pursues its own data collection effort, however, it will also become a shadow data source. As such, legal requirements and definitions will need to be finalized and possibly negotiated with each authoritative source.<sup>19</sup> This complication may further limit the feasibility of a national dataset.

### 4.1.3. Parcel Data Technology

In a 2010 report by the Federal Geographic Data Committee (FGDC), Cadastral Subcommittee, it was estimated that nearly 82 percent of the privately owned parcels in the United States have been automated with some type of mapping technology (GIS or Computer Aided Drafting [CAD]).<sup>20</sup> This 82 percent covers nearly 93 percent of the population but only about 55 percent of the land area (summarized in exhibit 21).

Technology has matured much faster than the institutions that manage the parcel data. Before the Internet explosion, a national parcel dataset was merely conceptual. The closest analogy to this undertaking would be the U.S. Census Bureau’s Topologically Integrated Geographic Encoding and

Referencing (TIGER) dataset that provides a legal, administrative, and statistical geography for managing all of the U.S. demographics and roads, road names, and address ranges.<sup>21</sup> With the availability and sophistication of data services, a likely scenario would be to have each authoritative or trusted source provide parcel data as a web service. Data could be provided or replicated instantaneously and virtually. Data could be accessed as needed, and the consumer’s mapping would be updated at the time that the data service is updated. Early studies (notably the first few NRC reports; see appendix A) estimating the level of effort for developing national parcel datasets assumed that acquiring the technology would require a costly and difficult effort. Yet, technology has developed to the point of being an enabling rather than prohibiting concern.

Completing the parcel data mapping in the United States and building and sustaining data services would come with costs. These costs, however, are not as dependent on the technology as they are on the ability to identify authoritative sources and the institutional willingness to participate. Institutional willingness at all levels of government—from federal leadership, to state commitment to aggregation of data, to local data producers to automate and provide access to data—remains the single largest obstacle. Given the current public fiscal climate, costs are still a major concern.

## 4.2. Nationwide Parcel Data Collection Requirements

In addition to the previously described constraints on the data sources, critical requirements on the federal agencies’ part further define the content and quality of the nationwide data collection.

**Exhibit 21. Summary of 2009 Review of National Parcel Data**

Statistic or Measure	Approximate Count
Total number of nongovernment-owned parcels in United States	150 million
Total number of parcels in a “GIS ready” format	123 million (82%)
Number of counties with parcels in a “GIS ready” format	1,600 (50%)
Percent of population living in areas with parcels in a “GIS ready” format	93%
Percent of U.S. land area with parcels in a “GIS ready” format	55%

<sup>19</sup> For background on HUD’s legal authority for and history of collecting data, see appendix B.

<sup>20</sup> Stage and von Meyer (2010).

<sup>21</sup> Although the census collects base data from local, state, and private sources, these data have strict standards and submission requirements. The enumeration data are collected by census and are not aggregated from other sources. It is also important to note the Census Bureau has invested significant resources in the development and maintenance of its datasets. The TIGER dataset maintains 211,267 block groups, each containing an average of 39 blocks. ([http://en.wikipedia.org/wiki/Census\\_block\\_group](http://en.wikipedia.org/wiki/Census_block_group)); this number was confirmed by personal communication with the Census Bureau on January 18, 2012.



- **Government-to-government.** The national aggregation scenario is focused on the coordination of data sharing among government entities. Basing this scenario on G2G data sharing overcomes the problematic topic of highly variable public access laws. A review of each state’s public records policies and legal issues on data sharing is beyond the scope of this work, but these access laws will call HUD’s status as a legal entity into question.
- **Annual updates.** The national and state aggregation scenarios presume that an annual update of the parcel data will be sufficient for most applications. This assumption acknowledges that in case-specific instances, such as responding to a particular neighborhood or a disaster, more current information might be necessary for that identified project area, which will require a current update for the targeted area.
- **Reasonable fees and agreements.** The national, state, and need-by-need aggregation scenarios presume that only reasonable fees and reasonable data-sharing agreements will be pursued. Reasonable data-sharing agreements are those with terms and conditions that can be reasonably met through normal data use and do not impose excessive restrictions on derivative uses. Localities with agreements that are untenable for HUD or that charge exorbitant fees will not be pursued for data aggregation.
- **Standardization.** All scenarios presume that HUD will want the data structured in a standardized format. Standardization accounts for converting the provided data into a standardized structure (such as uniform nomenclature or attribute format), but it does not necessarily include internally validating data, deriving a complete dataset where data are not readily available, or spatially manipulating or converting data. For example, land use codes assigned by local jurisdictions cannot be readily interpreted and converted to a single national standard land use coding scheme. As described in chapter 2, this constraint must be carefully weighed before the pursuit of any additional data collection efforts.
- **Likely attributes and spatial data.** All scenarios are based on obtaining the attributes that are most likely to be available and not the more difficult or improbable attributes, such as liens and easements. Attribute information is more critical than spatial data for HUD at this time, and mapping data—that is, GIS data—would be obtained where available.

### 4.3. Nationwide Parcel Data Collection Alternatives

A continuum of possibilities exists for developing a national parcel data system. Four scenarios are presented as possible strategies for addressing parcel data collection for HUD and other federal agencies. Each alternative is discussed in terms of the data, institutions, and technology requirements.

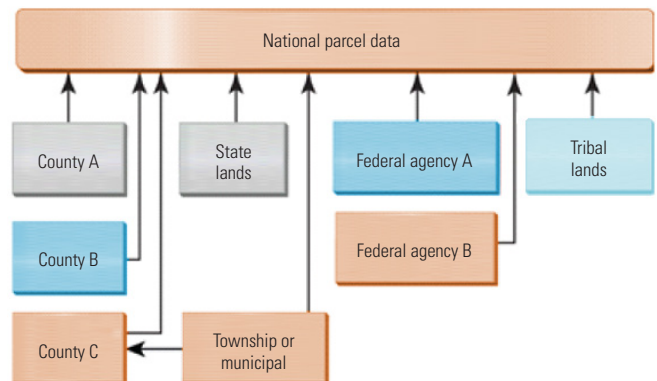
- National aggregation.
- Need-by-need aggregation.
- Private-sector purchase.
- State aggregation.

Three of these alternatives involve coordination with public-sector data producers and trusted sources of the data. The other option is a purchase option that relies on data aggregation by private vendors.

#### 4.3.1. National Aggregation

National aggregation in a county-by-county collection, following the same protocols instituted in this study, is one alternative. In this scenario, a designated federal agency would have a fully staffed organization that would manage and coordinate relationships with parcel data aggregators, standards development and support for access to relevant data systems. All parcel data in the United States would be accessible through this alternative, but the data would need to be captured in a single database or housed as a single agency dataset (exhibit 22).

Exhibit 22. National Parcel Data System Option





Some key properties of this approach are—

- **Data.** In a national aggregation alternative, all parcel data are available or accessible through a single access point, which may be a single dataset or system of datasets. The national parcel data system would make local, state, tribal, and federal parcel information available through web-based data services. Only existing automated data would be considered for this system. The data would be standardized by crosswalking each individual parcel data source against a standard format, similar to the approach of the current project.
- **Institutions.** The primary coordinating institution would be a federal agency. In some cases, states would be responsible for their individual areas. Federal parcel information would be contributed by either individual land management agencies or a single federal landownership database system. Coordination would involve providing guidance on standards, developing business plans for implementation, and promoting a national standard for use codes, sales information, and site addresses with all data producers.
- **Technology.** The technology behind the national aggregation approach would be database management tools. These tools would include those for automating data standardization, if possible; data quality and metadata tools; and tools for extracting, transforming, and loading data from all producers into a single dataset.

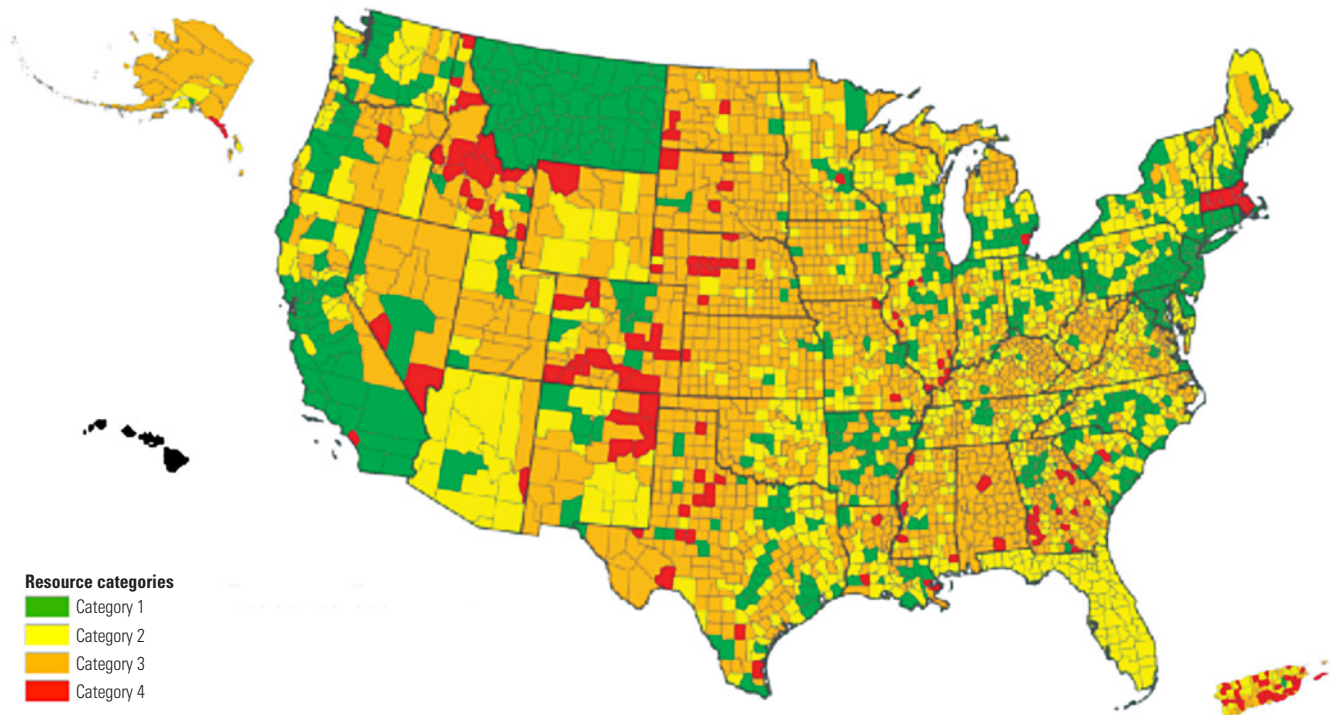
The resources required to develop this approach on a national level can be estimated based on the study of the 127 HUD-selected counties by using the same resources and level of effort for the four categories of county data described in chapter 3. For the 3,221 county-level governments in the United States and Puerto Rico, we assume the same variation in county categories based on county populations.

For example, we know that 1,229 of these jurisdictions have populations that are in the same range as that of the study’s 127 counties; at 12 hours per county for initial collection, that figure translates to a 14,748-hour level-of-effort. Exhibit 23 graphically depicts the categorization of all U.S. counties based on resource categories.

However, applying some qualitative judgment regarding the ability and willingness of each of the 3,221 jurisdictions, based on knowledge of a variety of counties across the country, provides a slightly more realistic estimate (exhibit 24). In the first year of complete national data collection, 45,653 hours would be required. By the fourth year, the annual effort would decrease to a 13,694-hour level of effort for maintenance activities, assuming current county data system capacity and hence little-to-no change in counties’ current ability and willingness to provide data.

The total cost for this 4-year effort, then, is estimated at \$21,873,200, not including subsequent maintenance costs.

Exhibit 23. U.S. County Classifications by Resource Categories



**Exhibit 24. Multiyear National Data Collection Level of Effort Estimates (Hours)**

County Category	Number of Counties	Y1	Y2	Y3	Y4
Category 1	581	2,324	1,162	581	436
Category 2	891	10,692	6,237	2,896	2,005
Category 3	1,584	28,512	20,592	10,296	7,128
Category 4	165	4,125	4,125	4,125	4,125
<b>Total</b>	<b>3,221</b>	<b>45,653</b>	<b>32,116</b>	<b>17,898</b>	<b>13,694</b>
<b>Estimated annual cost*</b>		<b>\$9,130,600</b>	<b>\$6,423,200</b>	<b>\$3,579,600</b>	<b>\$2,738,800</b>

Y = Year.

\* Assumes an average \$200-per-hour level-of-effort cost, including labor only.

Even with the advantage of repeat collection being easier than initial collection, the resources required to complete the national aggregation county by county could be excessive. Furthermore, this estimate does not include the internal agency costs for sustaining the computer systems to compile, maintain, protect, and provide access to the data. The large scale nature of this effort suggests that some additional thought needs to be put into the type and extent of its eventual use—that is, its benefits.

Very few, if any, decisions or production systems require all parcels in the nation at one time. Many applications have been identified that require or would benefit from parcel data, but none of those identified thus far require all of the parcel data at once. Most applications are either limited by geography, such as an area affected by a disaster, or by a condition, such as parcels in danger of foreclosure. With nearly 150 million records, maintained by more than 4,000 producers with an average turnover of about 15 percent annually, this configuration would have so many authorities and technological and budgetary impacts as to be essentially unattainable. This aggregation approach, therefore, is not recommended.

### 4.3.2. Need-by-Need Aggregation

In a need-by-need aggregation alternative, HUD would identify the priority counties for which parcel data would be most needed on an annual basis—an approach identical to the data collection and standardization effort of this study. HUD would contact each of these counties, and the resulting data would be aggregated, standardized, and housed in a centrally available data repository within HUD. This alternative is similar to the effort undertaken in this study but pursued over several years. Over time, a history of parcel information for the most requested data would develop, and with repeated requests the county and HUD would develop a relationship.

This alternative would enable HUD to build a relationship with the counties that are most affected by HUD programs and to deepen or expand the number of attributes and the quality of information for a limited set of counties. This data sharing between HUD and the counties is an important component of the ongoing relationship HUD has with its client counties. This

relationship can evolve into a synergistic relationship in which value-added information from analysis within HUD could assist the county and strengthen the HUD program management in the county. This alternative would enable the agency to have internal control over the data that the agency needs. One system within HUD could manage these data, and parcel data could be aggregated to any census or other geography for analysis or reporting. This limited dataset would be more manageable than a nationwide dataset and might be manageable within the current HUD infrastructure.

As is the case with the other alternatives, the parcel datasets would have metadata for each parcel record, and the dataset from which it was aggregated would be readily available so HUD users could tell at a glance the timeliness, veracity, and completeness of all parcel records. Each parcel record would link to its source data, the legal document that created and defined the record, and the authoritative institution that manages the data. This alternative is consistent with HUD’s current authorities for data collection and aggregation and most directly supports HUD’s missions. Another benefit of this approach for HUD is the ability to specify the type of data needed and, to the extent possible, their format, content, and file-submission processes.

The downside of this approach is that, in the case of disasters or unexpected immediate needs for data, the parcel data would not be readily available for counties not included in the repository. Parcel data from counties that are outside of the system would have to be specially collected or purchased. In all cases, this option can be reviewed based on the same core criteria as the other three options.

- **Data.** In a need-by-need aggregation alternative, parcel data are collected in a HUD-managed dataset. The data would be aggregated from identified counties, and historical information would be retained and available to HUD programs and projects. The data standardization issues would continue as they have in the current project—that is, each individual county’s data would be internally consistent, but the data content for some attributes, such as land use and assessment

basis, would still vary between counties. The general parcel data flow with need-by-need aggregation is identical to that shown for national aggregation in exhibit 20.

- **Institutions.** The primary managing institution would be HUD. Federal and tribal parcel information would be obtained from individual land management agencies. Public access and data-sharing agreements would be developed with each county as needed. Parcel status and contacts would need to be continuously updated and verified. Currently, the most efficient way to do so is to leverage the information contained in the national GIS inventory.<sup>22</sup> This leveraging will provide HUD and other federal agencies with current information on the status of parcel data automation and data availability. This activity would provide a means for tracking the availability and willingness for data sharing and would provide a means for tracking progress toward national parcel data availability. This tracking will save time and resources needed to identify contacts and web resources to support parcel data collection.
- **Technology.** The technology would be similar to that described previously for the state aggregation alternative, but it would be customized to meet HUD's specific needs and optimized for use by HUD programs, such as generalizing statistics into a census geography.

HUD could first identify counties that will be part of studies or analysis on an annual basis. This effort would continue to reach out to the counties defined in the current project but expand the number of counties. This step should also include a review of HUD programs to identify all agency parcel needs. In the first year, all programs that could use parcel data should be contacted, and the programs using parcel data identified and described. In future years, a followup questionnaire could be done to assemble added needs.

The resource requirements for this effort can be calculated similarly to the method identified in chapter 3 but depending on the counties that were included in the need-by-need selection and their categorization. In addition to the data collection and standardization effort, HUD could complete an analysis of their data priority needs to establish a core list of priority counties. These counties could then be checked for data comprehensiveness and availability.<sup>23</sup>

The need-by-need aggregation can be supplemented with vendor-purchased data if the data from the county are unavailable. Ultimately, this approach would solidify HUD's relationships with these most critical of counties and expand HUD's knowledge of the datasets, collection and standardization requirements, and data management requirements.

### 4.3.3. Private-Sector Collection

In a private-sector alternative, HUD would purchase all parcel data from existing private-sector (*shadow*) data aggregators but would then make it available publicly in aggregate form. Over time, it might be possible to establish a subscription for service with the vendors to provide the data needed on a scheduled and regularly updated basis. Many private-sector providers have commoditized parcel data. After the purchased data were accepted and ingested into HUD systems, one system within HUD could manage these data, and parcel data could be aggregated to any level of census geography for analysis or reporting. As with the previous alternative, this approach could be manageable within the current HUD infrastructure.

Private vendors' internal collection, standardization, publication, and delivery methods are proprietary and, therefore, unknown. However, some publicly available information sheds some light on the sources of these firms' methods for collecting and standardizing the data and the organizational infrastructure and operational resources required to offer their products. In general, larger private data vendors collect public records (such as tax rolls, property and lien files, GIS files, and criminal databases), third party proprietary data (such as mortgage transactions, mortgage securities data, and multifamily tenancy) from lenders and property owners, and third party secondary data (such as credit scores).<sup>24</sup> One of these firms employs more than 10,000 full- and part-time employees nationally to match individual public and financial records to tax rolls; clean data for specific individuals and parcels; compare data definitions to develop a common (and proprietary) taxonomy and adjust the data accordingly; and compare data on the same individual, household, or parcel. Private vendors created many of their proprietary processes for address tracking and data collection and standardization prior to any discussions of national parcel or assessor databases; the firm above began its property records business in 1959.

<sup>22</sup> Produced by the National States Geographic Information Council, the GIS Inventory is a tool for states and their partners to track data availability and the status of GIS implementation in state and local governments to aid the planning and building of spatial data infrastructures. See <http://www.gisinventory.net>.

<sup>23</sup> The NationalCad component of the GIS Inventory could be used to determine if specific counties' data are current and available, to check that the counties identified in the priority have correct contacts, and to document any limitations on access or data availability.

<sup>24</sup> For example, see <http://www.corelogic.com/about-us/data.aspx>.

For parcel data purchased from private providers, several additional factors should be considered.

- **Data content.** The private parcel data providers typically have a business need that drove their original aggregation and collection of data. This need will affect the content in terms of the number and types of attributes, which parcels are included in the dataset, and data currency. For example, data collected to support verification of land title may not have tax assessment and parcel use. Data collected to support agricultural analysis may not include urban sales.
- **Level of standardization.** Standardization processes can change the source information significantly. For example, standardization can involve changing locally provided data to increase the internal consistency of the information for a single dataset, but standardizing the data to a national specification so that the information can be compared across jurisdictional boundaries would require even more alteration to the original data. The processes used to create standard, consistent data files across jurisdictional boundaries are typically proprietary processes.
- **Currency.** The frequency of updates to vendor-sourced data will vary significantly from vendor to vendor. It may be important to understand the update cycle from the source and the frequency with which these updates are applied to the data in the vendor-sourced dataset.

In this alternative, HUD does not have a relationship with the counties and instead has a purchase arrangement with private-sector vendors. It might be more expedient, but these datasets are by their nature removed from the authoritative sources. Data use and subsequent distribution or use would be defined by purchase agreements and purchase license agreements.

As is the case with the other alternatives, the parcel datasets would have metadata for each parcel record, and the dataset from which it was aggregated would be readily available so HUD users could tell at a glance the timeliness, veracity, and completeness of all parcel records. Each parcel record would link to its source data, the legal document that created and defined the record, and the authoritative institution that manages the data.

The downsides of this approach are the lack of relationship with the client counties and that the investment in parcel data

would be a purchase that would have to be repeated each year with no subsequent savings. The lack of relationship with the counties also means that the data purchase would not support the further explorations of parcel data at the sources. As with the need-by-need alternative, in the case of disasters or unexpected immediate needs for data, the parcel data would not be readily available. Data from counties not in the original purchase would have to be purchased.

Using the review criteria, this alternative ranks in the following manner—

- **Data.** In a private-sector purchase, alternative parcel data are purchased from one or more vendors, and the content and format is a contract- or purchase-negotiated item. The data would be at least one step removed from the authoritative source, and standardization would be the result of the vendor's subjective analysis of the individual county datasets.
- **Institutions.** HUD would be the primary managing institution and would recognize only one agreement per vendor. HUD's relationship would be with the vendor and removed from the authoritative source of the data. No partnership or relationship would be established with the county client.
- **Technology.** The public-sector alternative technology requirements would be greatly reduced compared with those of other options, and they would include only the data management and distribution technology, eliminating the data standardization and collection tools. Technology requirements would include data archiving and backup processes and support, data security, and facilities for delivering the data packaged for both HUD internal network browsing and integration into HUD systems. This technology would be customized to meet HUD's specific needs and optimized for use by HUD programs by, for example, generalizing statistics into various census geographies.

The resource requirements for this option are based on past experiences by the study team with data purchases. On average, data purchases from private vendors for current data on the attributes requested in this project are approximately \$3,000 per county.<sup>25</sup> For all U.S. county-level jurisdictions, the total cost would equal approximately \$9,663,000—more than our national aggregation estimate for the first year—but with no reductions in subsequent years. This cost would be repeated each year.

<sup>25</sup> This estimate is an average based on recent purchases. Based on the volume in years and geographic coverage, and the detail of data purchased, however, the costs can and do vary widely.



#### 4.3.4. State Aggregation

The state aggregation alternative would have each state designated to serve as a trusted source that would aggregate the parcel information for the state. Ideally, these state systems would include representations of federal and tribal parcels in the state and of all state-owned and privately owned parcels. All parcel data for the state would be available in the state repository.

This alternative enables each state to account for the individual regulations and nuances of state regulations, and it recognizes the “states’ rights” authorization for property information management in the United States. Because the update cycle for real estate records is typically managed by state regulations, the timeliness of information in each state repository would be internally consistent.

From a national perspective, the state aggregation alternative could be a virtual collection, based on web services, that takes advantage of parcel aggregation from state, tribal, and federal sources. A designated federal agency would have a fully staffed organization that would manage the coordination, standards, and relationships among parcel data aggregators. All parcel data in the United States would be accessible through this alternative, but the data would not be captured through a single database or housed as a single agency dataset.

For data consumers, this alternative is ideal, providing the ability to identify which data are needed and have those data delivered to the desktop through web-based services. Metadata for each parcel, and the dataset from which it was aggregated, would be readily available so users could tell at a glance the timeliness, veracity, and completeness of all parcel records. Each parcel record would link to its source data, the legal document that created and defined the record, and the authoritative institution that manages the data.

The states could also benefit from this approach, given that many public operations need parcel information, including property value equalization, emergency response, state-owned land management, the coordination of census boundary annexation surveys, and redistricting, to name only a few. The states can also provide parcel publishing services to counties, reducing the cost burden on the individual counties and providing access to web services for smaller counties.

In this alternative, the states assume responsibility for the compilation and publication of locally compiled parcel data,

corners of common control, and state-owned land parcels and for making this information available for G2G data sharing through state-hosted services.<sup>26</sup> This alternative also supports and promotes the states to serve as the trusted data source for the state’s locally maintained parcel data. Although each state would still maintain slightly different parcel attribute and geometry standards, these variations would likely be modest and much less varied than those between all U.S. counties. This alternative has the following other advantages.

- All states have many business interests in parcel data that range across a wide variety of applications, including emergency response, fair and equitable valuation, public lands management, energy management, and environmental management, to name a few. This interest necessitates that the state has knowledge of who owns the land, surface rights, and subsurface rights, and to what use properties are being put.
- A sustainable system must be maintained through the daily business operations of local and state governments. State-hosted systems are more likely to become a part of daily decisionmaking within the states and, hence, are more sustainable.
- Many states already have or nearly have completed these types of systems. Some examples include Florida, which is described as a model in chapter 3; Montana, which has had a complete data-service-based system for many years; Indiana, which has assembled consistent parcel information statewide; and Arkansas, which includes not only access to standardized data but also a program to support automation of parcel data.

In a 2009 national review of statewide parcel aggregation, it was determined that many states have made significant progress in completing the state parcel datasets.<sup>27</sup>

- Ninety-five percent of states have identified a state cadastral contact.
- Seventy-eight percent of states have acquired or are acquiring an inventory of parcel information in their states.
- Forty-nine percent of states have parcel data in a GIS format in a state library.
- Five states have draft versions of a Parcel Data Business Plan, and thirteen states have versions in progress.

<sup>26</sup> This alternative is consistent with the Federal Geographic Data Committee Cadastral Subcommittee’s strategic vision for a national parcel system, described in appendix A.

<sup>27</sup> Stage and von Meyer (2010).



- Eighty-two percent of the parcels in the United States are in a format that can be used in a GIS. GIS parcels can readily be assumed to exist in all major metropolitan areas.

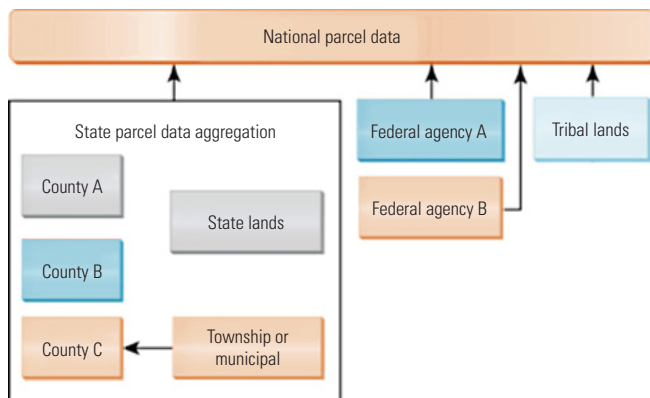
Since 2009, based on the personal communications obtained during this study, it has been observed by the authors that several states have made significant advances in statewide parcel data aggregation. For example, Massachusetts is in the process of collecting and automating its entire township parcel dataset. Wyoming and Idaho have both initiated statewide parcel programs. California, Nevada, and Utah have all made significant progress in establishing programs and relationships with counties to build ongoing state-aggregated parcel data. Arkansas has a program for standardizing and publishing data and is funding support for initial parcel mapping. In all cases, this aggregation option has the following characteristics—

- **Data.** In a state aggregation alternative, parcel data are collected into statewide datasets. Each state’s dataset would include the federal and tribal lands in the state, obtained from the appropriate federal agency, and locally generated datasets. Statewide attribution, such as parcel use codes, frequency of updates, assessment basis, and sales types, could be determined from state standards. Datasets may include more attributes to support state business processes. The datasets aggregated by the states are more likely to be consistent from county to county within the state. Exhibit 25 illustrates the general flow of parcel information with state aggregated data.

- **Institutions.** The primary managing institution would be a designated state agency. Federal parcel information would be contributed from individual land management agencies. Public-access and data-sharing agreements would be developed between the states and the local data producers, simplifying access to the local data through a single state agreement on behalf of the local producers. Local producers would remain the authoritative data source and, when more current data or additional attributes about parcels are needed, the data consumer would contact the authoritative local source
- **Technology.** The technology includes state-managed databases for attributes and spatial information; standardization technology—such as extract, transform, and load programs—for building the standardized data; data archiving and backup processes and support; data security; and facilities for delivering the data both through Internet browsing and data packaging and distribution. This technology would be repeated in each state or could be a standardized set of tools provided to each state.

Providing access to the nation’s parcel data through web services means that all of the nation’s parcel data would be available.<sup>28</sup> It would continue to be maintained by local data producers and would be standardized and aggregated by state. Parcel data would be available across state boundaries and as needed without the unsustainable aspects of a single national parcel database. In this alternative, federally managed lands could be either contributed to state data aggregation or aggregated by the federal agencies into a single or distributed dataset.

**Exhibit 25. National Parcel Data With State Aggregation Option**



Assembling needed parcel information from state holdings is significantly more resource-efficient than developing individual county datasets and, although in its infancy, is a likely trend. HUD could encourage states to build aggregated parcel datasets by making this activity grant eligible. Providing incentives such as grant funding to establish parcel aggregation programs and making these datasets available for G2G data sharing will accelerate current activities and encourage states that have not started such programs. HUD and other federal agencies could also encourage states to use the publication standards that have nearly been completed by the FGDC Cadastral Subcommittee.

We cannot currently estimate the costs of the data collection for 52 state-level jurisdictions because only 1 state was included in this study’s sample (Florida).<sup>29</sup> We can assume that it would be

<sup>28</sup> This alternative is a version of the Federal Geographic Data Committee Cadastral Subcommittee’s strategic vision for a national parcel system.

<sup>29</sup> We do know that the entire collection effort for the state of Florida required 20 person hours total. In general, assuming that all 52 state-level jurisdictions would require the same level of effort would yield 1,040 person hours for the first year’s effort. Using our \$200-per-person-hour multiplier, this effort would cost only \$208,000 compared with \$9,130,600 for a national county-by-county collection.

considerably less than collection from 3,221 counties, because of the reduced number of datasets and because state-level staff may be more consistently prepared for data requests. Additional and substantial savings would come from the standardization that has already occurred at the state level for all parcels within each state’s counties. States would require some preliminary funding to begin this transition, however. This issue is discussed further in the following section.

#### 4.3.5. Summary of Recommended Activities

The most efficient and cost-effective solution for HUD to gain access to the most current and authoritative parcel data is to continue to collect individual county data on a need-by-need basis while supporting state aggregation efforts. The state is the largest unit of consistent standards and procedures, as Florida, Montana, and other states are demonstrating. Although exceptions and special cases will always exist with real estate matters and hence parcel data, the state aggregation alternative has the best chance of developing consistency across local jurisdiction boundaries.

As a stopgap while states are developing aggregated parcel data holdings, HUD can pursue a need-by-need collection and aggregation. For example, should resources be available to continue the effort begun in this study, the initial 127 study counties would, for the most part, be moving into their second year collection. Additional counties can be included in future need-by-need collections. To encourage the states to participate in the state parcel data aggregation, HUD could consider making parcel data aggregation a Community Development Block Grant-eligible activity. As a part of this grant-eligible activity, HUD could specify the attributes, frequency of update, and data standards that the aggregation should include.<sup>30</sup> Increasing the

reliance on existing standards and encouraging state aggregations to adhere to those standards will improve the success of these systems.

The estimate of the resources required to help states aggregate parcel datasets is based on how far along the state is with current development efforts. On average, \$75,000 covers the portion of an information technology specialist’s salary at the state level to support the identified activities.<sup>31</sup> Based on this estimate, the resources required to catalyze this state-level activity are estimated at more than \$7.3 million, as detailed in exhibit 26.

HUD might trim this estimate by selecting only those states where parcel data are needed. In either case, however, HUD would also have to invest additional internal resources to develop the cross-state attribute definitions and data requirements (perhaps in coordination with the FGDC) and have to provide technical assistance to the states as they administer their grants.

Documenting the successes of the current parcel data collection effort and spotlighting the success of using state-aggregated data would also be advantageous. These findings could be presented to the National States Geographic Council (NSGIC) and offered in support of NSGIC’s Addresses for Nation and Parcels for Nation initiatives and the many states that are on the cusp of developing state parcel data holdings.

Increased coordination with the FGDC Cadastral Subcommittee could also increase HUD’s visibility in this arena and help promote efforts to share resource needs. The HUD attribute requirements could be proposed as additions to the current national core parcel data standard. This subcommittee includes representatives from states, counties, land surveyors, assessors,

**Exhibit 26. Estimates of Costs for Grants To Support State-Level Aggregation Efforts**

State Category	States	Cost	Total
States with data available and automated delivery capacity	2	\$0	\$0
States with data available but no automated delivery capacity	18	\$75,000	\$1,350,000
States with cadastral business plans but no data or capacity	18	\$150,000	\$2,700,000
States without cadastral business plans	14	\$235,000	\$3,290,000
<b>Total</b>	<b>52*</b>		<b>\$7,340,000</b>

\* Fifty states plus the District of Columbia and Puerto Rico.

<sup>30</sup> HUD would not be entering into unchartered territory in providing grants for this type of data collection encouragement. In the late 1970s through the 1980s, HUD provided grants to help develop the precursor data for the Census Bureau’s TIGER database (Sperling, 1995). A key difference between TIGER and a national parcel database, however, is that TIGER’s development coincided with the nascent growth in GIS technology, whereas automated proprietary property records were developed in the 1960s.

<sup>31</sup> For example, in Oregon, it was suggested that the state would provide \$1,000 to each county (\$36,000) and then use \$39,000 to support state website development and database management (State of Oregon, 2011).

and federal agencies. Increasing the dialogue among these participants could advance states' adoption of the standards and move toward aggregated parcel data systems.

The FGDC Cadastral Subcommittee dialogue would also increase the communication between HUD and other federal agencies that need parcel data. For example, the wildland fire community needs parcel data for all phases of wildland fire management. The U.S. Census Bureau needs parcel data to improve the vertical alignment of block and tract boundaries. Joining forces with other federal agencies with similar needs could result in cost savings for all participating agencies. Forming a working group on the FGDC Cadastral Subcommittee for federal G2G parcel needs could increase cross-agency participation.

Ultimately, as states complete their parcel data aggregation efforts, HUD could move to increase its own internal reliance on state systems, which would in turn reduce the requirements on internal data management and long-term system costs. To make a national database readily available for users, the cost of ongoing maintenance is estimated at between \$20,000 and \$50,000 per year, including the costs of backing up data, maintaining servers and services, and providing ongoing support and upgrades. Providing a national portal to link to state-hosted data services could be a shared activity among multiple federal agencies. As a longer term goal, HUD could either build or cooperate with other federal agencies to build a national access site, as well.

#### 4.4. Conclusion

This study has demonstrated that a national parcel data system is technically feasible. However, the resources required to make a core set of cross-county and cross-state data definitions and collection specifications operational would necessitate strong federal leadership, close coordination within and between states, and well-defined standards. Further work will be required to embark on this effort. Lessons learned in this study will help ensure refinements in resource projections as well as improve decisionmaking with regard to the most efficient and comprehensive ways to make this dataset a reality.

Implementing county outreach, performing data collection, and internal standardization for a large proportion of total U.S. parcels offered practical insight and highlighted some key challenges for moving forward. As stated in the findings of the National Research Council's 2007 report, *National Land Parcel Data: A Vision for the Future*, the financial and technical issues are only minor compared with the organizational and political ones. Overcoming organizational boundaries even among federal agencies has been difficult, as evidenced by the lack of a single inventory of federal lands.

With the challenge of thousands of counties or other governmental entities as potential producers of parcel data coupled with the historic fragmentation of land information and control in the United States, the organizational issues are complex. This report identified alternative data collection, standardization and update options for engaging in this national project. Whereas this study identified a state-focused effort as the most effective and sustainable long-term strategy for achieving the vision of a national parcel data system, in all cases, a coordinated and integrated national approach to parcel data will need to confront the following challenges—

- **Establishing federal leadership.** Leadership can take many forms. Clearly conveying what is needed and how it will be used should increase the data producer's understanding of the need for the data and would help define the level of effort for standardization across state and county boundaries. Committing agency resources to serve as a point of contact and creating incentives for states to aggregate and standardize data are additional steps toward demonstrated federal leadership.
- **Identifying benefits to data producers.** Benefits to the data producers for sharing parcel data are typically indirect. Clearly articulating real benefits of data sharing to the data producer improves the responsiveness to data requests. As indicated by the results of this project, counties are more willing to share their data if they can justify the workload for sharing the data and are informed of the purpose and use of the data. HUD was able to gain a relatively high level of cooperation for this data collection project, but to create a nationally consistent set of core parcel data useful for many agencies, HUD would need to develop a strong business case that includes a meaningful suite of incentives for the data producers.
- **Addressing confidentiality, cost, and collaboration.** While many data producers share their parcel data freely and provide documentation that can lead to data standardization, there are many notable exceptions. Impediments to data sharing can be based on an actual or perceived need to maintain the confidentiality of all or some of the parcel data, such as owner names or valuations; data sharing fees that can range from the cost of duplication to the more expensive data licensing fees; or a general unwillingness to share or collaborate with other agencies. Varying open records policies can further complicate confidentiality, costs, and collaboration.
- **Creating an infrastructure for data sharing.** As stated in the 2007 NRC report and further confirmed in this study, the institutional will and political hurdles to data sharing are

much greater than the technological issues. An infrastructure for data sharing includes the organizational structure, from local data producers to data aggregation and standardization conduits and then to systems for data accessibility, and the underlying technology for data sharing. The technology supporting the processing and accessing of the data continues to evolve, but if the organizational structure with its attendant policies and support mechanisms is in place, then a sustainable national parcel data system can be a reality.

Ultimately, this study has confirmed many of the NRC findings and added significant detail and implementation experience to

define achievable solutions for a national parcel data system. In particular, the study points to the needs for developing long-term leadership and institutional infrastructure for data collection. Furthermore, the study found that the challenge of standardizing data that are defined, collected, and validated differently by each data source is insurmountable as a long-term strategy without the participation and support of an institutionalized infrastructure for data sharing. In all cases, however, this study has provided additional focus to the vision of a nationally integrated land parcel dataset.

## Appendix A. History of Federal Interest in Local Parcel Data

Federal interest in local parcel data, as well as efforts to aggregate them, has a long history.

### Pre-2007 Studies

For more than three decades, the National Research Council of the National Academy of Sciences has investigated the federal role in the coordination of a national parcel database. This research began with the seminal 1980 report, *Need for a Multipurpose Cadastre*, conducted by the committee on Geodesy.<sup>32</sup> This report represented a landmark in the history of the automation of land records systems and is still considered by many to be the origin of the discussion of the benefits and implementation of automated land record systems. Specifically, the report states—

The increasing demand for land information at all levels of government and in the private sector, and the increasing involvement of federal agencies and programs in the development and maintenance of land-information systems, induced the federal agencies to request the National Research Council to define a federal role in the development of the multipurpose cadastre applicable on a national basis.<sup>33</sup>

The report also outlined the components of a multipurpose cadastre to include—

- A reference frame consisting of a geodetic network.
- A series of current, accurate, large-scale maps.
- A cadastral overlay delineating all cadastral parcels.
- A unique identifying number assigned to each parcel that is used as a common index to all land records in information systems.
- A series of land data files, each including a parcel identifier for purposes of information retrieval and linking with information in other data files.

The report's authors also reached critical conclusions regarding the contemporary state of practice that suggested a number of opportunities for future data collection, standardization, and aggregation.

- Current technology is adequate in most cases for the surveying, mapping, data collecting, and filing and dissemination of information.
- Advancement in computer applications, communication networks, and copying processes promise more efficient use of the multipurpose cadastre.
- The major obstacles in the development of a multipurpose cadastre are the organizational and institutional requirements.

A second report, *Modernization of the Public Land Survey System*, was published by the NRC in 1982.<sup>34</sup> The report focused on the specific needs for creating the multipurpose cadastre in the public-domain states, where the Public Land Survey System (PLSS) had been established, and provided a framework for mapping and describing many of the parcels. The problems identified in the report of the PLSS included lost corners and lack of good locational information. This report recommended a new federal agency to coordinate the geodetic, cadastral, and mapping activities necessary for the modernization of the PLSS that could, conceivably, be the baseline for a broader multipurpose parcel database. The report recommended the formation of an interagency working group with the “participation of all relevant federal agencies and interested groups at the state, local, and private sector levels to integrate the geodetic, cadastral, and mapping activities necessary for the modernization of the Public Land Survey System.”

A 1983 panel, resulting in a third publication titled *Procedures and Standards for a Multipurpose Cadastre*, built off of the 1980 report by addressing questions of how the public sector, especially local government, could carry out the earlier recommendations.<sup>35</sup> The report also laid out the technical

<sup>32</sup> NRC (1980).

<sup>33</sup> NRC (1980: v).

<sup>34</sup> NRC (1982).

<sup>35</sup> NRC (1983).



specifications for the components of a multipurpose cadastre that would be a key informational component to how government should fulfill its mission. The report lists 25 functional areas of government that could benefit from a complete inventory of all currently existing parcels and their legal identifiers. Property tax assessment, deed recordation, and planning were developed in the most detail. The report elaborated on the role of county land offices to include compiling standards for attributes, accuracy, update frequency, and completeness of records. It emphasized the need for the standards to be developed in cooperation with other jurisdictions.

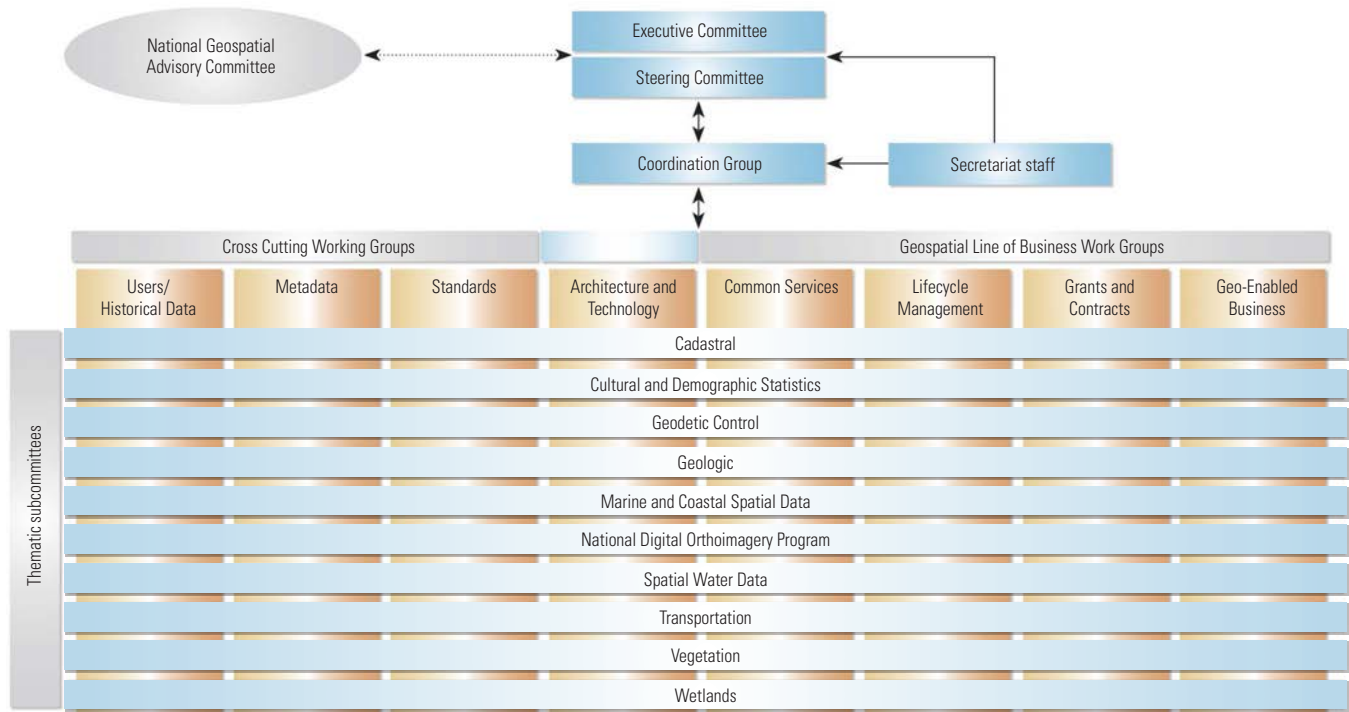
The panel also addressed the funding issue by calling for a program of federal grants to counties (or their equivalents) to cover about 40 percent of the cost for the multipurpose cadastre. They estimated that the cost of a matching federal program would be \$90 million per year over a 20-year period, for a total federal contribution of \$1.8 billion. In 2003 dollars, these figures represented \$165 million per year and \$3.3 billion total, respectively. The combined federal and local investment would be \$8.2 billion.<sup>36</sup>

Subsequently, in 2002, the charge and structure of the Federal Geographic Data Committee were refined by Office of

Management and Budget (OMB) Circular A-16 regarding the Coordination of Geographic Information and Related Spatial Data Activities (discussed in appendix B). Originally released in 1990, the revision provided direction for federal agencies to coordinate approaches for the electronic development of the National Spatial Data Infrastructure (NSDI) and the organization of the FGDC. This revision included naming the deputy director for management of the OMB as vice-chair. The FGDC’s structure is visually depicted in exhibit A-1.

As a multitiered entity involving numerous federal participants that focuses on thematic subcommittees, the FGDC has a very active Cadastral Subcommittee. This group has made significant progress in the development of standards and coordination with stakeholders. Both OMB Circular A-16 and the FGDC define cadastral data as the “geographic extent of past, current, and future right, title, and interest in real property, and the framework to support the description of that geographic extent. The geographic extent includes survey and description frameworks such as the Public Land Survey System, as well as parcel-by-parcel surveys and descriptions.” Cadastral data

Exhibit A-1. FGDC Organization



<sup>36</sup> This estimate is in line with recent estimates—including those of this study—of \$7 to \$8 billion for the creation of a nationwide multipurpose cadastre.

are considered a framework layer along with the other five framework themes (ortho imagery, elevation, hydrography, governmental unit, and geodetic control). Stewardship of the cadastral layer is delegated to the Department of Interior's (DOI's) Bureau of Land Management (BLM).

In 2003, the U.S. Department of Housing and Urban Development sponsored an important NRC study, *GIS for Housing and Urban Development*, which addresses many of the issues relating directly to the role of HUD in the coordination of parcel data.<sup>37</sup> The panel provided an in-depth analysis of how parcel-level data could be used to meet the mandates of HUD's various programs. The report stated that "HUD has a unique ability to introduce local priorities into national dialogs and to provide support and encouragement so that local data meet national standards for inclusion in a National Spatial Data Infrastructure (NSDI). The development of a parcel-level layer for metropolitan areas is particularly important to HUD, to the communities HUD serves, and to national initiatives, including the NSDI and other federal data initiatives."<sup>38</sup>

The panel assembled a table of HUD's programmatic interests and how a parcel-based GIS could address those issues. The report's analysis describes many ways that HUD could apply GIS to analyze and interpret the results of projects and studies. Much of the GIS information referenced in that table would be enhanced with the addition of locally generated parcel data. The panel also introduced the concept of an Urban Spatial Data Infrastructure (USDI) that would serve as a complement to the NSDI. In order to make the USDI a functional program, the panel recommended that HUD take a lead role and offered the following recommendations for developing mechanisms to accept and integrate relevant, locally derived data and to georeference the data for integration in the agencywide GIS.

- HUD should spatially enable local data by performing address matching of individual records at the finest scale using geographic coordinates.
- HUD should select, tabulate, analyze, and map relevant housing variables through a GIS at multiple relevant geographic scales (census block, block group, and tract; place, county, and metropolitan area).
- The Office of Policy Development and Research (PD&R) should take the lead within HUD in efforts to integrate

grantee and other data at the parcel, neighborhood, municipality, school and school district, metropolitan area, state, and national levels.

## 2007 NRC Report

The NRC Mapping Science Committee conducted a major update of the original 1980 multipurpose cadastre report starting in 2006 and culminating in the 2007 report, *National Land Parcel Data: a Vision for the Future*.<sup>39</sup> The work was sponsored by the BLM, the U.S. Census Bureau, the FGDC, the Department of Homeland Security (DHS), and ESRI (a private GIS company). The report generally represents the most comprehensive analysis of the current state of parcel data and has served as the focal point for major discussion on the topic.

Since its publication, the report's findings and recommendations have been discussed extensively. In many ways, the report confirms many of the findings from the one published 27 years earlier. As the executive summary states—

The committee concluded that a national land parcel dataset is necessary, timely, technically feasible, and affordable. Although the benefits and needs for nationally consistent parcel data are much more clear and urgent than in 1980, there has been little progress toward the recommendations of the 1980 report. While a great deal of parcel data has been digitized at the local level, 30 percent of individual parcels still need to be converted, and there has been little progress toward an integrated national set. Many of the technical barriers have been overcome, so the remaining challenges are primarily organizational.<sup>40</sup>

The committee found that a significant digital parcel data divide exists between various counties. In many parts of the United States, parcel data only exist as lines on paper maps stored in a local courthouse. While about 70 percent of the tax parcels in the United States now exist in digital form, the remaining 30 percent are in the roughly 2,000 most rural counties. Although these counties have fewer total parcels, they also do not have adequate financial resources to convert their data to digital form. On the other side of the divide, many urban areas are covered by two or three versions of parcel data, and often anyone with a simple web browser can anonymously retrieve

<sup>37</sup> NRC (2003).

<sup>38</sup> NRC (2003: 4).

<sup>39</sup> NRC (2007a).

<sup>40</sup> NRC (2007a).

information about the ownership, taxes, and value of any parcel by owner name or street address. Many communities routinely align parcel boundaries using digital aerial photographs that precisely display fences, driveways, sidewalks, hedges, and other features that align with property boundaries. In fact, there are parcel data programs that reflect real-time changes in real estate transactions or new street addresses through field-based, global positioning system-enabled, hand-held computers.

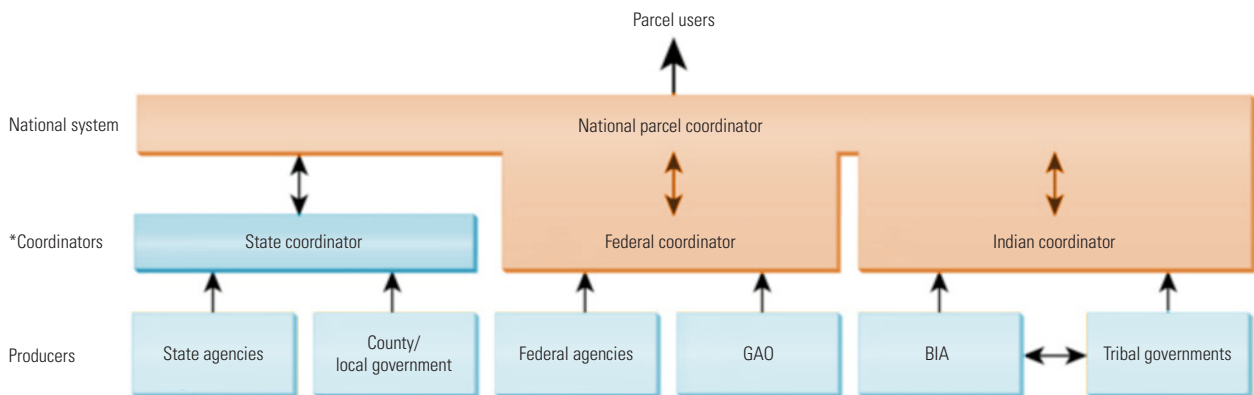
The committee’s vision for the national land parcel dataset is of a distributed system of land parcel data that are housed with the appropriate data stewards but accessible through a central, web-based interface. It would have a minimum set of attributes, and development and integration of the national dataset would be overseen by a national coordinator working with coordinators for federal lands, tribal lands, and each state. This dataset would serve as the cadastral data layer of the NSDI.

The envisioned system would link a series of distributed servers maintained by local and state governments, as depicted in exhibit A-2. In a virtual environment, the system could seamlessly assemble accurate and timely parcel information for any part of the nation—analogue to *just-in-time manufacturing*, in which parts needed for assembly are obtained when they are needed rather than stored in a large warehouse filled with inventory. Each parcel would need to be treated as a unique entity, and

local government officials would maintain the information about each parcel. Local parcel data stewards would share only geographic coordinates that define the geometry of each parcel and a minimal set of attributes including street address, unique parcel identification number (PIN), a generalized category of ownership, and metadata. For federal lands, a federal land parcel coordinator would improve the way the U.S. government manages parcel data for the lands that are federally managed.

Since most parcel data is developed by entities other than the U.S. government, the development of a nationally consistent set of data is complex. Such a comprehensive approach would require the establishment of a national land parcel coordinator (separate from the federal coordinator overseeing federally administered lands) who would be the initial point of contact for all levels of parcel data. This coordinator would build relationships with state and local governments to establish an unfettered access to a comprehensive set of parcel data linked to a unique identification system. This level of intergovernmental coordination is proposed by other current initiatives at both the state and national levels, such as the National States Geographic Information Council (NSGIC) Fifty States Initiative to improve statewide coordination of geospatial information technologies, OMB’s Geospatial Line of Business, and the current Data.gov and Geospatial Platform initiatives to improve geospatial data coordination across the federal government.<sup>41</sup>

**Exhibit A-2. Proposed New Coordination for National Parcel Dataset**



BIA = Bureau of Indian Affairs. GAO = Government Accountability Office.

\* Coordinators provide a range of guidelines and resources to all primary custodians. They may also need to do the basic production work for smaller and more isolated custodians.

Source: Adapted from NRC (2007)

<sup>41</sup> The report also recommended reviewing—

- Whether the Bureau of Land Management should be the lead federal agency.
- How Congress and the Census Bureau should explore modifying Title 13 so that building addresses and coordinates can be made public.
- If the national coordinator could administer an intergovernmental funding program for developing and maintaining parcel data, including incentives to participate for those counties with fully developed systems and financial support for those who do not.
- The requirements placed on local governments for putting parcel geometry and a very limited set of attributes into the public domain, possibly as a minimum requirement for receiving federal funds directly associated with property, such as disaster relief.

## Post-2007 Efforts

Since the 2007 update, a variety of studies and preliminary efforts have begun, most notably with regard to disaster management. In fact, a separate 2007 NRC study was prepared for DHS that focused on the importance of geospatial data when responding to disasters.<sup>42</sup> Although this report does not suggest that DHS assume the lead role in the maintenance of parcel data, it emphasizes the importance of parcel data when dealing with the use, value, and ownership of property. In particular, the report noted the need to tap into the “data on the ownership of land parcels, or cadastral data, [which] provide a particular and in some ways extreme example of the problems that currently pervade the use of geospatial data in emergency management.” The report also noted that, “as with many other data types, it is not so much the existence of data that is the problem, as it is the issues associated with rapid access.”<sup>43</sup>

### Creation of the National Geospatial Advisory Committee

The National Geospatial Advisory Committee (NGAC) was formed in early 2008 to provide advice and recommendations to the FGDC on the management of federal geospatial programs. The Committee is a federal advisory committee sponsored by the DOI and operated under the Federal Advisory Committee Act. The NGAC, which reports to the chair of the FGDC, has been charged with the mission “to provide strategies regarding the creation, management and dissemination of cohesive geospatial data, information and knowledge to enable commercial, academic, and nonprofit organizations and all levels of government.”

### Mortgage Crisis Stakeholders Meeting

More recently, a 1-day meeting was held in Washington, D.C., in May 2009, cosponsored by the FGDC Subcommittee on Cadastral Data and the International Association of Assessing Officers. The meeting focused on the use of parcel data to address the research issues related to the mortgage crisis. The meeting attracted more than 50 attendees, including several members of the HUD staff and representatives from the Federal Reserve. At the time, NGAC was still evaluating the NRC recommendations and endorsed the request for a FGDC-sponsored mortgage and economic recovery working group meeting to

address the use of land parcel data for monitoring, evaluation, and management of financial and mortgage issues. During the session, Michael Howell, deputy administrator of the OMB and cochair of the FGDC Steering Committee, noted—

We need to be open to innovative and creative ways to address this complex set of problems and take advantage of new tools and capabilities to develop effective responses. Parcel data is an excellent case in point. I think you will see from some of the examples today the powerful capabilities that land parcel data can provide when combined with other datasets and analytical tools and technology.<sup>44</sup>

Some key recommendations from the discussion include—

- **Add the local parcel ID (PIN) to the Home Mortgage Disclosure Act (HMDA) data.** The PIN provides the unique key for linking property to a wide range of attribute information. It is also a critical part of all legal transactions. It would not be a burden for a lender to include the PIN when submitting an HMDA report because it is included in the documentation used to process a mortgage. The PIN provides the link to the local government property assessor’s data, enabling the mortgage to be accurately located for mapping and analysis of conditions in any neighborhood. This PIN also provides an excellent way to track first time homebuyers and individuals taking advantage of the new Hope for Homeowners and Neighborhood Stabilization Programs. The addition of the PIN would place the transactions into a common, location-based information platform that could be accessed through a set of web services. After this platform has been created, refinements can be made to use records from public agencies or from the private sector.
- **Develop a parcel early warning system.** Much as the Centers for Disease Control and Prevention tracks early warning indicators for public health, a national system is needed that can use early warning indicators of financially distressed housing and mortgage markets. The stakeholders meeting identified a range of indicators that need to be further refined to identify the most appropriate indicators that could be mapped monthly at the census tract spatial resolution to provide a dashboard, or quick overview, of national trends. By using a system like the one deployed

<sup>42</sup> NRC (2007b).

<sup>43</sup> Stage, von Meyer, and Ader (2005) echoed this sentiment, arguing that the cadastral data that can provide the most accurate information in support of emergency management is difficult to access because of a number of factors, including data distribution agreements, data purchase requirements, and widely disparate data formats.

<sup>44</sup> FGDC Subcommittee on Cadastral Data (2009).



by the New York Times or other commercial services (such as Zillow, RealtyTrac, First American Spatial Solutions, and so on), analysts could establish rules that would trigger reports on exceptional or outlier conditions to identify those census tracts that are trending toward problems, so they could be targeted for the collection of additional detail. The system could track census tract geography with monthly data updates displayed on a dashboard map that can be scrolled through time with alerts for certain indicators (for example, percentage delinquent mortgage or averaged credit score trends by tract).

- **Complete the standardization and availability of parcel data nationwide.** Although nearly 82 percent of parcels are available in a digital GIS format, these existing data need to be standardized, and the publication and availability of core data need to be implemented in jurisdictions that currently have limited capabilities or limited distribution policies. More than 1,000 counties do not have digital parcel data at all, however. These remaining jurisdictions, which are generally found in persistently impoverished rural areas, need to have their parcels converted to a digital format. These three recommendations are achievable with current technology. They would provide a significant improvement in the available information to address the current mortgage crisis and could be achieved quickly.

The FGDC's 2009 annual report highlighted the excellent work of the FGDC Cadastral Subcommittee in conducting the research on the mortgage crisis, providing updates to the existing national inventory of cadastral data, providing updated parcel data to the wildland fire community, and working to establish sustainable standardized parcel datasets at the state level. Working with the BLM, the subcommittee also participated in the development of a standardized PLSS, FGDC cadastral data content standards, and publication guidelines. It also appeared that the FGDC was going to take a serious look how it might actively engage in the development of a national parcel database.

In 2010, the FGDC Subcommittee on Cultural and Demographic Data published *United States Thoroughfare, Landmark, and Postal Address Data Standard with Maintenance* to further provide one standard for meeting the diverse address data management requirements for local address administration, postal and package delivery, emergency response (and navigation), administrative record keeping, and address data aggregation.

The FGDC also surveyed agency staff in 2011 regarding their use of address data files and the need for a consistent geocoding service that would involve point-level addresses rather than address interpolation as performed with the Topologically Integrated Geographic Encoding and Referencing (TIGER) database.<sup>45</sup>

Finally, responding to a request for feedback from NGAC regarding the recommendations that it had made to the FGDC during its first 3 years of operation, the FGDC noted in the summer of 2011 that clear challenges in federal agency authority continue to prohibit the development of a robust national parcel database. In particular, the DOI had conducted a review of legal authorities to conduct land parcel coordination activities based on the 2007 NRC report and the assignment of lead responsibility to the BLM via OMB Circular A-16. However, the DOI did not have statutory authority or funding to provide national parcel coordination as described in the NRC report. The BLM has continued to provide active leadership of the FGDC Cadastral Subcommittee. To date, then, no DOI office or division has the statutory authority or funding to develop a nationwide parcel database despite recommendations otherwise.

### Current Efforts

FGDC member agencies, however, continue to pursue individual initiatives. For example, in June 2011, HUD presented to the NGAC regarding that agency's unique role in parcel information gathering during the post-Hurricane Katrina disaster recovery in the Gulf of Mexico and related challenges and incentives for developing comprehensive databases (including data quality, property situs addresses, address matching, and enhanced capabilities to link federal, state, and local data sources).

Note also that several federal agencies have been licensing the data from commercial firms. For example, during the public comment period at the June 2011 NGAC meeting, representatives from CoreLogic, Inc., reported that HUD, DHS, the Small Business Administration, the General Services Administration, the Federal Deposit Insurance Corporation, and the Federal Bureau of Investigation were currently licensing their spatial and nonspatial parcel data. These commercial purchases occur despite a 2004 Government Accountability Office recommendation that the Director of the OMB should strengthen the agency's oversight actions to more effectively coordinate federal geospatial data and systems acquisitions and thereby reduce potentially redundant investments.

<sup>45</sup> [http://www.nsgic.org/public\\_resources/2011AC\\_FGDC-Address-Questionnaire.pdf](http://www.nsgic.org/public_resources/2011AC_FGDC-Address-Questionnaire.pdf).



At the same time, interest has increased in the legislative branch. The Congressional Research Service issued its first report, “Issues Regarding a National Land Parcel Database,” in 2009.<sup>46</sup> An update of that report was published in May 2011. These reports, which borrow heavily from the NRC and NGAC materials, provide an extremely useful review of the issues surrounding the topic. They also summarize the core administrative challenge noted by the FGDC.

Although BLM is the steward of federal land parcel data and coordinator for cadastral data under the FGDC, a 2007 National Research Council (NRC) report found that a coordinated approach to federally managed parcel data did not exist. Legislation that would address some of the issues for creating a national cadastre (H.R. 1620, the Federal Land Asset Inventory Reform Act of 2011) was introduced in the 112th Congress. Similar legislation was introduced in the 111th Congress but was not enacted.

On April 15, 2011, Representative Kind introduced the Federal Land Asset Inventory Reform Act of 2011 (H.R. 1620) to require the Secretary of the Interior to develop a multipurpose cadastre of federal *real property*. The legislation defines *cadastre* as an inventory of real property and defines *federal real property* as land, buildings, crops, forests, or other resources still attached to or within the land; improvements or fixtures permanently attached to the land; or structures on it. The bill would require the Secretary to coordinate with the FGDC pursuant to OMB Circular A-16, integrate the activities under the legislation with similar cadastral activities of state and local governments, and participate in establishing standards and protocols that are necessary to ensure interoperability of the geospatial information of the cadastre for all users. Similar legislation was introduced in the House of Representatives in the 111th Congress and in previous Congresses. H.R. 1620 was referred to the House Natural Resources Committee, Subcommittee on Energy and Mineral Resources, but it has not yet been acted on by the subcommittee.

Similarly, some debate (although little legislative action) has occurred regarding e-government interactions. Section 216 of P.L. 107–347, the E-Government Act of 2002, called for facilitating the development of common protocols for geographic information to promote collaboration and use of standards and to reduce redundancy among federal agencies. Authorization for appropriations under the act expired at the end of fiscal year 2007. If the E-Government Act were amended, Section 216 could be expanded to include language for a national cadastre, as proposed in H.R. 1620, for designating Executive Office of

the President leadership for all federal geospatial activities, as recommended by NGAC, or for amending Title 13 of the U.S. Code to enable broader sharing of address data for its inclusion in a national land parcel database.

Some additional stakeholders have argued that the OMB could take a stronger role in the FGDC through more active enforcement. Thus, an administrative option for creating a national land parcel database, at least for the federal lands, would be to enforce OMB Circular A-16 more rigorously. This action would likely mean that the OMB would take a true oversight and coordination role and enforce compliance with the circular through its power to affect the budgets of the participating departments and agencies. The NGAC recommended this action, and further recommended a Geographic Information Officer within each department or agency with responsibility under the FGDC. NGAC also recommended a geospatial leadership and coordination function in the Executive Office of the President, which would elevate the profile of the geospatial enterprise within the administration and presumably signal a higher priority for coordinating geospatial activities in the federal government.

The supplemental guidance does not address the cadastral theme directly (nor does it address other themes directly), but it sets forth its goal of a portfolio-centric model that “cures the single agency, stovepipe model by applying consistent policy, improved organization, better governance, and understanding of the public to deliver outstanding results.” Moreover, the supplemental guidance recognizes that federal investments in geospatial data, which would include land parcel data, “were largely uncoordinated and often lacked transparency, and sometimes resulted in data deficiencies, lack of standardization, inefficient use of resources, lack of interoperability, or inability to share data.” Whether and how the new guidance will affect how land parcel data are acquired, managed, and coordinated at a national level is not yet known.

## Summary

This review indicates that, at least since the publication of the 1980 NRC report, considerable interest has focused on the role of the federal government in the coordination of parcel data. Interest was further piqued following the publication of the 2007 NRC report. The current crisis in the mortgage markets and congressional actions involving remedies has further intensified this interest. Even though the FGDC was enthusiastic about supporting a federal initiative to coordinate parcel data

<sup>46</sup> Folger (2009).

in 2009, the official policy remains that the DOI did not have statutory authority or funding to provide national parcel data collection as recommended in the NRC report. Furthermore, under OMB Circular A-16, the BLM remains the steward for the cadastre layer. This layer is interpreted to be federal land ownership and the cadastral reference data. The FGDC Cadastral Subcommittee continues to operate as the focal point for coordinating national parcel and cadastral framework activities and federally managed lands' cadastral information.

Administrative options have also been proposed to achieve the vision for a land parcel database described in the 2007 NRC report, including a distributed system of land parcel data housed with the appropriate data stewards but accessible through a web-based interface. Some recommend that the OMB and the DOI take a stronger hand in enforcing the requirements of OMB Circular A-16 and Executive Order 12906, which created the FGDC and instigated efforts to create the NSDI. NGAC also recommended a geographic information officer within each federal department or agency and a geospatial leadership and coordination function in the Executive Office of the President. The Obama Administration issued supplementary guidance to OMB Circular A-16 on November 10, 2010, which could address some of the same issues raised in the NRC report, particularly regarding data sharing, coordination, and funding.

The federal government's efforts to coordinate its geospatial activities, through the FGDC and the development of the NSDI, include a strong emphasis on land parcel data. Coordinating all land parcel data, the bulk of which is produced for local and regional needs, remains even more of a challenge. The FGDC Cadastral Subcommittee has made significant progress in establishing standards and coordinating cadastral data, according to the NRC. Yet, tremendous implementation challenges remain. As the NRC observed, despite nearly 20 years of effort at coordinating geospatial information and land parcel data—

... one could conclude that the United States has a comprehensive approach to parcel data. However, a detailed analysis of the situation suggests the opposite. ... It is difficult to ascertain the status of parcel data within the various federal agencies, and it appears that none of the federal land management agencies have a comprehensive and complete parcel dataset for the lands they manage. ... There is also evidence that many federal agencies that do not manage lands are acknowledging that they need parcel data to fulfill their missions and, in the absence of a national means to access the data nationwide, are creating datasets to meet their particular needs, often without coordination with other federal agencies that may have needs for the same or similar data.

## Appendix B. HUD Authority, Data Needs, and Activity

Over the past decade, the U.S. Department of Housing and Urban Development's role in coordination of parcel data has been much discussed. The National Research Council, in its report, *GIS for Housing and Urban Development*, offered the most comprehensive and strongest recommendation. Under that plan, it was proposed that HUD would operate a parcel-based Urban Spatial Data Infrastructure, although such a system has not been implemented to date.

This exploratory data collection study, the County Data Records Project, acknowledges HUD's interest in a national collection of locally generated parcel data. In fact, the original request for proposal in 2009 acknowledged the fact that the United States does not currently maintain a national parcel database. Secondly, it also specifically acknowledged that parcel-level data are important for monitoring the mortgage crisis. Third, it expressed its intent to use locally maintained parcel data to support evaluations and monitoring of its programs, such as the Neighborhood Stabilization Program. Fourth, HUD also acknowledged that it has an interest in becoming the long-term steward of housing information based on parcel-level data. For these reasons, a review of HUD's authority and activities is provided.

### HUD Authority To Collect Parcel Data

Defining and managing the government's authority to act is a central premise for governance in the United States; federal agencies must have a statutory authority to act and conduct regulatory and supporting activities. Federal authority, as opposed to budgetary authorization, is generally the power of a federal agency or its administrators to carry out the terms of the law creating the agency and to administer and implement regulations, laws, and government policies. Three primary authorities were found related to HUD's ability to request, collect, and maintain parcel information. These authorities are Office of Management and Budget Circular A-16, Title 12 Banks and Banking National Housing Act, and the Dodd-Frank Wall Street Reform and Consumer Protection Act. The following citations are from those authorizations, with emphasis added.

### OMB Circular A-16

This circular provides direction for federal agencies that produce, maintain, or use spatial data either directly or indirectly in the fulfillment of their mission. This circular establishes a coordinated approach to electronically develop the National Spatial Data Infrastructure and establishes the Federal Geographic Data Committee.

The circular was revised in 2002 from the 1990 version to reflect changes in technology, further describe the components of the NSDI, and assign agency roles and responsibilities for development of the NSDI. The revised circular names the deputy director for management of the OMB as vice-chair of the FGDC.<sup>47</sup>

This circular defines 11 federal agency responsibilities related to using geospatial data, and it defines four roles for theme lead agencies. HUD is currently the theme lead for housing, and as such one of its responsibilities is to provide leadership and facilitate the development and implementation of a plan for nationwide population of each data theme. Plans will include the development of partnership programs with states, tribes, academia, the private sector, other federal agencies, and localities that meet the needs of users; address human and financial resource needs; identify needs for standards, metadata, and the clearinghouse; and advance a timetable for the development of NSDI data themes.

Under the current OMB Circular A-16, released by the OMB in December of 2010, HUD is listed as the steward for housing data. The circular specifically states—

Housing: Department of Housing and Urban Development (HUD) HUD's database maintains geographic data on homeownership rates, including many attributes such as HUD revitalization zones, location of various forms of housing assistance, first-time homebuyers, underserved areas, and race. Data standards have not yet been formalized.

Unlike cadastral data, however, housing is not considered to be a framework layer, and little effort has been made to develop a functioning working group or data content standard, although HUD staff does participate on the FGDC Cadastral Subcommittee.

<sup>47</sup> From OMB web pages available at [http://www.whitehouse.gov/omb/circulars\\_a016\\_rev](http://www.whitehouse.gov/omb/circulars_a016_rev).

In the 2010 Circular A-16 supplement guidance, the OMB emphasized the need for portfolio management of National Geospatial Data Assets (NGDAs), which includes identifying existing geospatial datasets that may be considered NGDAs, determining if they should become part of the NGDA Portfolio, and managing them. This portfolio management approach could change the way data stewardship is delegated and potentially funded, including pooling data to an alternative authority.

It is also significant that at the October 2011 FGDC steering committee meeting, the executive director of the FGDC introduced the outline for a reorganization and consolidation of the A-16 data themes (exhibit B-1). This proposed reorganization could dramatically change the way the federal government approaches parcel and housing data. By consolidating the current 34 data themes to these 17, potential exists for considerable confusion. The housing data theme, which is of most interest to HUD, would be eliminated, and it is not clear where housing would fall in the new taxonomy. For example, land use and land cover have traditionally (that is, U.S. Geological Survey Anderson Categories) meant broad land cover categories at the highest level (that is, Level I as “urban”) and specific land use for individual structures (Level IV as “residential or commercial”).

Conceptually, if the NSDI had a theme that consisted of every structure in the United States, the structure’s address and land use could be attributes of each feature. Note that the U.S. Census Bureau’s Master Address File Structure Points are the closest approximation to such a comprehensive geospatial database; those data, however, are protected under Title 13 from release outside the bureau. Within the new taxonomy, the category “real property” is federally restricted to the General

Services Administration’s Real Property Inventory of federal assets (buildings and property) and directly overlaps with the cadastre theme. Consequently, no category would logically include housing.

**Title 12 Banks and Banking National Housing Act**

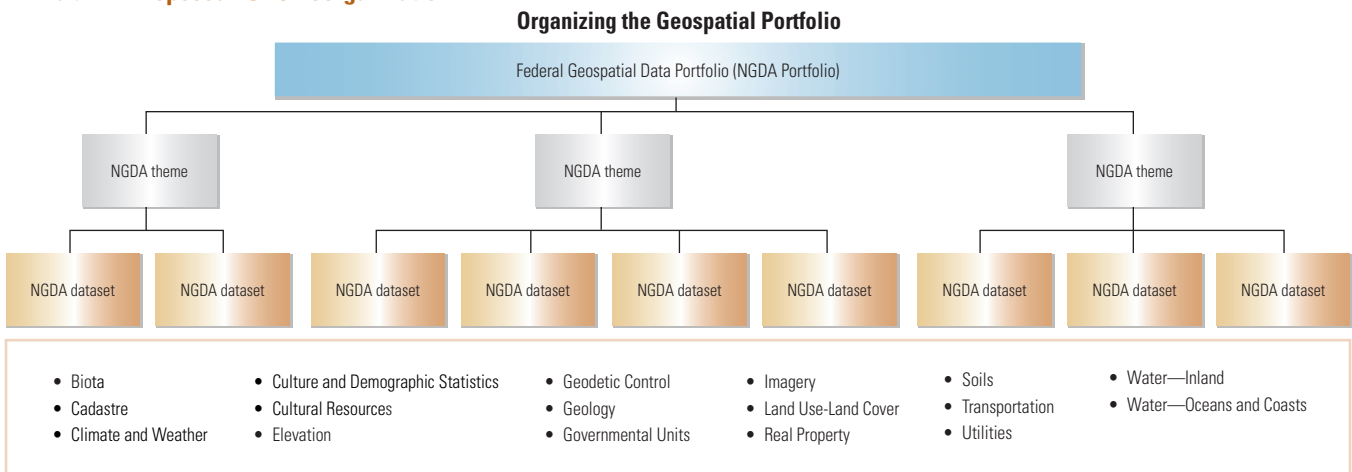
Under this act (12 USC 1701), Section 1701z-1 states that “The Secretary of Housing and Urban Development is authorized and directed to undertake such programs of research, studies, testing, and demonstration relating to the mission and programs of the Department as he determines to be necessary and appropriate.” This language might provide sufficient space to include the collection of parcel data for research and evaluation purposes.

**Dodd-Frank Wall Street Reform and Consumer Protection Act**

The Dodd-Frank financial reform legislation, Section 1446—Study of Defaults and Foreclosures—states the following—

The Secretary of Housing and Urban Development shall conduct an extensive study of the root causes of default and foreclosure of home loans, using as much empirical data as are available. The study shall also examine the role of escrow accounts in helping prime and nonprime borrowers to avoid defaults and foreclosures, and the role of computer registries of mortgages, including those used for trading mortgage loans. Not later than 12 months after the date of the enactment of this Act, the Secretary shall submit to the Congress a preliminary report regarding the study. Not later than 24 months after such date of enactment, the Secretary shall submit a final report

**Exhibit B-1. Proposed FGDC Reorganization**



FGDC = Federal Geographic Data Committee. NGDA = National Geospatial Data Assets.  
Source: FGDC (2011)

regarding the results of the study, which shall include any recommended legislation relating to the study, and recommendations for best practices and for a process to identify populations that need counseling the most.

Section 1447—Default and Foreclosure Database—further stipulates that—

The Secretary of Housing and Urban Development and the Director of the Bureau, in consultation with the Federal agencies responsible for regulation of banking and financial institutions involved in residential mortgage lending and servicing, shall establish and maintain a database of information on foreclosures and defaults on mortgage loans for one- to four-unit residential properties and shall make such information publicly available, subject to subsection (e).

Finally, Section 1094—Related to the Home Mortgage Disclosure Act Amendments—made substantial changes to the content and format of HMDA reports that include property parcel numbers. The interpretation of the phrase “any other information the agency may require” might hinge on the mission of the divisions within HUD and how these divisions could use parcel and housing information from local governments.

## HUD Data Needs

Parcel data have many potential uses within HUD. For example, parcel data could assist the following HUD entities.

- **Real Estate Assessment Center (REAC).** REAC’s mission is to provide and promote the effective use of accurate, timely, and reliable information assessing the condition of HUD’s portfolio; to provide information to help ensure safe, decent, and affordable housing; and to restore the public trust by identifying fraud, abuse, and waste of HUD resources. REAC’s product is accurate, credible, and reliable information assessing the condition of HUD’s housing portfolio. To deliver a quality product, REAC depends on the successful partnership of people and technology. At the heart of this technology is an Internet database of comprehensive and objective information drawn from existing government systems and from an ongoing program of property inspections, analysis of financial and management reports, and resident surveys.
- **Office of Policy Development and Research.** PD&R was established in 1973 as an office headed by an assistant secretary. The statutory authority for PD&R’s research

activities is found in Title V of the 1970 Housing Act, which authorizes programs of “research, studies, testing, and demonstrations relating to the missions and programs of the Department.” Research priorities have differed from administration to administration, with varying mixes of housing studies, housing technology research, demonstrations, HUD program evaluations, and policy reports. PD&R’s research activities are designed to have immediate relevance to the policy issues facing the secretary and his principal staff.

PD&R provides a valuable service to researchers and the public by expanding the availability of statistics on housing and urban development (<http://www.huduser.org>). In addition to the American Housing Survey and State of the Cities data systems, PD&R makes available (1) unique data on the Low-Income Housing Tax Credit Program; (2) annual publications of fair market rents and estimates of Area Median Family Income, which are used as a standards for determining rent and subsidy levels in HUD and other federal and state housing programs; and (3) a variety of other publications on the characteristics of families assisted under HUD programs.

- **Housing Scorecard.** Each month, HUD and the U.S. Department of the Treasury produce a monthly scorecard on the health of the nation’s housing market. The scorecard incorporates key housing market indicators and highlights the effect of the administration’s unprecedented housing recovery efforts, including assistance to homeowners through the Federal Housing Administration and the Home Affordable Modification Program.

## HUD Parcel Data Activity

In addition to representation on the FGDC, HUD has taken on geocoding services that could become part of a broader geospatial platform. HUD’s enterprisewide Geocode Service Center (GSC) provides a single source of geocoding services across the department—reducing costs, increasing quality, and providing opportunities to create new products and services. The GSC has developed a standard set of address data quality and geocoding services that provide a consistent, high-quality method for enhancing address information and assigning geospatial codes and coordinate information. HUD’s geocoding infrastructure has also supported a number of data linking and matching capabilities, including the use of local parcel data. Ongoing geoenabling of HUD’s data will improve program performance and support policy and decisionmaking.



## Appendix C. Key Terms and Concepts

Despite significant interest at the national level, major obstacles exist to the development of a national parcel database, not the least of which is counties' submission of data in a uniform format. A common understanding of terms and concepts is critical to overcoming this challenge. Discussions about parcel data are difficult, because the everyday use of the terms describing property ownership can be confounded with technical terms in legal practices and the technology that is used to manage these large datasets. Although professional agreement on their definitions exists across counties, the implementation of these definitions varies wildly.

The term *parcel data*, in the context of this report, refers to two components, (1) the attributes about a parcel (ownership, value, and use) and (2) the geometry of the parcels that is created by automated mapping, either through Computer Aided Drafting or GIS. The attributes for parcels are typically managed by an assessor and are stored in a proprietary complex database that can contain hundreds of fields to describe parcel value, conditions, and characteristics. These systems are often referred to as *CAMA* (Computer Aided Mass Appraisal) files. The CAD or GIS component typically includes relatively accurate (although not survey accurate) geometric renderings of parcels that are linked to the *CAMA* files. It is common and accepted practices to have a distinct GIS department maintain responsibility for these data. When combined, these databases are a very powerful tool for conducting analysis on the parcels.

*Authoritative data* come directly from the creator or authoritative source. They are the most current and accurate data and are often vetted according to official rules and policy. The data have a known accuracy and lineage, and in some cases they are certified. For some, the authoritative data are referred to as the *primary* data source. Often, an authoritative source has statutory authority or requirements for collecting and maintaining the data. The term *data steward* is often used interchangeably with authoritative source. The distinction is that the authoritative source is the organization that has the assigned authority, and the data steward is the entity in that jurisdiction that is responsible for creating, maintaining, and providing the data, which in smaller jurisdictions can be one and the same but in larger jurisdictions can be one of a jurisdiction's many specialty offices.

*Trusted data* are published data available from a trusted source. Trusted data are typically accessible through the Internet for viewing or downloading or as a web service. The authoritative

source has an agreement with the trusted source for publication. In some cases, the trusted source may be the same as the authoritative source. In other cases, a third party, such as a state or regional agency, may assemble data from many authoritative sources, aggregate their data, and serve as the trusted source for publication in that region.

Data are managed in an organization's production environment for the purpose of carrying out its daily business operations. These operations include very complex databases to support many different business activities. The data are structured to optimize the operation and maintenance workflows. *Publication data*, by contrast, are a small subset of the *production data*. These data are essentially a report in the form of a set of digital data that is combined in as simple a way as possible for efficient distribution, web publishing, or data service. Publication data formats and requirements do not in any way put requirements on the production environment.

The *timeliness* of parcel data can be divided into two types, the certified roll and the working roll. Knowing the data source and where it is in the life cycle of those data is important. The *certified roll* is the official, assessed value of the properties in the county that is provided to the treasurer or tax collector. A snapshot of the current status is taken (often in January, but it varies from state to state), and those data are scrubbed and verified until a countywide assessment of property values can be made. This process can take 6 to 9 months, meaning a certified roll is usually 1 year old at the time of publication and as many as 2 years old by the time the next certified roll is released. The *working roll* includes the ongoing production data that are updated as new properties and values are accumulated. The working roll may not be complete, and many counties do not publish or share their working roll. The update cycle for the working roll can be daily or annually.

The *assessment cycle* is the time between updates to property valuations. Most of the larger counties (with populations of 50,000 or more) are on a 1-year assessment cycle, but some smaller counties may be on a 2- or even 3-year cycle. Most states specify a minimum cycle for updates to assessments, typically 3 to 5 years. The primary responsibility of a local government assessor is to provide a certified roll of values for properties to the treasurer or tax collector. The *mil rate*, the tax rate per \$1,000 of value, is computed against the certified roll values to determine the tax bill amount. The state often provides requirements with specific deadlines for delivering the

certified roll to the treasurer or tax collector and for occasional summary or full records reporting to the state. Within a state, state regulations cause many similarities in scheduling and timing. The differences in state regulations cause much greater variance between states.

Accessing parcel data requires knowledge of additional terms. The *public records laws* vary from state to state and, when left to local governments to interpret and implement, from county to county. In general, public records laws allow a cost of reproduction that usually is less than \$100. Many counties waive fees for government-to-government sharing by requiring only some form of agreement. Many local government data are available and many are freely available from the Internet. Some states are inclined to support charging more than cost-recovery fees, but these states are few, and many counties in these states still provide reasonable access to their data because they see it as a public service. One of the confounding aspects of public records is the case in which counties or states have entered into an exclusive contractual arrangement with a private-sector vendor to create a parcel GIS layer or to publish the tax information at no cost to the county or state, an agreement that allows the vendor to charge on the assumption that the vendor

can make a profit. Typically, no exception exists for G2G data sharing with this type of access. The experience with the U.S. Department of Housing and Urban Development project collection indicates that these counties make up less than 1 percent of all counties.

Local governments also often require *data-sharing agreements*. The authoritative source wants data requestors to come to them or to a trusted data source so that agreements and disclaimers are signed. The reasons for this requirement are practical. Parcel data are highly valued and an important part of most local government operations, which is why fees are often charged for collecting them. A problem also occurs when *shadow data*, which are data that have been acquired from the county and then modified in some way, are mistaken for current data. It is not uncommon for questions to be asked of, or even legal action to be taken against, the authoritative source based on information in shadow data. To minimize confusion and reduce data clarification costs, the authoritative source often requests a data-sharing agreement. Ensuring that professionally accepted definitions were used and understood in this study by the respondent counties was only the first of many challenges.

## Appendix D. Data Collection Status of Study Counties

Exhibit D-1. Status of Collection for Selected Study Counties (1 of 3)

County	State	Final Data Collection Status (Reason)	Fee Purchase Requirement	MOU Requirement (MOU Source)
Jefferson	AL	No data collected (high fee)	NA	NA
Maricopa	AZ	Collected	NA	NA
Pima	AZ	Collected	NA	NA
Alameda	CA	Collected	\$20	NA
Contra Costa	CA	Collected	NA	NA
Fresno	CA	Collected	NA	NA
Kern	CA	Collected	NA	NA
Los Angeles	CA	Collected	\$1,293	NA
Orange	CA	No data collected (high fee)	NA	NA
Riverside	CA	Collected	NA	NA
Sacramento	CA	Collected	NA	NA
San Bernardino	CA	Collected	\$256	NA
San Diego	CA	Collected	\$125	NA
San Joaquin	CA	Collected	NA	NA
Solano	CA	Collected	\$170	NA
Stanislaus	CA	Collected	NA	NA
Tulare	CA	Collected	NA	NA
Adams	CO	Collected	NA	NA
Denver	CO	Collected	NA	NA
El Paso	CO	No data collected (high requirements)	NA	NA
Brevard	FL	Collected	NA	NA
Broward	FL	Collected	NA	NA
Collier	FL	Collected	NA	NA
Duval	FL	Collected	NA	NA
Escambia	FL	Collected	NA	NA
Hillsborough	FL	Collected	NA	NA
Lake	FL	Collected	NA	NA
Lee	FL	Collected	NA	NA
Manatee	FL	Collected	NA	NA
Marion	FL	Collected	NA	NA
Miami Dade	FL	Collected	NA	NA
Orange	FL	Collected	NA	NA
Osceola	FL	Collected	NA	NA
Palm Beach	FL	Collected	NA	NA
Pasco	FL	Collected	NA	NA
Pinellas	FL	Collected	NA	NA
Polk	FL	Collected	NA	NA
Sarasota	FL	Collected	NA	NA
Seminole	FL	Collected	NA	NA
St. Lucie	FL	Collected	NA	NA
Volusia	FL	Collected	NA	NA
Augusta-Richmond	GA	Collected	\$100	NA
Chatham	GA	Collected	\$1,000	NA
Clayton	GA	Collected	\$1,100	NA
Cobb	GA	Collected	NA	NA
DeKalb	GA	No data collected (high requirements)	NA	NA
Fulton	GA	Collected	\$25	NA
Gwinnett	GA	Collected	NA	NA
Muskogee	GA	No data collected (nonresponsive)	NA	NA
Cook	IL	No data collected (nonresponsive)	NA	County
DuPage	IL	Collected	NA	HUD
Kane	IL	Collected	NA	County

Exhibit D-1. Status of Collection for Selected Study Counties (2 of 3)

County	State	Final Data Collection Status (Reason)	Fee Purchase Requirement	MOU Requirement (MOU Source)
Lake	IL	Collected	\$52	NA
McHenry	IL	Collected	NA	NA
St. Clair	IL	Collected	NA	County
Will	IL	Collected	NA	NA
Winnebago	IL	Collected	NA	NA
Allen	IN	Collected	NA	NA
Delaware	IN	Collected	NA	NA
Elkhart	IN	Collected	NA	NA
Hamilton	IN	Collected	NA	NA
Howard	IN	Collected	NA	NA
Lake	IN	Collected	NA	NA
Madison	IN	Collected	NA	NA
St. Joseph	IN	Collected	NA	NA
Vanderburgh	IN	Collected	NA	NA
Louisville/Jefferson	KY	Collected	\$652	HUD
East Baton Rouge	LA	No data collected (unavailable)	NA	NA
Orleans	LA	Collected	NA	NA
Hampden	MA	No data collected (unavailable)	NA	NA
Plymouth	MA	No data collected (unavailable)	NA	NA
Worcester	MA	City of Worcester data collected only	\$150	NA
Baltimore City	MD	Collected	NA	NA
Baltimore County	MD	Collected	\$350	NA
Montgomery	MD	Collected	\$350	NA
Prince Georges	MD	Collected	\$350	NA
Genesee	MI	Collected	NA	NA
Ingham	MI	Collected	\$100	NA
Kent	MI	Collected	NA	County
Macomb	MI	No data collected (unavailable)	NA	NA
Oakland	MI	No data collected (nonresponsive)	NA	NA
Washtenaw	MI	Collected	NA	NA
Wayne	MI	No data collected (nonresponsive)	NA	NA
Anoka	MN	Collected	NA	NA
Dakota	MN	Collected	NA	NA
Hennepin	MN	Collected	NA	County
St. Louis	MO	Collected	\$15	NA
Jackson	MS	Collected	NA	NA
Mecklenburg	NC	Collected	NA	NA
Bergen	NJ	Collected	NA	NA
Essex	NJ	No data collected (unavailable)	NA	NA
Hudson	NJ	Collected	NA	NA
Passaic	NJ	Collected	NA	NA
Union	NJ	Collected	NA	NA
Clark	NV	Collected	\$200	NA
Nassau	NY	No data collected (nonresponsive)	NA	NA
New York	NY	Collected	NA	NA
Orange	NY	Collected	NA	NA
Suffolk	NY	No data collected (nonresponsive)	NA	NA
Butler	OH	Collected	NA	NA
Cuyahoga	OH	Collected	NA	NA
Franklin	OH	Collected	NA	NA
Hamilton	OH	Collected	NA	NA
Lake	OH	Collected	NA	NA
Lorain	OH	Collected	NA	NA
Montgomery	OH	Collected	NA	NA
Stark	OH	Collected	NA	NA
Summit	OH	Collected	NA	NA
Allegheny	PA	Collected	\$104	NA
Lehigh	PA	Collected	\$1,000	NA

**Exhibit D-1. Status of Collection for Selected Study Counties (3 of 3)**

County	State	Final Data Collection Status (Reason)	Fee Purchase Requirement	MOU Requirement (MOU Source)
Philadelphia	PA	Collected	\$100	NA
York	PA	Collected	\$1,194	County
Greenville	SC	Collected	\$500	County
Richland	SC	No data collected (nonresponsive)	NA	NA
Davidson	TN	No data collected (nonresponsive)	NA	NA
Hamilton	TN	Collected	NA	NA
Knox	TN	No data collected (high fee)	NA	NA
Shelby	TN	Collected	NA	NA
Bexar	TX	Collected	NA	NA
Dallas	TX	Collected	NA	NA
El Paso	TX	Collected	NA	NA
Harris	TX	Collected	NA	NA
Hidalgo	TX	Collected	NA	NA
Tarrant	TX	Collected	\$62	NA
Fairfax	VA	Collected	NA	NA
Prince William	VA	Collected	NA	NA
Milwaukee	WI	Collected	NA	NA

MOU = memorandum of understanding. NA = not applicable.

Note: Does not include additional counties in which data were collected beyond the original study request.



## Appendix E. Original Parcel Attributes

Exhibit E-1 describes the individual parcel attributes captured by the project. Characteristics of each of the parcel data elements, also known as data documentation, were also collected to correctly interpret the structure, format, and content of the data elements.

### Exhibit E-1. Parcel Attributes, Descriptions, and Comments (1 of 2)

Parcel Attribute	Attribute Description	Attribute Discussion	Additional Metadata
Parcel identifier	Parcel identifier assigned and managed by the local authority	<ol style="list-style-type: none"> <li>1. Some counties recycle parcel identifiers periodically. We will collect the year associated with the identifier to preserve its uniqueness.</li> <li>2. Parcels may split over time, and we will document how split parcels are identified.</li> <li>3. We will also create a national parcel ID by appending the FIPS code to the parcel ID.</li> </ol>	<ol style="list-style-type: none"> <li>1. Are parcel identifiers recycled?</li> <li>2. How are split parcels identified?</li> </ol>
Parcel identifier year	Year associated with parcel identifier	The year the parcel number first appears on the assessment rolls or in the local government datasets.	<ol style="list-style-type: none"> <li>1. Calendar year or fiscal year?</li> </ol>
Site address	Address of the parcel location	<ol style="list-style-type: none"> <li>1. This is one of the harder attributes to collect. Many tax assessment systems do not maintain high-quality site address because assessors are more interested in the mailing address of the owner.</li> <li>2. Site addresses may appear in inconsistent fields or the comments fields.</li> <li>3. Some counties do not collect site addresses for all units in a parcel; for example, site address of only the first unit in a duplex may appear in the data.</li> <li>4. We will attempt to obtain the site address from other county departments as much as possible.</li> </ol>	<ol style="list-style-type: none"> <li>1. Number and format of address fields</li> </ol>
Subdivision or neighborhood	Subdivision or neighborhood identifier	Beyond site address, we will also collect subdivision and neighborhood information if available.	
Owner type	Indicator that property is owner-occupied or a rental property	For some counties, this attribute may not exist in a form desired by HUD. When available, we will collect other fields, such as tax exempt code or homestead credit indicator that can be used to establish owner type.	
Structure type	Type of home (for example, manufactured, single-family, condominium, townhouse, and so on)	We may need to collect multiple data fields from different tables to create a separate field for structure type.	
Assessed value	Assessed value of property taxed by the state	Assessed values obtained from state sources are likely to be normalized across the state. We will document the basis of value and note any tax cap or state normalization. We will collect the total assessed value and the assessed values of the land and improvements. Some counties may not keep separate assessed values for land and improvements. In some states, notably California, assessed values are also capped at the local level. These exceptions were noted in the data documentation.	<ol style="list-style-type: none"> <li>1. Basis of value</li> <li>2. Normalized?</li> <li>3. Tax cap?</li> </ol>
Assessment date	Date property was assessed	Not all parcels in the data may have been assessed in the same year or cycle. In most tax rolls, another data field also exists called "certification date"—the date when the entire roll is verified and certified (not the assessment date).	<ol style="list-style-type: none"> <li>1. Assessed date or certified date?</li> </ol>
Assessment timeframe	Frequency of assessment (for example, annual or biennial)	Sometimes counties may perform assessments more or less frequently than the state-mandated assessment timeframe.	<ol style="list-style-type: none"> <li>1. Special assessment timeframe?</li> </ol>
Sales price	Value of sale	Given that each parcel can be sold multiple times, we will collect the historical sales prices that are available. Full-disclosure states should have this information available, but it will be difficult to collect from nondisclosure states. We will document the reliability of this information. We will also collect data on the type of sales. The level of sales information in the assessment system may vary from county to county and it may not be feasible to receive sales information from secondary sources. In addition, sales information from secondary sources may be tied to addresses rather than parcel identifier and may be difficult to use.	<ol style="list-style-type: none"> <li>1. Maximum number of sales recorded?</li> <li>2. Any years excluded?</li> </ol>

**Exhibit E-1. Parcel Attributes, Descriptions, and Comments (2 of 2)**

Parcel Attribute	Attribute Description	Attribute Discussion	Additional Metadata
Sales type	Type of sale (for example, arm's length, foreclosure, deed in lieu, and so on)	See previous discussion of sales price.	
Sales date	Year of transaction	Some counties may have the transaction date, whereas others may have the recorded date of the sale.	1. Type of sales date
Area of lot	Lot size	This information is available from the GIS data.	
Area of structure	Square footage of living area	We will also collect unit for area, if available, so that it will be possible to express the area in a standard unit. This information may not be automated, however.	
Structure area units	Units of structure area	See previous discussion of structure area.	
Land use code	How the parcel is used	Three potential candidates exist for this attribute. 1. Current use determined by a land planning office or zoning department. 2. Tax use classification code used in classification for assessment and valuation. 3. Future or planned use as expressed in the master plan or economic development plan. From our experience, this attribute is difficult to standardize.	1. Code description
Easements	Current easements of the property	This attribute is handled differently across the nation. It is frequently not available, or when available, it may only be captured as a map layer rather than an attribute by the utility department. It is certainly a difficult attribute to standardize.	
Condition of property	Current condition of the property	This attribute is not always available. We will also collect the year when condition was determined and the year when the property was built.	1. Letter grade, number grade, or text field?
Condition year	Year associated with condition of property	See previous discussion on property condition.	
Year built	Year property was built	See previous discussion on property condition.	
Multifamily unit	Structure with more than four units	Multifamily unit may already be defined by the structure type attribute. This attribute may also need to be reconciled with structure type.	
Liens	All liens on the property	Tax delinquency liens may not be available through standard data collection sources. When possible, we will also collect available information such as linkage to the lien document, lien amount, lien holder, and year of the lien.	

## Appendix F. Final Parcel Attributes

**Exhibit F-1. Final Attributes in Primary Attribute File**

Variable Name	Label
1 FIPS	County FIPS code
2 Parcelid	Parcel identifier
3 AltID	Alternate parcel identifier
4 GNISID	Placeholder field—Left Blank
5 ParcelYr	Year or date associated with parcel identifier
6 Street_address	Site street address
7 City	Site city
8 ZIP	Site ZIP Code
9 Nbhd	Neighborhood name or code
10 Subdivision	Subdivision name or code
11 Loc_owner_type	Locally defined parcel owner type
12 Structure_code	Type of structure (coded)
13 Structure_desc	Type of structure (description)
14 AssessedVal	Total actual value of property
15 AssessedVal_land	Actual value of property land
16 AssessedVal_bldg	Actual value of property structure(s)
17 AssessedVal_date	Date or year of actual value appraisal
18 TaxVal	Actual or tax-assessed value of property
19 TaxVal_land	Actual or tax-assessed value of property land
20 TaxVal_bldg	Actual or tax-assessed value of property structures
21 TaxVal_date	Date or year of tax assessment
22 Add_Val_Flag	Additional valuation information flag (= 1 if additional valuation data are available in raw data, = 0 otherwise)
23 SalePrice	Price of most recent sale
24 SaleType	Type of most recent sale
25 SaleDate	Date of most recent sale
26 SaleAtt1	Miscellaneous attributes of most recent sale (for example, deed type)
27 SaleAtt2	Miscellaneous attributes 2 of most recent sale (for example, land versus land/structure sale)
28 LotArea	Area of parcel
29 Area Units	Units in which area of parcel is measured
30 ImpArea	Area of structure(s) on parcel
31 ImpAreaUnits	Units in which area of structure on parcel is measured
32 ImpYr	Year structure(s) on parcel were constructed
33 Loc_land_use_code	Locally defined land use or activity associated with parcel (coded)
34 Loc_land_use_desc	Locally defined land use or activity associated with parcel (description)
35 PropClass	Locally defined property classification
36 PropCond	Condition of structure(s) on property or quality of structure(s) construction
37 PropCondYr	Year or date associated with property condition or quality
38 MultiFam	Multifamily structure indicator flag (= 1 if contains four or more units, = 0 otherwise)
39 Occupancy	Information regarding who occupies a parcel
40 Dup_Sale_Flag	Duplicate sales record flag (= 1 if sales data file contains duplicate parcel-transfer record excluded from primary attribute file, = 0 otherwise)
41 Dup_Rec_Flag	Duplicate record flag (= 1 if duplicate parcel identifier, = 0 otherwise)

**Exhibit F-2. Final Attributes in Sales File**

Variable Name	Label
1 FIPS	County FIPS code
2 Parcelid	Parcel identifier
3 AltID	Alternate parcel identifier
4 SalePrice	Price of sale
5 SaleType	Type of sale
6 SaleDate	Date of sale
7 SaleAtt	Miscellaneous attributes of sale (for example, deed type)
8 SaleAtt2	Miscellaneous attributes 2 of sale (for example, land versus land/structure sale)

**Exhibit F-3. Attributes Dropped From Study**

Variable Name	Label	Reason for Dropping
1 Std_owner_type	Standardized version of locally defined owner type	Did not receive enough information or data.
2 Std_structure_desc	Standardized structure description (land only or land and structure)	Too diverse. Could not produce effective standardized field.
3 Std_land_use_desc	Standardized land use description (residential or nonresidential)	Too diverse. Could not produce effective standardized field.
4 Asd_val_use_code, Asd_val_use_desc	Method of property valuation (for example, market, cost, income)	Relatively consistent within each county. Captured in metadata.
5 AssessedVal_Cycle	Cycle of property valuation	Relatively consistent within each county. Captured in metadata.
6 Easement	Easements on parcel	Did not receive enough information or data.
7 Liens	All liens on property, including tax delinquency liens	Dropped from analysis in initial research design.
8 Census_Tract	Census tract containing parcel	Not in initial research design; did not receive enough information or data.
9 Tract_Year	Census year (for example, 2000 or 2010) for Census Tract	Not in initial research design; did not receive enough information or data.
10 Foreclosure	Foreclosure sale flag (= 1 if foreclosure sale, = 0 otherwise)	Not in initial research design; did not receive enough information or data. (Note: foreclosure sales captured in SaleType attribute.)
11 Fore_Date	Date property went into foreclosure	Not in initial research design; did not receive enough information or data.
12 ImpArea	Area of structure(s) on parcel	Initially planned to use ImpArea as an alternate parcel identifier; improvement area, however, is rarely captured in sales data.

## Appendix G. Parcel Data Comprehensiveness Findings

**Exhibit G-1. Dataset Types Collected (HUD-Selected Counties Only) (1 of 3)**

County	State	Attribute	Sales	GIS	Historical	Attribute, Sales, GIS	Attribute and Sales Only	Attribute and GIS Only	GIS With Sales Only	GIS Only	Attribute Only
Maricopa	AZ	1	1	1	0	1	0	0	0	0	0
Pima	AZ	1	1	0	0	0	1	0	0	0	0
Alameda	CA	1	0	1	0	0	0	1	0	0	0
Contra Costa	CA	1	0	1	0	0	0	1	0	0	0
Fresno	CA	1	0	1	0	0	0	1	0	0	0
Kern	CA	1	1	1	1	1	0	0	0	0	0
Los Angeles	CA	1	0	1	0	0	0	1	0	0	0
Riverside	CA	1	1	1	0	1	0	0	0	0	0
Sacramento	CA	1	0	1	0	0	0	1	0	0	0
San Bernardino	CA	1	0	1	0	0	0	1	0	0	0
San Diego	CA	1	0	1	1	0	0	1	0	0	0
San Joaquin	CA	1	0	1	0	0	0	1	0	0	0
Solano	CA	1	1	0	0	0	1	0	0	0	0
Stanislaus	CA	1	1	1	0	1	0	0	0	0	0
Tulare	CA	1	0	0	0	0	0	0	0	0	1
Adams	CO	1	1	1	0	1	0	0	0	0	0
Denver (City)	CO	1	1	1	0	1	0	0	0	0	0
Brevard	FL	1	1	1	1	1	0	0	0	0	0
Broward	FL	1	1	1	1	1	0	0	0	0	0
Collier	FL	1	1	1	1	1	0	0	0	0	0
Duval	FL	1	1	1	1	1	0	0	0	0	0
Escambia	FL	1	1	1	1	1	0	0	0	0	0
Hillsborough	FL	1	1	1	1	1	0	0	0	0	0
Lake	FL	1	1	1	1	1	0	0	0	0	0
Lee	FL	1	1	1	1	1	0	0	0	0	0
Manatee	FL	1	1	1	1	1	0	0	0	0	0
Marion	FL	1	1	1	1	1	0	0	0	0	0
Miami-Dade	FL	1	1	1	1	1	0	0	0	0	0
Orange	FL	1	1	1	1	1	0	0	0	0	0
Osceola	FL	1	1	1	1	1	0	0	0	0	0
Palm Beach	FL	1	1	1	1	1	0	0	0	0	0
Pasco	FL	1	1	1	1	1	0	0	0	0	0
Pinellas	FL	1	1	1	1	1	0	0	0	0	0
Polk	FL	1	1	1	1	1	0	0	0	0	0
Sarasota	FL	1	1	1	1	1	0	0	0	0	0
Seminole	FL	1	1	1	1	1	0	0	0	0	0
St. Lucie	FL	1	1	1	1	1	0	0	0	0	0
Volusia	FL	1	1	1	1	1	0	0	0	0	0



Exhibit G-1. Dataset Types Collected (HUD-Selected Counties Only) (2 of 3)

County	State	Attribute	Sales	GIS	Historical	Attribute, Sales, GIS	Attribute and Sales Only	Attribute and GIS Only	GIS With Sales Only	GIS Only	Attribute Only
Augusta-Richmond	GA	1	1	0	1	0	1	0	0	0	0
Cobb	GA	1	1	0	0	0	1	0	0	0	0
Fulton	GA	1	1	0	0	0	1	0	0	0	0
Gwinnett	GA	1	1	1	0	1	0	0	0	0	0
Chatham	GA	1	1	0	0	0	1	0	0	0	0
Clayton	GA	1	1	0	1	0	1	0	0	0	0
DuPage	IL	0	0	1	0	0	0	0	0	1	0
Kane	IL	1	0	1	0	0	0	1	0	0	0
Lake	IL	1	1	1	0	1	0	0	0	0	0
McHenry	IL	0	0	1	0	0	0	0	0	1	0
St. Clair	IL	1	1	1	0	1	0	0	0	0	0
Will	IL	1	1	0	0	0	1	0	0	0	0
Winnebago	IL	1	1	1	1	1	0	0	0	0	0
Allen	IN	1	1	0	0	0	1	0	0	0	0
Delaware	IN	1	0	1	0	0	0	1	0	0	0
Hamilton	IN	1	0	1	1	0	0	1	0	0	0
St. Joseph	IN	1	0	1	1	0	0	1	0	0	0
Vanderburgh	IN	1	1	1	1	1	0	0	0	0	0
Elkhart	IN	1	1	0	1	0	1	0	0	0	0
Lake	IN	1	1	0	1	0	1	0	0	0	0
Madison	IN	1	1	0	1	0	1	0	0	0	0
Jefferson	KY	1	1	1	0	1	0	0	0	0	0
New Orleans	LA	0	0	1	0	0	0	0	0	1	0
Worcester	MA	1	0	1	0	0	0	1	0	0	0
Baltimore	MD	1	1	1	1	1	0	0	0	0	0
Baltimore (city)	MD	1	1	1	0	1	0	0	0	0	0
Montgomery	MD	1	1	0	1	0	1	0	0	0	0
Prince George's	MD	1	1	1	1	1	0	0	0	0	0
Genesee	MI	1	0	1	0	0	0	1	0	0	0
Ingham	MI	1	1	1	0	1	0	0	0	0	0
Kent	MI	1	1	1	0	1	0	0	0	0	0
Oakland	MI	1	1	0	0	0	1	0	0	0	0
Washtenaw	MI	1	0	1	0	0	0	1	0	0	0
Anoka	MN	1	1	1	1	1	0	0	0	0	0
Dakota	MN	1	1	1	0	1	0	0	0	0	0
Hennepin	MN	1	0	0	0	0	0	0	0	0	1
St. Louis	MO	1	1	1	0	1	0	0	0	0	0
Jackson	MS	1	1	1	0	1	0	0	0	0	0
Mecklenburg	NC	1	1	1	0	1	0	0	0	0	0
Bergen	NJ	1	0	0	0	0	0	0	0	0	1
Hudson	NJ	1	0	0	0	0	0	0	0	0	1
Passaic	NJ	1	0	0	0	0	0	0	0	0	1

Exhibit G-1. Dataset Types Collected (HUD-Selected Counties Only) (3 of 3)

County	State	Attribute	Sales	GIS	Historical	Attribute, Sales, GIS	Attribute and Sales Only	Attribute and GIS Only	GIS With Sales Only	GIS Only	Attribute Only
Union	NJ	1	1	0	0	0	1	0	0	0	0
Clark	NV	1	0	0	0	0	0	0	0	0	1
Orange	NY	1	1	1	1	1	0	0	0	0	0
New York City	NY/NJ	1	1	0	0	0	1	0	0	0	0
Butler	OH	1	1	1	1	1	0	0	0	0	0
Franklin	OH	1	1	1	0	1	0	0	0	0	0
Cuyahoga	OH	1	1	0	1	0	1	0	0	0	0
Hamilton	OH	1	1	1	0	1	0	0	0	0	0
Lake	OH	1	1	0	0	0	1	0	0	0	0
Lorain	OH	1	0	1	0	0	0	1	0	0	0
Montgomery	OH	1	1	1	0	1	0	0	0	0	0
Stark	OH	1	0	1	1	0	0	1	0	0	0
Summit	OH	1	1	1	0	1	0	0	0	0	0
Allegheny	PA	1	1	0	0	0	1	0	0	0	0
Lehigh	PA	1	1	1	0	1	0	0	0	0	0
York	PA	1	1	1	0	1	0	0	0	0	0
Philadelphia	PA	1	1	0	0	0	1	0	0	0	0
Greenville	SC	1	1	1	0	1	0	0	0	0	0
Hamilton	TN	1	1	1	0	1	0	0	0	0	0
Shelby	TN	1	1	1	0	1	0	0	0	0	0
Bexar	TX	1	0	1	0	0	0	1	0	0	0
Dallas County	TX	1	0	1	1	0	0	1	0	0	0
El Paso	TX	1	0	0	0	0	0	0	0	0	1
Harris	TX	1	0	1	1	0	0	1	0	0	0
Hidalgo	TX	1	0	1	0	0	0	1	0	0	0
Tarrant	TX	1	1	1	1	1	0	0	0	0	0
Fairfax	VA	1	1	0	0	0	1	0	0	0	0
Prince William	VA	1	1	0	1	0	1	0	0	0	0
Milwaukee	WI	1	1	1	0	1	0	0	0	0	0

Exhibit G-2. Parcel Attributes Collected (HUD-Selected Counties Only) (1 of 4)

County	State	Parcel ID	Alternate Parcel ID	Year Associated w/ Parcel ID	Parcel Street Address	Parcel City	Parcel State	Parcel Zip Code	Parcel Neighborhood	Parcel Subdivision	Locally Defined Owner Type	Structure (Code)	Structure (Description)	Parcel Value	Value of Parcel Land	Value of Parcel Buildings	Date of Value Determination	Sale Price	Sale Type	Date of Sale	Additional Sale Attributes (1)	Additional Sale Attributes (2)	Parcel Area	Parcel Area Units	Building Area	Building Area Units	Improvement Year	Local Land Use (Code)	Local Land Use (Description)	Property Classification	Property Condition	Year of Property Condition	Multifamily Property Flag	Owner Occupancy Flag
Maricopa	AZ	1	0	0	1	1	1	1	0	1	0	1	0	1	1	1	0	1	0	1	1	0	1	1	1	1	1	1	1	0	0	1	1	
Pima	AZ	1	1	1	1	1	1	0	1	1	0	0	0	1	1	0	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	0	0	
Alameda	CA	1	1	1	1	1	1	0	0	0	1	0	1	1	1	1	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	
Contra Costa	CA	1	1	1	1	1	1	0	0	0	1	0	1	1	1	1	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	
Fresno	CA	1	1	1	0	0	1	0	0	1	0	1	1	1	1	1	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	
Kern	CA	1	1	1	1	1	1	1	0	1	0	0	0	1	0	0	0	1	1	0	0	0	1	1	1	1	1	0	1	0	0	0	0	
Los Angeles	CA	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0	0	0	0	0	1	1	0	0	0	1	1	0	0	0	0	
Riverside	CA	1	1	1	1	1	1	0	0	1	1	0	1	1	1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	
Sacramento	CA	1	1	1	1	1	1	0	0	1	1	0	1	1	1	1	0	0	0	0	0	0	1	1	0	0	0	1	1	0	0	0	0	
San Bernardino	CA	1	1	1	1	1	1	1	1	0	1	0	1	1	0	0	1	1	1	1	0	0	1	1	0	0	0	1	1	0	0	0	0	
San Diego	CA	1	1	1	1	1	1	0	0	1	1	0	1	1	1	1	0	0	0	0	0	0	1	1	1	0	0	1	1	0	0	0	0	
San Joaquin	CA	1	1	1	1	1	1	0	0	0	1	0	1	1	1	1	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	
Solano	CA	1	0	1	1	1	1	0	0	1	0	0	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	0	0	0	1	1
Stanislaus	CA	1	1	1	1	1	1	0	0	0	1	0	1	1	1	1	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0
Tulare	CA	1	1	1	1	1	1	0	0	1	1	0	1	1	1	1	0	0	0	0	0	0	1	1	0	0	0	1	1	0	0	0	0	0
Adams	CO	1	1	0	1	1	1	1	1	1	0	0	1	1	1	1	0	1	0	1	1	0	1	1	1	1	1	1	1	1	1	0	0	0
Denver (city)	CO	1	0	1	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0	1	1	0	0	0	1	1	1	1	1	1	0	0	0	0
Brevard	FL	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1
Broward	FL	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	0	0	1
Collier	FL	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1
Duval	FL	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1
Escambia	FL	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1
Hillsborough	FL	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1
Lake	FL	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1
Lee	FL	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1
Manatee	FL	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1
Marion	FL	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1
Miami-Dade	FL	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1
Orange	FL	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1
Osceola	FL	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1
Palm Beach	FL	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1
Pasco	FL	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1

Exhibit G-2. Parcel Attributes Collected (HUD-Selected Counties Only) (2 of 4)

County	State	Parcel ID	Alternate Parcel ID	Year Associated w/ Parcel ID	Parcel Street Address	Parcel City	Parcel State	Parcel Zip Code	Parcel Neighborhood	Parcel Subdivision	Locally Defined Owner Type	Structure (Code)	Structure (Description)	Parcel Value	Value of Parcel Land	Value of Parcel Buildings	Date of Value Determination	Sale Price	Sale Type	Date of Sale	Additional Sale Attributes (1)	Additional Sale Attributes (2)	Parcel Area	Parcel Area Units	Building Area	Building Area Units	Improvement Year	Local Land Use (Code)	Local Land Use (Description)	Property Classification	Property Condition	Year of Property Condition	Multifamily Property Flag	Owner Occupancy Flag	
Pinellas	FL	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1	
Polk	FL	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	0	1	
Sarasota	FL	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	0	0	0	1		
Seminole	FL	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	0	0	0	1		
St. Lucie	FL	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	0	0	0	1		
Volusia	FL	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	0	0	0	1		
Augusta-Richmond	GA	1	1	1	1	0	1	0	1	0	1	0	0	1	0	0	1	1	1	0	1	0	1	1	1	1	0	0	0	1	0	0	1	0	
Chatham	GA	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	0	1	0	1	1	1	0	0	0	0	0	1	0	1	0	0	1	0	
Clayton	GA	1	0	1	1	0	1	1	0	0	0	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	0	1	0	0	0	
Cobb	GA	1	0	1	1	1	1	0	0	0	0	1	0	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	0	1	1	0	0	0	
Fulton	GA	1	0	1	1	1	1	0	1	0	0	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	0	1	1	0	0	0	
Gwinnett	GA	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	0	0	0	1	1	1	1	1	0	1	1	0	0	0	0	
DuPage	IL	1	0	0	1	1	1	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	1	1	0	0	0	1	0	1	0	0	0	
Kane	IL	1	0	1	1	1	1	1	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	1	
Lake	IL	1	0	0	1	1	1	1	1	1	0	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	
McHenry	IL	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
St. Clair	IL	1	0	1	1	1	1	1	0	1	0	0	0	1	1	1	0	1	0	1	1	0	1	1	0	0	0	1	1	1	0	0	0	1	
Will	IL	1	0	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1	0	1	1	1	1	0	0	0	0	0	1	0	0	0	0	1	
Winnebago	IL	1	0	1	1	1	1	1	0	0	1	0	0	1	1	1	0	1	1	1	1	0	1	1	0	0	0	1	1	1	0	0	0	1	
Allen	IN	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0	0	0	0	0	1	1	0	0	1	1	0	0	0	0	0	
Delaware	IN	1	1	0	1	0	1	1	1	0	1	0	1	1	1	1	0	0	0	0	0	0	0	1	1	1	1	1	1	0	1	1	0	0	0
Elkhart	IN	1	1	1	0	0	1	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	1	1	0	0
Hamilton	IN	1	1	1	1	1	1	1	0	1	0	0	0	1	1	1	1	0	0	0	0	0	0	1	1	0	0	0	1	1	0	0	0	0	
Lake	IN	1	1	1	0	0	1	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	1	1	0	0
Madison	IN	1	1	1	0	0	1	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	1	1	0	0
St. Joseph	IN	1	1	0	1	1	1	1	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
Vanderburgh	IN	1	1	0	1	1	1	1	1	0	0	0	0	1	1	1	1	1	0	1	1	0	1	1	0	0	1	1	0	1	1	0	1	0	
Jefferson	KY	1	0	0	1	1	1	1	0	0	0	0	0	1	1	1	1	1	0	1	0	0	1	1	1	1	1	1	1	1	0	0	1	0	
New Orleans	LA	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Worcester	MA	1	1	0	1	1	1	1	0	1	0	0	1	1	1	1	0	0	0	0	0	0	0	1	1	0	0	0	1	1	0	0	1	0	0
Baltimore	MD	1	0	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	0	1	1	0	1	
Baltimore (city)	MD	1	0	1	1	1	1	1	1	0	0	0	0	1	1	1	0	1	0	1	0	0	0	0	1	1	1	1	1	0	0	0	1	1	

Exhibit G-2. Parcel Attributes Collected (HUD-Selected Counties Only) (3 of 4)

County	State	Parcel ID	Alternate Parcel ID	Year Associated w/ Parcel ID	Parcel Street Address	Parcel City	Parcel State	Parcel Zip Code	Parcel Neighborhood	Parcel Subdivision	Locally Defined Owner Type	Structure (Code)	Structure (Description)	Parcel Value	Value of Parcel Land	Value of Parcel Buildings	Date of Value Determination	Sale Price	Sale Type	Date of Sale	Additional Sale Attributes (1)	Additional Sale Attributes (2)	Parcel Area	Parcel Area Units	Building Area	Building Area Units	Improvement Year	Local Land Use (Code)	Local Land Use (Description)	Property Classification	Property Condition	Year of Property Condition	Multifamily Property Flag	Owner Occupancy Flag
Montgomery	MD	1	0	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	0	1	1	0	1	
Prince George's	MD	1	0	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	0	1	1	0	1	
Genesee	MI	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0		
Ingham	MI	1	0	1	1	1	1	1	0	0	0	0	0	1	1	1	0	1	1	1	1	1	1	0	0	0	0	0	1	0	0	0		
Kent	MI	1	0	1	1	1	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0		
Oakland	MI	1	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	
Washtenaw	MI	1	0	1	1	1	1	1	1	1	0	1	0	1	0	0	0	0	0	0	0	0	1	1	0	0	1	1	0	1	0	0	0	
Anoka	MN	1	0	0	1	1	1	1	1	1	0	0	1	1	1	1	0	1	0	1	0	0	1	1	1	1	1	0	1	0	0	0	0	
Dakota	MN	1	0	0	1	1	1	1	0	0	0	0	0	1	1	1	0	1	0	1	0	0	1	1	1	1	1	0	1	0	0	0	0	
Hennepin	MN	1	0	0	0	1	1	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0
St. Louis	MO	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0
Jackson	MS	1	1	1	1	1	1	1	1	0	0	0	0	1	1	1	0	1	0	1	1	0	1	1	0	0	1	0	0	0	1	0	0	0
Mecklenburg	NC	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0
Bergen	NJ	1	0	1	1	0	1	1	0	0	0	1	0	1	1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0	0
Hudson	NJ	1	0	1	1	0	1	1	0	0	0	1	0	1	1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0	0
Passaic	NJ	1	0	1	1	0	1	1	0	0	0	1	0	1	1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0	0
Union	NJ	1	1	0	1	0	1	0	0	0	0	1	0	1	1	1	0	1	0	1	0	0	1	1	0	0	1	1	1	1	0	0	0	0
Clark	NV	1	1	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0	1	0	1	0
Orange	NY	1	0	1	1	0	1	1	1	0	0	1	0	1	1	1	1	0	1	1	0	0	0	0	1	1	1	1	0	0	1	0	0	0
New York City	NY	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0
Butler	OH	1	1	0	1	1	1	1	1	0	0	0	0	1	0	0	0	1	1	1	0	0	1	1	1	1	1	1	1	0	1	0	0	0
Cuyahoga	OH	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0
Franklin	OH	1	0	0	1	1	1	1	1	0	0	1	1	1	1	1	0	1	0	1	0	0	1	1	1	1	1	1	1	1	1	0	0	1
Hamilton	OH	1	1	0	1	1	1	1	1	0	0	0	0	1	1	1	0	1	1	1	1	1	1	1	0	0	0	1	1	0	0	0	0	1
Lake	OH	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0
Lorain	OH	1	1	1	1	1	1	1	1	0	0	0	1	1	0	0	1	0	0	0	0	0	1	1	1	1	1	1	0	1	1	1	1	1
Montgomery	OH	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1
Stark	OH	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0
Summit	OH	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	0	1	0	1	1	0	1	1	0	0	0	1	1	1	0	0	0	0
Allegheny	PA	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	0
Lehigh	PA	1	1	0	1	1	1	1	0	1	0	1	1	1	1	1	0	1	0	1	0	0	1	1	1	1	1	1	1	1	0	0	1	0
Philadelphia	PA	1	1	1	1	0	1	0	0	1	0	1	0	1	1	1	1	0	1	0	1	0	0	1	1	0	0	1	1	0	0	0	0	0



Exhibit G-2. Parcel Attributes Collected (HUD-Selected Counties Only) (4 of 4)

County	State	Parcel ID	Alternate Parcel ID	Year Associated w Parcel ID	Parcel Street Address	Parcel City	Parcel State	Parcel Zip Code	Parcel Neighborhood	Parcel Subdivision	Locally Defined Owner Type	Structure (Code)	Structure (Description)	Parcel Value	Value of Parcel Land	Value of Parcel Buildings	Date of Value Determination	Sale Price	Sale Type	Date of Sale	Additional Sale Attributes (1)	Additional Sale Attributes (2)	Parcel Area	Parcel Area Units	Building Area	Building Area Units	Improvement Year	Local Land Use (Code)	Local Land Use (Description)	Property Classification	Property Condition	Year of Property Condition	Multifamily Property Flag	Owner Occupancy Flag
York	PA	1	0	0	1	0	1	1	1	0	0	1	0	1	0	0	0	1	1	1	0	0	1	1	1	1	1	1	0	0	1	1	0	0
Greenville	SC	1	1	1	1	1	1	1	0	1	0	0	0	1	1	1	0	1	0	1	0	0	1	1	1	1	0	1	1	1	0	0	1	0
Hamilton	TN	1	0	0	1	1	1	1	1	1	0	1	0	1	1	1	0	1	1	1	1	0	1	1	0	0	0	1	1	1	0	0	0	0
Shelby	TN	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	
Bexar	TX	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	1
Dallas County	TX	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0
El Paso	TX	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	1
Harris	TX	1	0	1	1	1	1	1	1	0	0	1	1	1	1	1	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	0	0	0
Hidalgo	TX	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	1
Tarrant	TX	1	1	1	1	1	1	1	1	0	0	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	
Fairfax	VA	1	0	0	1	1	1	1	1	0	0	1	0	1	1	1	0	1	0	1	0	0	1	1	1	1	1	1	0	1	0	0	1	0
Prince William	VA	1	0	1	1	1	1	1	1	0	0	1	1	1	1	1	0	1	0	1	0	0	1	1	1	1	1	1	0	0	0	0	0	
Milwaukee	WI	1	1	1	1	1	1	1	0	1	0	0	1	1	1	1	0	1	1	1	0	0	1	1	0	0	0	1	1	0	1	0	0	0

## Appendix H. Original Feasibility Questions

This table provides the subcomponents, rating measures, information to collect, and source of collection for the six feasibility measures described in the original research design.<sup>48</sup>

**Exhibit H-1. Instrument and Data Source for Feasibility Measures (1 of 2)**

Feasibility Measure	Subcomponents	Rating	Information To Collect	Source	Notes
<b>1. Acquisition effort</b>	a. How many contacts were needed to identify the primary or lead contact for the initial data collection?	Initial hours, ongoing hours	Count	Track	Is the contact for ongoing data collection the same person or position as the one with whom contacted was established initially?
	b. How many different datasets are needed from the county to build the complete set of attributes for the initial data collection?		Count	Track	
	c. What is the estimated number of hours needed for initial data collection?		Hours	Estimate	
	d. How many contacts will be needed to identify the primary or lead contact for ongoing collection?		Count	Track, estimate	
	e. How many different datasets will be needed from the county to build the complete set of attributes for ongoing collection?		Count	Ask county	
	f. What is the estimated number of hours required for each contact?		Hours	Estimate	
	g. Can data requests be automated for the future, for example, from an e-mail or a web service?		Yes/no	Ask county	
<b>2. Acquisition requirements</b>	a. What are the initial costs for acquiring the data for HUD's use?	Dollar value + agreement type	Dollar value	Ask county, track	Is it possible that HUD could add the provisioning of parcel data to its grant programs? Could any agreement for obtaining the parcel data be embedded in the grant contracts to avoid having to sustain and track the status and language in agreements? Part of the discussion of the feasibility for the collection of the data from these identified counties will review the details of the agreements that may have been needed and provide options for dealing with these terms in the HUD contracts in the future.
	b. What are the additional initial costs for public domain distribution at the summary level?		Dollar value	Ask county, track	
	c. What is the additional cost for subsequent acquisition for HUD's use?		Dollar value	Ask county, track	
	d. What is the additional cost for subsequent acquisition for public domain distribution at the summary level?		Dollar value	Ask county, track	
	e. What type of data sharing agreement is needed?		G2G data sharing, MOU, liability release, contract, license agreement, letter, or nothing	Ask county, track	

<sup>48</sup> A seventh feasibility measure—the catchall “other considerations”—is not listed here.

Exhibit H-1. Instrument and Data Source for Feasibility Measures (2 of 2)

Feasibility Measure	Subcomponents	Rating	Information To Collect	Source	Notes
<b>3. Update frequency</b>	a. How frequently are parcel attributes updated?	Annually, quarterly, monthly, weekly, daily, or continuously	Annually, quarterly, monthly, weekly, daily, or continuously	Ask county, track	Understanding the tax assessment update cycle is important to understanding the timeliness of reported attributes. A discussion of the various cycles (assessment, taxation, and valuation or revaluation) will be included in the final report. Part of the data collection effort will collect information from the counties and states on these various cycles and key reporting dates.
	b. How frequently are these updates posted to a distributable dataset?	Annually, quarterly, monthly, weekly, daily, or continuously	Annually, quarterly, monthly, weekly, daily, or continuously	Ask county, track	
	c. What is the frequency of GIS update?	Annually, quarterly, monthly, weekly, daily, or continuously	Annually, quarterly, monthly, weekly, daily, or continuously	Ask county, track	
	d. Are all the necessary components updated in the same cycle?	Yes/no	Yes/no	Ask county, track	
	e. In what month or months is the tax roll certified?	Month name(s)	Month name(s)	Ask county	
<b>4. Availability</b>	a. How many of the required attributes are included in the dataset for the county? Note provided attributes in crosswalk document.	Percent of attributes, percent of coverage	Count/percent	Track	In the long term, will HUD have to sustain an external FTP site that can be used as dropoff point for datasets? Does any security concern exist with this type of a site? Do any virus or worm issues emerge with injecting the datasets? For example, will availability be limited because the datasets have to be cleaned before they can be injected into HUD's systems?
	b. How many of the optional attributes are available in the dataset for the county?	Count/percent	Count/percent	Track	
	c. What is the coverage of the county with these attributes; for example, is the site address on all of the parcels or a subset of the parcels?	Percent	Percent	Track	
	d. What is the county's ability to provide the data to HUD?	Has FTP, needs FTP, mails DVD	Has FTP, needs FTP, mails DVD	Ask county, track	
<b>5. Standardization effort</b>	a. What is the format of the dataset(s) as provided by the county?	Initial hours, ongoing hours, overall source	Open, proprietary	Track	What is the best format to document and propagate crosswalks? These exchanges should be able to be developed as standard extract, transform, and load routines. Is it the best way to propagate the use of the crosswalks in HUD? Capturing these crosswalks as Microsoft Word documents or in a database or spreadsheet is another option.
	b. How many datasets were provided by the county?	dataset complexity rating	Count	Track	
	c. What is the variability of the content of the data?	Field formatted same throughout, field separated into components with varied structure or formats, and so on	Field formatted same throughout, field separated into components with varied structure or formats, and so on	Track	
	d. How many hours does it take to do the initial crosswalk and standardization?	Hours	Hours	Track	
	e. How many hours are estimated to do subsequent standardization?	Hours	Hours	Estimate	
<b>6. Collection sustainability</b>	a. Is the county planning any new software acquisitions that will affect the format or content of the datasets provided to HUD?	Generally easier trend, no trend, generally harder trend	Yes/no	Ask county	What is the overall sustainability for HUD to continue to collect data from the identified 127 counties? What is the expected annual workload and technological requirements to support this effort? What can HUD expect in the next 2 years, the next 5 years, and the next decade?
	b. Might any known impending retirements or changes in the work force or elected officials have a foreseeable effect on data availability?	Yes/no	Yes/no	Ask county	
	c. How formally is the data sharing arrangement with HUD established?	Institutional, inter personal relationship, and so on	Institutional, inter personal relationship, and so on	Track	

FTP = file transfer protocol. GTG = government-to-government. MOU = memorandum of understanding.

## Appendix I. HUD-Approved Data Sharing Agreement Template

### COUNTY DATA RECORD PROJECT MEMORANDUM OF UNDERSTANDING

---

This Memorandum of Understanding is made and entered into as of the date appearing on the signature page, by and between the **XXX, INSERT ADDRESS**, (hereinafter referred to as “County”) and the **US Department of Housing and Urban Development**, 451 7<sup>th</sup> St. SW Washington DC, 20410-6000 (hereinafter referred to as “HUD”).

#### Section 1: Purpose

1. HUD requests County’s parcel level information on ownership, values and characteristics of residential land parcels from Counties receiving the Neighborhood Stabilization Program (NSP) Grants, to support a HUD project called the County Data Record Project (“Project”).
2. A national non-proprietary aggregation of local land parcel data does not exist, and therefore this pioneering project will provide HUD with unprecedented access to data such as home sales, foreclosures and tax assessments for many counties. Future HUD evaluations using this parcel data will position HUD to respond effectively and efficiently to economic and natural disasters. Furthermore, parcel data will also be useful for effective management of NSP and other HUD programs
3. To accomplish the Project, HUD has contracted with Abt Associates and its subcontractors, Smart Date Strategies (SDS) and Fairview Industries to compile parcel data including GIS data from a select group of counties.

#### Section 2: Authority

Title V of the 1970 Housing Act (12 U.S.C. §1701z) authorizes the Secretary to undertake programs of research, testing, and demonstrations relating to the mission of the Department.

NOW THEREFORE, the County and HUD agree to data sharing on the terms and conditions set forth as follows.

### **Section 3: Contributions of Parties**

1. The County agrees to share the parcel level data with HUD in a form and manner mutually acceptable to the parties under this MOU.
2. County's parcel level data to be shared with HUD will include:
  - a. electronic data files with attribute information describing ownership, sales history, value, location and other relevant information; and
  - b. any associated GIS files to provide location information
3. County shall share the parcel level data with HUD free or nominal charge if necessary.

### **Section 4: Term and Termination**

1. The Project shall commence prior to August 1, 2011, and the parcel level data shall be transferred by the County to HUD on or after August 1, 2011. The parties may continue to perform the Project under the terms of this MOU until terminated by either party under this MOU by an advance notice of thirty (30) days in writing.
2. This MOU will terminate immediately if HUD or County fails to comply with any provision of this MOU. Each party shall have an opportunity to respond to the other party in writing within thirty (30) days of a termination based upon failure to comply with any provision of this MOU, and upon a showing satisfactory to the other party, that in fact no violation actually occurred, the MOU may be reinstated. Otherwise, the MOU may be terminated by either party upon the giving of thirty (30) days written notice.

### **Section 5: Use of Parcel Level Data**

1. HUD will use the county level parcel data for:
  - a) assessing the feasibility of collecting parcel data from counties on a regular basis; and
  - b) standardizing the collected parcel data, both attributes and GIS files, and using it to assess the needs for and effectiveness of NSP and other HUD programs.
2. HUD shall not a) resell any collected parcel data; or b) redistribute the collected data outside of the HUD research community, except in an aggregate geographic level of census tract or higher.



3. Except as required or authorized by law, HUD is not authorized, and agrees not to resell or redistribute, publish, relinquish possession of, transfer, lease, copy, or allow others to copy or sell to others to use, or grant any license to another with regard to collected parcel data without prior written approval of County.
4. This parcel data will not be used to find the location of any individual's residence for any purpose that a reasonable person might find intrusive, threatening, or a violation of their privacy.

### **Section 6: Project Benefit to County**

1. HUD agrees to provide the standardized data to County after the completion of this Project. Counties may benefit from having standardized parcel data to respond to Stafford Act requests for emergency loans and recovery dollars during disaster response.

### **Section 7: Subcontracts**

1. Any subcontractors and outside associates or consultants contracted by HUD in connection with Data Record Project will be covered by this MOU and shall be limited to individuals or firms specifically authorized in writing by HUD. Any substitutions in or additions to subcontractors, associates or consultants shall be subject to the prior written approval of County. HUD shall bind all subcontractors to the provisions of this MOU, and said subcontractors are directly responsible to HUD.

### **Section 8: Notice**

1. All notices, demands or other writings permitted or required by the terms of this MOU shall be deemed to have been fully given, made or sent when made in writing and deposited in the United States Mail, registered and postage prepaid, and addressed to the Contract Administrators as follows:

County:

HUD: **Ashaki Robinson Johns or Jon Sperling**  
US Department of Housing and Urban Development  
Office of Policy Development and Research  
Suite 8120  
Washington, DC 20410

The address to which any notice, demand or other writing may be given or sent to any party may be changed by written notice given to the other party.

**Section 9: Expenses of the Parties**

This MOU does not obligate funds, personnel, services, or other resources of either department to the other. Each party to the MOU will bear its own expenses in connection with the preparation, negotiation, and execution of the MOU, and neither party shall be liable to the other party for such expenses.

This MOU is an expression of intent only, that each party acts as an independent party with respect to the performance of duties under the MOU and does not represent that it is an employee or agent of another party to the MOU.

This MOU does not give a third party any benefit, legal or equitable right, remedy or claim under any agreement contained herein.

**Section 10: Authorization to Sign**

1. The persons signing below on behalf of the parties hereby personally warrant and guarantee that they have the authority to sign this MOU and to bind the entity for which they are signing this MOU.

In Witness whereof this MOU is accepted and entered into on the date indicated below.

**County of XXXXXXXX**

By: \_\_\_\_\_

Date: \_\_\_\_\_

Approved as to Form:

\_\_\_\_\_

Date: \_\_\_\_\_

**United States Department of Housing  
and Urban Development**

By: \_\_\_\_\_

Raphael Bostic

Office of Policy Development and Research

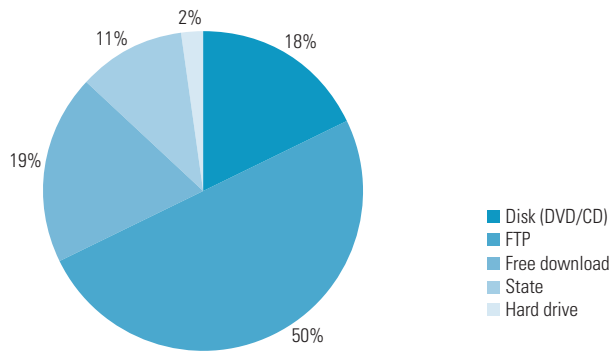
Date: \_\_\_\_\_

## Appendix J. Feasibility Findings

Because a significant number of the study counties either elected not to respond to the feasibility questions, responded only to a portion of the questions, or could not be reached to determine response (for example, among counties whose data are publicly accessible), the feasibility findings are provided only as preliminary evidence for future feasibility and were used to assist in the qualitative categories described in chapter 3.

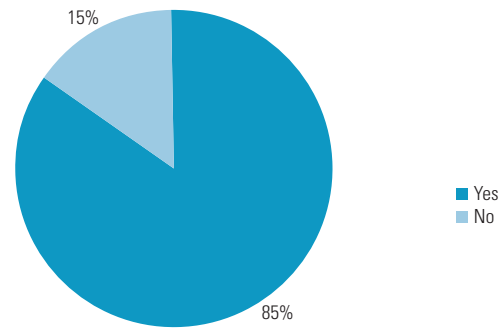
Exhibit J-1 describes the method by which counties offer or prefer to deliver datasets after requests, regardless of the types of data or attributes requested. Exhibit J-2 shows the number of counties that maintain prepackaged parcel datasets to deliver on request. These prepackaged datasets typically do not include all of the attributes that a special data request may solicit, as was the case in this study.

**Exhibit J-1. Method for Accessing Parcel Data**



Note: N = 102. Three jurisdictions responded "NA" and 22 did not respond.

**Exhibit J-2. Prepackaged Parcel Dataset**



Note: N = 99. Two jurisdictions responded "NA" and 26 did not respond.

**Exhibit J-3. Parcel Data Fees, Restrictions, and Waivers**

Question	Responses					
	Yes (N)	%	No (N)	%	NA or Blank (N)	%
<b>Fees</b>						
Are there fees for the data?	45	35	58	46	24	19
If there are fees, can they be waived?	28	22	19	15	80	63
Can the fees be waived with an agreement?	14	11	18	14	95	75
Would there be reoccurring costs for obtaining the data in the future?	25	20	69	54	33	26
<b>Use restrictions and agreements</b>						
Are there use restrictions?	39	31	63	50	25	20
Do we need a data sharing agreement?	16	13	88	69	23	18
Is there a standard agreement for HUD to review?	13	10	90	71	24	19

Notes: N = 127 counties. "NA" responses and nonresponses ("Blank") are included above though, in sum, these ranged from 18 to 75% depending on the question. All questions were asked of each county regardless of response to a previous question.

**Exhibit J-4. Currency of Data**

Question	Responses										
	Current* (Fall 2011) (N)	%	As of Jan 2011 (N)	%	Before Jan 2011 (N)	%	NA or Blank (N)		%		
How current is the dataset?	52	41	10	8	6	5	59		46		
How often is it updated?	Annually (N)	%	Quarterly (N)	%	Monthly (N)	%	Daily (N)	%	NA or Blank (N)	%	
	23	18	4	3	15	12	28	22	57	45	
How often is it certified?	Annually (N)	%	NA or Blank (N)							%	
	70	55	57							45	

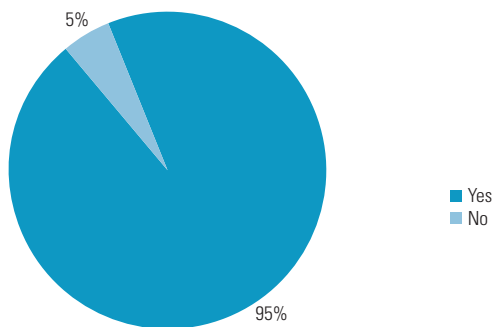
Notes: N = 127 counties. "NA" responses and nonresponses ("Blank") are included above though, in sum, these ranged from 18 to 75% depending on the question. All questions were asked of each county regardless of response to a previous question.

**Exhibit J-5. Tax Roll Profile**

Question	Responses													
	Q1 (N)	%	Q2 (N)	%	Q3 (N)	%	Q4 (N)	%	NA or Blank (N)		%			
In what quarter is the tax roll certified?	16	13	4	3	31	24	24	19	52		41			
Is a working roll available?	Yes (N)	%	No (N)	%	NA or Blank (N)							%		
	48	38	44	35	35							27		
Are there any data timeliness or completeness issues we should be aware of for the working roll?	Yes (N)	%	No (N)	%	NA or Blank (N)							%		
	20	16	65	51	42							33		
Assessment cycle (reappraisal)	Biannual (N)	%	Annual (N)	%	Random (N)	%	NA or Blank (N)					%		
	1	1	75	59	1	1	50					39		
Do you have historical records that you can provide?	Yes (N)	%	No (N)	%	NA or Blank (N)							%		
	50	39	67	53	10							8		
How many years of historical data can we obtain? (years)	1 (N)	%	2 (N)	%	3 (N)	%	4 (N)	%	5 (N)	%	> 5 (N)	%	NA or Blank (N)	%
	24	19	10	8	17	13	4	3	2	2	13	10	57	45

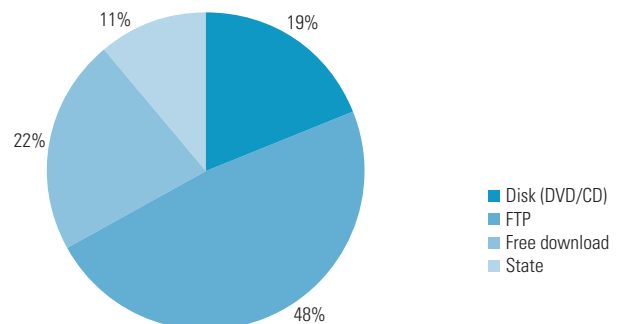
Notes: N = 127 counties. "NA" responses and nonresponses ("Blank") are included. Respondents answering "0 years" are included in "NA or Blank." All questions were asked of each county regardless of response to a previous question.

**Exhibit J-6. Prepackaged GIS Dataset**



Note: N = 97. Three jurisdictions responded "NA" and 27 did not respond.

**Exhibit J-7. Method for Accessing GIS Data**



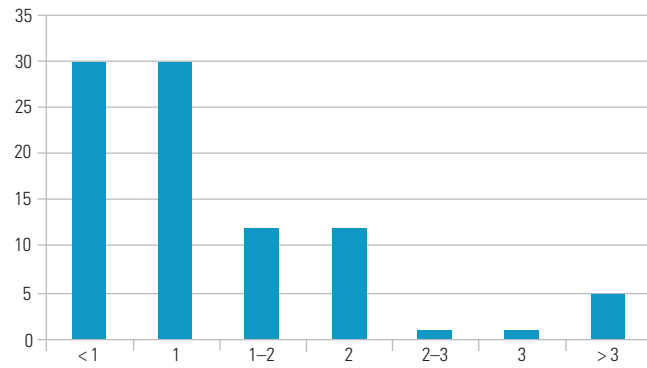
Note: N = 98. Four jurisdictions responded "NA" and 25 did not respond.

**Exhibit J-8. GIS Fees, Restrictions, and Waivers**

Question	Responses					
	Yes (N)	%	No (N)	%	NA or Blank (N)	%
<b>Fees</b>						
Are there fees for the GIS data?	38	30	60	47	29	23
Can the GIS fees be waived with an agreement?	21	17	19	15	87	68
<b>Use restrictions and agreements</b>						
Are there use restrictions?	34	27	63	50	30	23
Does the data sharing agreement for the tax roll also cover the GIS data?	21	17	19	15	92	73

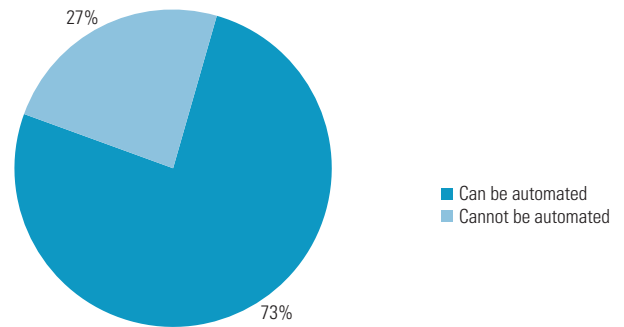
Notes: N = 127 counties. "NA" responses and nonresponses ("Blank") are included above. All questions were asked of each county regardless of response to a previous question.

**Exhibit J-9. Level of Effort for Collection (Hours)**



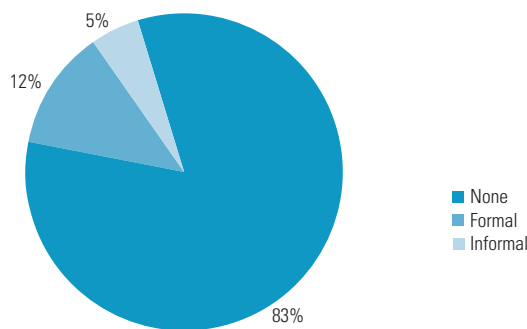
Notes: N = 91. Data for 14 jurisdictions were obtained through alternative sources (such as states), 20 were not collected, and 2 had extenuating collection at the time of feasibility analysis.

**Exhibit J-10. Future Automation Possibility**



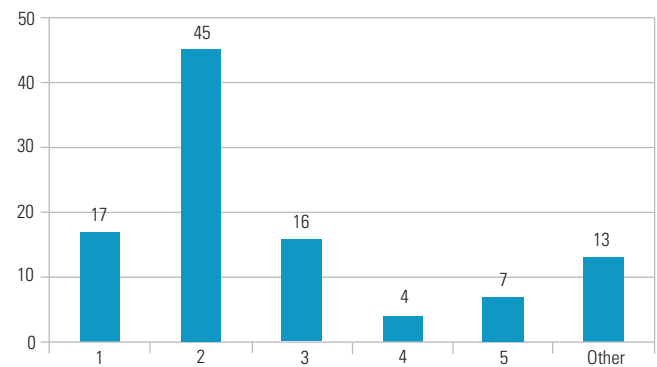
Notes: N = 94. Eight jurisdictions responded "NA" and 25 did not respond.

**Exhibit J-11. Data Sharing Agreement Type**



Notes: N = 76. Ten jurisdictions responded "NA" and 41 did not respond.

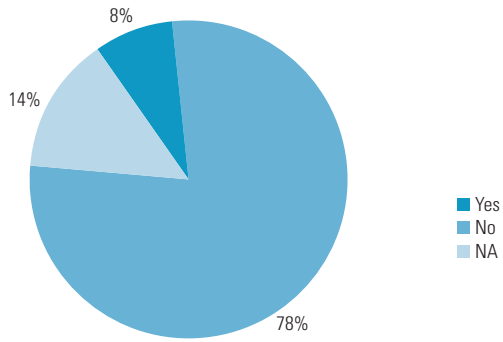
**Exhibit J-12. Level of Effort for Initial Contact (Number of Attempts)**



Notes: N = 102. Three jurisdictions responded "NA" and 22 did not respond.

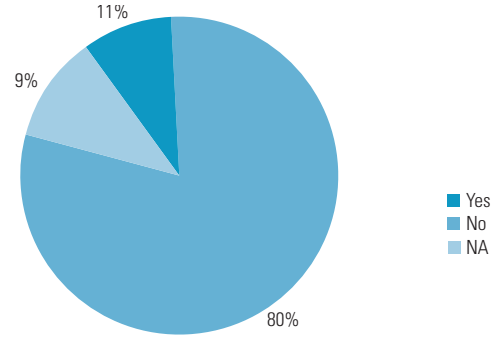


**Exhibit J-13. Future Plans for New Software**



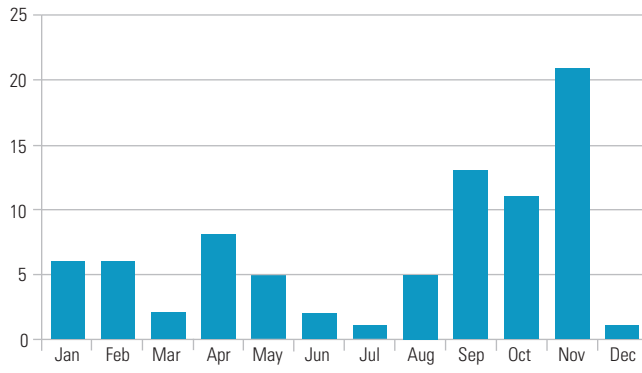
Notes: N = 91. Thirty-six jurisdictions did not respond.

**Exhibit J-14. Impending Retirements or Workforce Changes**



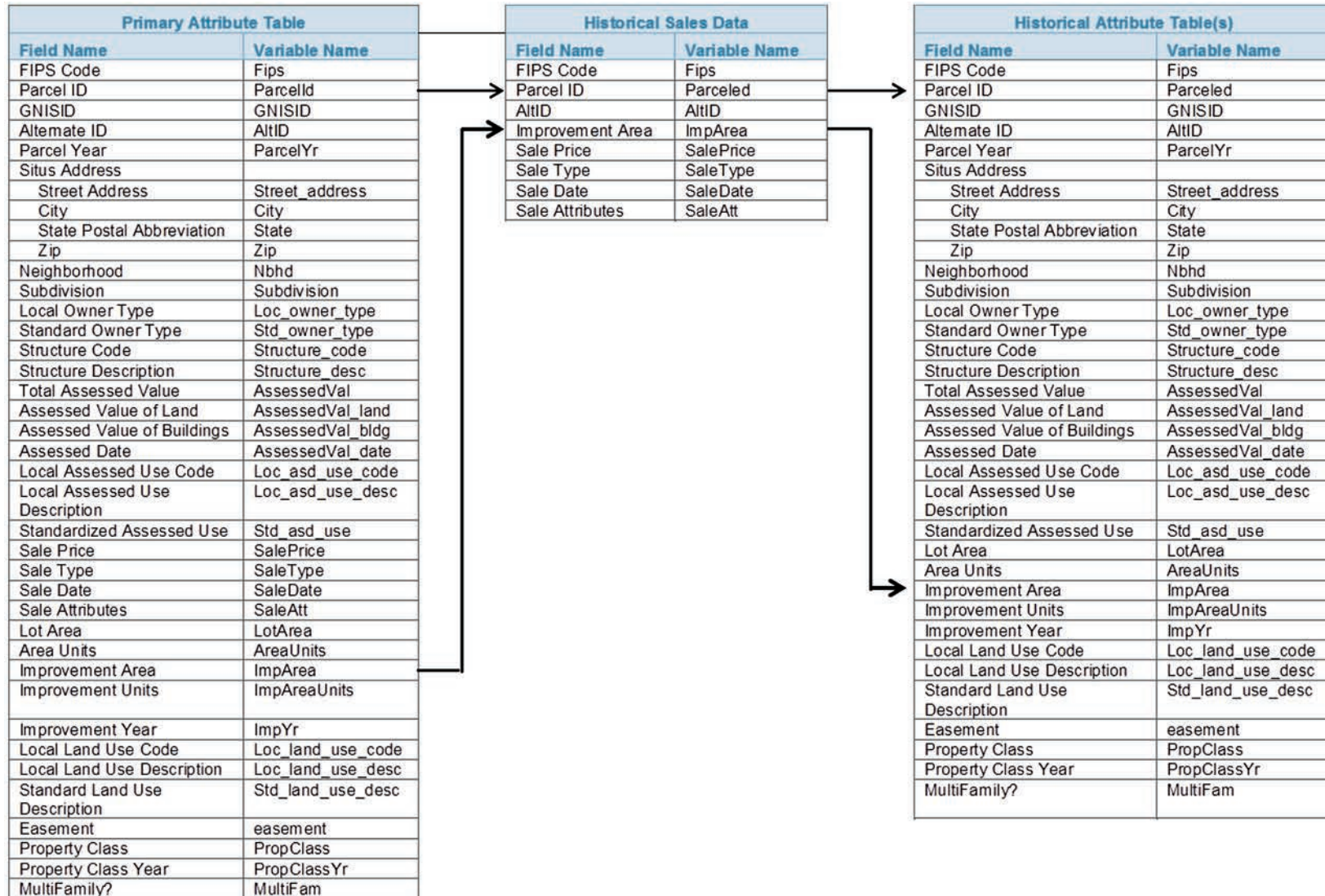
Notes: N = 92. Thirty-five jurisdictions did not respond.

**Exhibit J-15. Month of Tax Bill Mailing**



## Appendix K. Preliminary Standardized Database Shell and Format

Exhibit K-1. Preliminary Standardized Database Shell and Format



## References

- Federal Geographic Data Committee (FGDC) Subcommittee on Cadastral Data. 2009. "Land Parcel Data for the Mortgage Crisis: Results of the Stakeholders Meeting." Available at [http://www.nationalcad.org/data/documents/Land\\_parcel\\_data\\_for\\_the\\_mortgage\\_crisis\\_-\\_stakeholders\\_meeting\\_findings.pdf](http://www.nationalcad.org/data/documents/Land_parcel_data_for_the_mortgage_crisis_-_stakeholders_meeting_findings.pdf) (last accessed [date]).
- Folger, Peter. 2009. *Issues Regarding a National Land Parcel Database*. Congressional Research Service Paper 7-5700. Washington, DC: Congressional Research Service. Also available at <http://www.fas.org/sgp/crs/misc/R40717.pdf>.
- National Research Council (NRC). 1980. *Need for a Multipurpose Cadastre*. Washington, DC: National Academy of Sciences.
- . 1982. *Modernization of the Public Land Survey System*. Washington, DC: National Academy of Sciences.
- . 1983. *Procedures and Standards for a Multipurpose Cadastre*. Washington, DC: National Academy of Sciences.
- . 2003. *GIS for Housing and Urban Development*. Washington, DC: National Academy of Sciences.
- . 2007a. *National Land Parcel Data: A Vision for the Future*. Washington, DC: National Academy of Sciences.
- . 2007b. *Successful Response Starts With a Map: Improving Geospatial Support for Disaster Management*. Washington, DC: National Academy of Sciences.
- Sperling, Jon. 1995. "Development and Maintenance of the TIGER Database: Experiences in Spatial Data Sharing at the U.S. Bureau of the Census." In *Sharing Geographic Information*, edited by Harlan J. Onsrud and Gerard Rushton. New Brunswick, NJ: Center for Urban Policy Research: 377–396.
- Stage, David, and Nancy von Meyer. 2010. "An Assessment of Parcel Data in the U.S.: 2009 Survey Results." Available at [http://www.nationalcad.org/data/documents/2009\\_Assessment\\_of\\_Parcel\\_Data\\_in\\_US.pdf](http://www.nationalcad.org/data/documents/2009_Assessment_of_Parcel_Data_in_US.pdf) (last accessed January 15, 2012).
- Stage, David, Nancy von Meyer, and Bob Ader. 2005. "Parcel Data and Wildland Fire Management." Available at <http://www.nationalcad.org/data/documents/Parcel%20Data%20and%20Wildland%20Fire%20Management%20Final%20Oct%202005%20c.pdf> (last accessed January 15, 2012).
- State of Oregon. 2011 (September). Personal communication with Fairview Industries.