



Office of Financial Research
Working Paper #0012
December 5, 2013

Common Ground: The Need for a Universal Mortgage Loan Identifier

Matthew McCormick¹ and Lynn Calahan²

¹ Office of Financial Research, matthew.mccormick@treasury.gov

² Office of Financial Research, lynn.calahan@treasury.gov

The Office of Financial Research (OFR) Working Paper Series allows staff and their co-authors to disseminate preliminary research findings in a format intended to generate discussion and critical comments. Papers in the OFR Working Paper Series are works in progress and subject to revision.

Views and opinions expressed are those of the authors and do not necessarily represent the positions or policy of OFR, Treasury, or any of the agencies represented by contributors. Comments and suggestions are welcome and should be directed to the authors. OFR Working Papers may be quoted without additional permission.

www.treasury.gov/ofr

Common Ground: The Need for a Universal Mortgage Loan Identifier

Matthew McCormick¹ and Lynn Calahan²

¹Office of Financial Research, matthew.mccormick@treasury.gov

²Office of Financial Research, lynn.calahan@treasury.gov

Abstract

The U.S. mortgage finance system is a critical part of our nation's financial system, representing 70 percent of U.S. household liabilities. It is also highly complex, with many finance channels, participants, and regulators. The data produced by this system reflect that complexity; unfortunately, no single identifier exists to link the major loan-level mortgage datasets. The establishment of a single, cradle-to-grave, universal mortgage identifier that cannot be linked to individuals using publicly-available data would significantly benefit regulators and researchers by enabling better integration of the fragmented data produced by the U.S. mortgage finance system. Such an identifier could additionally serve as the foundation of a system that could benefit private market participants, as long as such a system protected individual privacy.

The authors gratefully acknowledge helpful comments from and discussion with Cornelius Crowley, William Nichols, and Linda Powell of the Office of Financial Research (OFR); Ren Essene and Ron Borzekowski of the Consumer Financial Protection Bureau; Theresa DiVenti, William Doerner, Robin Seiler, and Valerie Smith of the Federal Housing Finance Agency; Therese Scharlemann of the Department of the Treasury; Marc Rodriguez of the Board of Governors of the Federal Reserve System; Michael Nixon, Kurt Usowski, William Reeder, and Edward Szymanoski of the Department of Housing and Urban Development; Thomas Mayock of the Office of the Comptroller of the Currency; Harriet Newburger of the Federal Deposit Insurance Corporation; Rolaine Bancroft and Robert Errett of the Securities and Exchange Commission; and others. Any errors or omissions are the responsibility of the authors alone.

Views and opinions expressed are those of the authors and do not necessarily represent the positions or policy of OFR, Treasury, or any of the agencies represented by contributors. Comments and suggestions are welcome and should be directed to the authors.

1. Introduction

The recent housing crisis exposed a number of data gaps, risk-management failures, and shortcomings in operational controls throughout the mortgage finance system, including problems connecting origination data to performance data, tracking loan modifications, and verifying loan underwriting practices. Mitigating these problems is an important step toward returning the mortgage finance system to a healthy state. This paper examines the potential of the universal adoption of a single mortgage identification standard to help address these problems by enabling better integration of the fragmented data that is produced by the U.S. mortgage finance system.

The size (\$9.4 trillion outstanding as of Q1 2013), complexity, and fragmented nature of this system and its regulation are significant.¹ Given these factors, it is difficult to accurately identify patterns when identification requires connecting various sources of data without clear, consistent, and unambiguous identification of the key component: the mortgage itself. However, although a number of identification systems exist, no single universal identifier is shared across all government agencies or across all major private entities.² Although market participants cross-reference their identifiers as a matter of necessity, ambiguities and inconsistencies in definitions often make cross-referencing difficult and inaccurate. As with legal entity identifiers, “Simply put, having a multitude of identifiers adds layers of complexity, increases the potential for errors, and results in redundant efforts.”³

Shared identifiers are a public good that benefit market participants, regulators, and researchers. Were a universal mortgage identifier adopted, researchers and regulators would be better able to understand the mortgage finance system from a systemic risk perspective, and regulators would have a better understanding of the mortgage finance system from the perspective of compliance and prudential supervision. This improved understanding could yield better policy research, which could in turn drive better policymaking, thereby benefitting the public. Additionally, firms would benefit from improved risk management, lower costs of integration among trading partners and within large lending facilities, and lower compliance burdens from reporting requirements. Finally, contingent on the development of appropriate protections for individual privacy, purchasers of mortgage securities could potentially benefit from greater transparency.

The mortgage industry has been exploring the issue of unique loan identification, in part within voluntary consensus standards bodies, as defined under Office of Management and Budget (OMB) Circular A-119.⁴ One such organization is the Mortgage Industry Standards Maintenance Organization (MISMO), a standards-setting organization currently administered by the Mortgage Bankers Association. Many participants in MISMO have noted that a lack of a clear, open, and unifying system of identification, independent of any one private entity, has created problems for consumers, lenders, and others.^{5,6} The efforts of a MISMO working group have recently culminated

¹ Board of Governors of the Federal Reserve System, Z.1 Flow of Funds Accounts, June 6, 2013.

² An overview of existing and proposed mortgage identification regimes is provided in the appendix.

³ Bottega and Powell (2011).

⁴ Office of Management and Budget. Circular No. A-119 Revised. February 10, 1998.

⁵ Sokolowski (2012).

in the release of a white paper that discusses the utility of a unique loan identifier from the perspective of industry.⁷ Ideally, a standard developed by a voluntary consensus standards body would also meet the needs of government, including privacy requirements, which would allow for its adoption per OMB Circular A-119. Similarly, other private-sector efforts to create identifiers, such as the American Securitization Forum’s Loan Identification Number Code (ASF LINC), have grown out of a perceived need to “improve information flows to investors” in the wake of the recent housing crisis.⁸

Also since the crisis, the public sector and experts have noted the need for a unique mortgage loan identifier to enable cradle-to-grave monitoring of mortgages for financial stability purposes.⁹ Congress in the Dodd-Frank Wall Street Reform and Consumer Protection Act (Dodd-Frank Act) required regulators to write rules that involve identification of different types of entities and products, including mortgage loans. Specifically, Section 1094 of Dodd-Frank amended the Home Mortgage Disclosure Act (HMDA) to allow the newly created Consumer Financial Protection Bureau (CFPB) to mandate, “as [it] may determine to be appropriate, a universal loan identifier.”^{10,11}

The task of designing a universal identifier will be a difficult one. Careful consideration must be given to privacy concerns. As noted below, a mortgage identifier would have to be designed to prevent market participants from re-identifying individuals. No links from public documents to mortgage identifiers should be allowed. Otherwise the identifier could be used to identify individuals, rendering all datasets containing the identifier personally-identifiable information. Such a designation would create concerns about the use of individual data in the private sector and trigger burdensome requirements for government researchers using the data.

The creation of a universal identifier also raises tactical questions, such as the timing of assignment, and the structure and governance of any entities issuing identifiers or coordinating them. Other important issues include determining the parties that should have access to the identifier; the documents (if any) that must or should be allowed to carry the identifier; how to ensure use of the identifier at each step of the mortgage life cycle; how to ensure identifier integrity; and how to develop mechanisms to connect simultaneous or sequential liens taken out by a borrower on a particular property.

The next section of this paper provides a stylized overview of the mortgage finance system and the fragmented nature of its regulation, and discusses the major federal loan-level mortgage data collection efforts and disclosure requirements. Section 3 describes the goals that regulators and others hope to achieve by adopting a single common mortgage identifier. Section 4 describes the

⁶ Panchuk (2012).

⁷ MISMO (2013).

⁸ ASF (2009).

⁹ Wachter (2010).

¹⁰ Codified at 12 USC § 2803(b)(6)(G).

¹¹ The Dodd-Frank Act additionally provides for other unique identifiers, such as a SAFE Act identifier (Section 1094) and unique identifiers relating to loan brokers or originators (Section 942).

implications those goals have for the structure of an identifier and identification system, and Section 5 concludes.

2. The Supply Chain: The Residential Mortgage Life Cycle and Mortgage Data

The current state of residential mortgage finance in the United States is a result of an accretion of structures and regulations, designed to address different sets of problems, dating back to the Great Depression.¹² As a result, legal and regulatory authority is distributed among agencies at local, state, and federal levels, and there exist four major funding channels (bank portfolios and three securitization channels, discussed below), each with its own set of regulators. This fragmentation is reflected in the data produced by this collection of entities.

The Mortgage Life Cycle

To make clear the scope of the problem faced by regulators, academics, and market participants in developing and maintaining high-quality data, we begin with a stylized overview of the mortgage finance system and its many data-generation processes that are not linked by a common identifier.

Application to acceptance

For purchase loans, the mortgage supply chain begins with a home being made available for sale, typically through a state-licensed realtor, who may make the listing available on a regional Multiple Listing Service. A potential buyer then applies for a mortgage loan either through a mortgage broker or directly from a lender, both of which are licensed and regulated at the state level. Since 2008, the licensing of state-regulated entities has been coordinated by the Conference of State Bank Supervisors' Nationwide Mortgage Licensing System (NMLS).

Bank-chartered mortgage loan originators are additionally regulated by federal bank regulators and the CFPB and must be registered in the National Mortgage Licensing System and Registry, a modified version of NMLS, which Congress required to be created in 2008 to serve as a federal registry for mortgage loan originators, including state-regulated entities. The Dodd-Frank Act revised HMDA to require loan originators to submit information about mortgage applications to the CFPB or their primary federal regulator. Currently, the Federal Financial Institutions Examination Council collects this information on behalf of the federal regulatory agencies and contracts with the Federal Reserve Board (FRB) for the HMDA data collection.

When a potential borrower applies for a loan, the lender checks the applicant's credit with one or more consumer credit reporting agencies, which are primarily regulated by the CFPB and the Federal Trade Commission. An independent appraisal of the property is typically performed by a state-licensed appraiser, although automated valuation models are used in some cases, and another party performs the property inspection.¹³ At different stages during this process, the

¹² U.S. Department of Housing and Urban Development (2006).

¹³ The 2010 Interagency Appraisal and Evaluation Guidelines state that "an automated valuation model (AVM), in and of itself does not meet the Agencies' minimum appraisal standards" (Federal Register, Vol. 75, No. 237, p. 77453). An AVM can, however, be used as a part of an evaluation, which serves as a substitute for an appraisal, provided the transaction satisfies certain appraisal exemption requirements.

borrower is provided with required disclosures, including disclosures required under the Truth in Lending Act (TILA) and Real Estate Settlement Procedures Act (RESPA), which detail the terms of the mortgage and a good-faith estimate of closing costs.

Acceptance to origination

Parties to a mortgage officially enter into the contract at closing. At this time the title to the property is transferred via a deed, and two important documents are generated: the promissory note signed by the borrower and the mortgage. The mortgage, an instrument that varies in form and name from state to state, transfers an interest in the property to the lender. The title is a collection of rights to the property and the deed is a document describing the land and establishing the transfer of title.

With the exception of Hong Kong and parts of Ireland, the United States is alone among industrialized nations in not relying on property registration to establish ownership. Instead, this country uses a common-law system of chain of title and recordation of legal instruments affecting title.¹⁴ As a result, data on the ownership of real property and associated liens are maintained in local recorder offices. The difficulties that can arise in establishing a clear chain of title have resulted in the development of the title insurance industry, which does not exist in most of the developed world.

Beyond the mortgage, note, and title insurance, other documents are often generated: the property or mortgage may additionally be covered by flood insurance (provided through the Federal Emergency Management Agency's National Flood Insurance Program), private mortgage insurance, homeowner's (hazard) insurance, and mortgage life insurance. Each of the latter three insurers is subject to state regulation. At closing, all these origination documents are collected and collated by a closing agent, who ensures that the documents are in proper order and legally signed, and that funding and proceeds are distributed according to the terms within the documents.

Servicing

After origination and closing, the borrower's primary contact is with the mortgage servicer, which is responsible for collecting payments from borrowers, addressing delinquencies and loan modifications, and in some cases, conducting foreclosures. Mortgage servicing can be performed by the loan originator, but mortgage servicing rights and their associated cash flows are also often sold to other parties. Servicers are regulated at the state and federal levels, with additional monitoring often performed by entities that hold or guarantee the mortgages they service.¹⁵

Several types of servicers may be involved. There is generally a primary servicer, who is usually responsible for collecting payment from the borrower and remitting the proceeds to the appropriate parties (such as tax payments, or interest and principal if the servicer is not also the lender), and for conducting the foreclosure process when necessary. In many cases, these

¹⁴ Ireland is transitioning to a property registration system.

¹⁵ Servicers are regulated at the federal level by prudential regulators or the CFPB, depending on the institution's size and charter.

activities are contracted out to sub-servicers. Also, a master servicer often serves as an intermediary between primary servicers, and the trustee and issuer.

Loan disposition

Mortgages in the United States can either be held in portfolio by a bank, or securitized and sold through one of three major securitization channels.

Banks often use mortgages in their portfolios as collateral in order to borrow from one of the Federal Home Loan Banks.¹⁶ Information about loans in a bank portfolio is of interest to bank examiners.¹⁷

However, only about 10 percent of mortgages were held in bank portfolios in 2012.¹⁸ Most mortgages were instead sold as whole loans or securitized through one of the following three channels:

- Loans insured by the Department of Veterans Affairs, Federal Housing Administration, and United States Department of Agriculture are eligible to be pooled and sold as securities by private issuers with a guarantee from the Government National Mortgage Association (Ginnie Mae), a government corporation in the Department of Housing and Urban Development. These loans accounted for 20 percent of originations in 2012.¹⁹
- Loans that meet conforming loan limits and other requirements are eligible to be sold to the Federal National Mortgage Association or Federal Home Loan Mortgage Corporation (Fannie Mae and Freddie Mac, respectively). Fannie and Freddie pool mortgages, guarantee pools for a fee, and sell securities representing various aspects of those pools on the secondary market. They are regulated by the Federal Housing Finance Agency (FHFA). These loans accounted for 69 percent of originations in 2012.²⁰
- Loans can also be bundled and sold by private issuers without any government or agency guarantee. These securities are typically called private label securities (PLS). In 2012, PLS made up less than 1 percent of residential mortgage-backed securities issuances, compared to almost 36 percent in 2005 and 2006.²¹

In any of these three forms of securitization, loans are grouped into pools, which are held by trusts that in turn issue securities. The form of these trusts can vary significantly. Some, usually those guaranteed by Ginnie Mae, Fannie Mae, and Freddie Mac (together, the Agencies), though also some PLS, are simple pass-through structures wherein payments from the pools are simply

¹⁶ The FHLBs provide liquidity to the primary mortgage market and are overseen by the Federal Housing Finance Agency.

¹⁷ Regulation of banks and credit unions depends on their charter; they may be regulated by the Federal Deposit Insurance Corporation, the Federal Reserve System, the Office of the Comptroller of Currency, the National Credit Union Administration, and state bank regulators.

¹⁸ Inside Mortgage Finance, Issue 2013-4, February 1, 2013, p. 4.

¹⁹ Ibid.

²⁰ Ibid.

²¹ Ibid.

distributed pro-rata to the shareholders. More complicated structures enable the issuance of multiple securities with different exposures to risk from prepayment, and in the case of PLS, default. The Agencies and PLS issuers all have requirements (underwriting, operational, documentation, and reporting) of the lenders that sell mortgages to them or whose loans they guarantee.

The activities of the Agencies are determined by Congress in their charters, which limit the types and characteristics of loans that they can purchase (in the case of Fannie Mae and Freddie Mac) and guarantee. As a result, secondary market funding for mortgages that do not meet the criteria established for Agency loans can only be provided through issuance of PLS; PLS may also be issued when pricing in the market favors private-label execution over Agency execution. Although the market for new issuances not guaranteed by the Agencies has collapsed since the crisis, activity in this market may rebound in the future. All public offerings of PLS must be registered with the Securities and Exchange Commission (SEC), unless otherwise exempt from registration. If the offering is registered, the prospectus used to offer the PLS must comply with the SEC's initial and ongoing disclosure requirements, known as Regulation AB. Regulation AB contains the SEC's disclosure requirements for publicly offered asset-backed securities, including requirements for disclosure about the mortgages and other underlying assets. Additionally, ongoing disclosures for such publicly offered PLS, including the performance of the underlying assets, are required to be filed with the SEC.²² In 2010, the SEC proposed additional requirements to Regulation AB that would, among other things, require issuers to disclose asset-level information at the time of securitization and on an ongoing basis.²³

Loan termination and modification

Though some mortgage loans are paid down over the original term, most terminate either as a voluntary or involuntary prepayment, and loans are sometimes modified with the intention of preventing the latter.²⁴

The two most significant causes of voluntary prepayments are refinances, when the borrower takes out a new mortgage loan on the same property and uses it to pay off the outstanding loan, and the sale of a home.²⁵ When a refinance loan is originated, the identity of the loan being refinanced is not currently captured. Involuntary prepayments, in contrast, are generally the result of foreclosure or home-forfeiture actions, such as a short sale or deed-in-lieu of foreclosure.

²² Such disclosures must now be filed generally for the life of the security. See "Suspension of the Duty To File Reports for Classes of Asset-Backed Securities Under Section 15(d) of the Securities Exchange Act of 1934" SEC Release No. 34-65148 (August 27, 2011) [76 FR 52549]. Irrespective of the SEC disclosure requirements, some individual loan performance information may be available from private data vendors for a fee.

²³ In addition to asset-level disclosure requirements, the SEC also proposed significant revisions to the disclosure, offering process, and reporting requirements for ABS issuers. See "Asset-Backed Securities", SEC Release No. 33-9117 (April 7, 2010) [75 FR 23328].

²⁴ For example, the Securities Industry and Financial Markets Association's Mortgage Prepayment Projection Tables as of December 3, 2012, imply that half of a pool of conventional 30-year fixed-rate mortgages originated in 2011 with a coupon of 3.5 percent would be expected to prepay within 35 months given the yield curve as of that date. Alternatively, the historical assumption used by FHFA as to the average life of a mortgage loan is 10 years.

²⁵ Other causes, such as the death of an occupant covered by mortgage life insurance, or the borrower simply accelerating payments, are less common.

Mortgage modifications are changes to the original terms of a mortgage, typically to keep a loan in “current” status or restore it to that status. Although rare before the financial crisis, mortgage modifications have become relatively common for distressed properties.²⁶

Junior liens

Junior, or subordinate, liens are placed on properties when consumers borrow against additional equity in their homes without refinancing the existing mortgage. These liens are typically home equity loans, also known as “second mortgages,” which have amortization structures much like traditional first mortgages, or home equity lines of credit, which are revolving lines of credit secured by homes. The existence of second liens can be significant both for risk management purposes and when handling modifications to first liens.

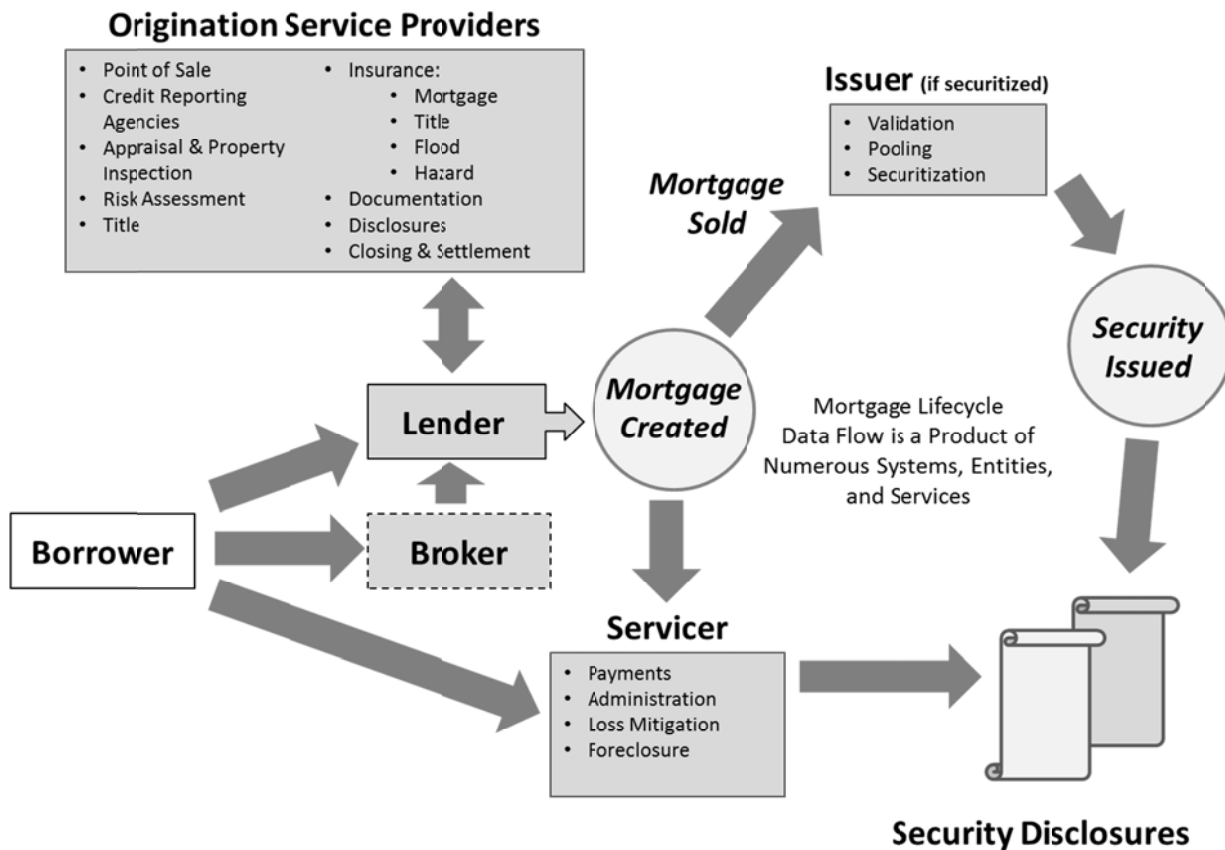
²⁶ For example, among loans covered by the OCC Mortgage Metrics report for the fourth quarter of 2012, 164,676 modifications and trials plans were initiated, compared to 156,773 foreclosures.

Illustration

In Figure 1 below, the arrows represent flows of information and documents. In the origination process, numerous service providers are necessary to fulfill the requirements of loan creation, each with its own systems, processes, and identification methods. The lender must integrate all the data generated by these disparate service providers to process a mortgage. Many of the identification numbers related to services rendered are lost through the process. Some of the remaining identification data are passed on if the loan is sold to an issuer; some are not. As a result, any entity trying to track down all of the underlying information known at the time of origination and connect it to performance data or a particular mortgage pool can be quickly stymied.

A consistent identification process would unite all of the pieces, producing efficiencies in data integration and service fulfillment.

Figure 1: Mortgage Life Cycle Illustration



Source: OFR analysis

Mortgage Datasets

Government agencies engage in a number of ongoing loan-level collections of mortgage data. A list of these datasets is in the appendix. These datasets are used for prudential supervision of regulated entities; to ensure compliance with regulations related to consumer protection, antidiscrimination, and affordable housing; to support the development and evaluation of housing policies; and to enable researchers and market participants to better understand the housing finance market.

These datasets contain significant overlap. A single mortgage could potentially appear in more than half the datasets listed in the appendix. Despite this overlap, no single mortgage identifier can link all of these datasets. As a result, information such as the property address or other borrower-specific information must be used in merging datasets. Even with this additional information, matching datasets is slower and more resource-intensive than it would be with a common identifier.

Aside from datasets maintained by the federal government and by private industry for internal business purposes (such as the portfolio information of banks and mortgage insurers), three additional sources of loan-level mortgage information are aggregated and used by private industry:

1. public records, including records of deeds, titles, mortgages, liens, and court filings;
2. servicing data, including details about mortgage origination and ongoing performance; and
3. credit histories, including information about borrowers' credit lines and mortgages, but lacking in detail about properties and mortgage terms.

3. Objectives of a Universal Mortgage ID

The establishment of a universal mortgage identifier (UMID) would have several major benefits. First, it would enable more effective supervision by regulators (prudential, consumer protection, and financial stability) through reliable dataset matching. Second, it would assist market participants with an industry-wide basis for workflow management and data transfer between firms, which could help in turn to improve risk management practices. This benefit would be particularly powerful if the industry established the UMID through a consensus standards body. Third, it would improve information flows to the investing public.

However, these goals must be balanced against a strong commitment to protect individual privacy and uphold relevant privacy laws, which must take precedence over realizing any particular benefit. For that reason, the implementation of a UMID must be handled with more care than in the case of identifiers designed to address non-personal information, like the Legal Entity Identifier (LEI).

Enabling more effective regulation, supervision, and policy research

Regulators would benefit from the ability to merge datasets faster and more accurately by aligning regulatory reporting IDs between agencies and with government-sponsored securitization

entities. In addition to the cost benefits of less resource-intensive dataset matching and the data quality improvements from more accurate matches, faster matches can be viewed as another potential improvement to data quality, because they provide researchers and policymakers with a better understanding of current market conditions.

Consumer compliance regulators would be able to link origination to servicing and performance data, allowing them to better understand industry practices, fair lending, the impact of disclosures on mortgage performance, and other areas. Further, if the UMID were associated with loan documents, consumer protection regulators would have an easier time assessing compliance with the Truth in Lending Act and the Real Estate Settlement Procedures Act.

From a systemic risk perspective, a UMID would be useful in tracking the propagation of risks through the financial system by helping link outstanding debt to the entities bearing potential losses due to risks related to credit, interest rates, and servicing. Furthermore, by better linking origination practices and parties to later outcomes, researchers could achieve a better understanding of how risks are created. Finally, the ability to easily identify and remove duplication in datasets that partially overlap would give regulators a better overall view of the health of the mortgage market.

A UMID would also provide significant benefits to policy researchers, by quickly and accurately linking the many datasets currently in use. By bringing origination data together with data on loan performance, refinances and modifications, and disposition, researchers will better understand the current and potential impacts of policies on borrowers and investors.

Finally, a UMID could potentially assist in other important issues, like matching first and second liens or linking purchase loans to later refinances to obtain a holistic view of patterns in borrowing, if augmented by appropriate reporting structures. However, a UMID by itself cannot provide a definitive solution to all identification issues in mortgage finance. For example, as liens are fundamentally related to real property, tasks such as lien matching may be more effectively addressed by considering the standardization of parcel identification, a matter outside the scope of this paper.

Providing a basis for document management and improved risk management

As illustrated in Section 2, the life cycle of a mortgage loan produces significant volumes of documentation and data. Managing this documentation can be a challenge, one made only more difficult when documents are transferred from an entity using one identification scheme to another entity with its own identification scheme. By associating all documents and data relevant to a loan with a UMID, document management and information interchange between firms and business units can be simplified, making regulatory compliance easier and lowering costs for market participants. Better management of data and documents also decreases operational risk by reducing the potential for error. Further, a UMID would present the opportunity to establish a system that supports document versioning, revisions in key documents, and the recording of events such as mortgage modifications.

To improve information interchange between firms and to improve risk management, a UMID should be structured as a true public good that becomes commonly used by market participants and available to all with no encumbrances on its use other than those necessary to protect borrower privacy. In this context, a UMID would become a building block of a more efficient housing finance system.

Protecting privacy

Significant consideration must be given to individual privacy, given the inherent risks of re-identification associated with any universal identifier system. These risks place critical constraints on the structure, governance, and use of a universal mortgage identifier system. Many definitions of privacy exist, most notably Brandeis's definition of it as "the right to be let alone" and Westin's definition of it as "the claim of an individual to determine what information about himself or herself should be known to others."^{27,28} The latter definition is of particular importance, as Westin's work in this arena influenced the passage of the Privacy Act of 1974, which, along with the Fair Credit Reporting Act, affords the most significant individual protections surrounding the collection, use, and transmission of personal information in the mortgage finance arena.

Mortgage transactions pose a potential privacy problem because investors in the secondary market have a legitimate reason to want to know private characteristics of borrowers, such as credit scores, but because the universe of potential investors consists effectively of the general public, it is essential that the individual borrowers not be identifiable. This problem aligns with the federal government concept of "personally-identifiable information" (PII), defined as "information which can be used to distinguish or trace an individual's identity, such as their name, social security number, biometric records, etc. alone, or when combined with other personal or identifying information which is linked or linkable to a specific individual, such as date and place of birth, mother's maiden name, etc."²⁹

The existence of a UMID can work either to protect personal privacy, by allowing direct matching of datasets for research and supervision without requiring that the individual be identified in an intermediate step, or to degrade privacy protections, if a UMID appeared both on a publicly filed mortgage note, as currently occurs with some common identifiers, and in commercially available servicing data. Additionally, if associated with an individual, a universal identifier would allow additional information to be associated with an individual than would otherwise be available, to the extent that any other information associated with that identifier were both available and not already linked with PII.

To address this issue, any potential UMID system must be designed to prevent the re-identification of individuals, particularly by preventing public disclosure of information linking the identifier to documents or datasets that identify borrowers by name or other identifying features, or that could be used in combination with other information to re-identify borrowers. Additionally, any reference data associated with the UMID would have to pertain only to the loan and not contain PII.

²⁷ Warren and Brandeis (1890).

²⁸ Westin (2003).

²⁹ Office of Management and Budget Memorandum M-07-1616.

Further, any connection between a public document and a mortgage identifier would be unacceptable, because the identifier could be directly traced to an individual's identity, making all datasets containing the identifier PII. Given that such a connection would place a significant additional burden on regulators and researchers in the federal government, protecting privacy is important in both its own right and as a matter of the practical utility of a UMID.

4. Properties of a Universal Mortgage ID

A universal mortgage ID should follow the best practices that have been developed in the establishment of other identifier systems. Although some specific issues surrounding a mortgage identifier differ from previous efforts, such as the LEI, many of the core issues remain the same:³⁰

1. *Scope of Coverage*

The identifier should ideally cover all single-family and multifamily residential mortgages. To be of the greatest use, it should be assigned as early as possible in the process, i.e., at application, regardless of whether the loan is approved.

2. *Structure of the Identifier*

- a. *One-to-one relationship to mortgages*

The identifier should be unique: one, and only one, mortgage should be assigned to any particular identifier. The converse also holds: one, and only one, identifier should be assigned to any particular mortgage. This property, a one-to-one mapping between identifiers and mortgages, allows for unambiguous identification of a particular mortgage.

- b. *Persistence*

The identifier should remain with the mortgage until the loan is terminated. Thus, it must persist over time regardless of the holder of the loan or any modifications made to it.

- c. *Extensibility*

For an identifier to be useful indefinitely, it must allow for growth in the number of identifiers issued, without having to reuse identifiers, which would violate the uniqueness criterion.

- d. *Neutrality*

Identifiers should be neutral: no information should be encoded in the identifier itself. This is important for persistence, extensibility, and in the case of mortgages, privacy. An identifier containing embedded information can be vulnerable to changes over time if that information changes (for example, an identifier that embeds a borrower's name), undermining persistence. Similarly, embedding information can limit the number of bits of information practically available, limiting extensibility — an identifier that embedded a ZIP code, for example, would only be useful so long as the ZIP code with the greatest number of mortgages over

³⁰ The properties noted here are influenced significantly by Bottega and Powell (previously referenced), as well as the CPSS-IOSCO "Report on OTC Derivatives Data Reporting and Aggregation Requirements" of January 2012, and the Commodity Futures Trading Commission's final rule, "Swap Data Recordkeeping and Reporting Requirements," published January 13, 2012.

time did not run out of identifiers. In addition, an identifier with embedded personal information, or other embedded information that could be used to re-identify individuals, would pose a threat to privacy.

e. Reliability

For a universal mortgage identifier to be adopted by market participants, its reliability would have to be ensured. The assignment mechanism must be robust, the identifier should not conflict with other systems that may be in use, and the assignment mechanism and any reference databases must be independent of any entity that could fail in the future.

f. Open Standard

The identifier should be based on an open, voluntary consensus standard.³¹

3. Public Availability

The use of the identifier must be free of any contractual restrictions. All parties in the mortgage market, regulatory agencies, and the research community should be able to use the identifier system freely. This does not imply that registration of an identifier must be free, or that information other than basic reference information must be freely available; it means only that registration of identifiers must be openly available and that use of the identifier system must be freely available, without licensing costs or restrictions.

4. Privacy Protection

Beyond simply not embedding information in the identifier itself, the identifier system must be designed to prevent re-identification, as discussed earlier.

5. Incentive Compatibility

To help ensure that market participants have an incentive to invest in maintaining a robust system of identification, they should benefit from using the identifier in their regular course of business. Agreement among regulators about using a single identifier for reporting can help encourage coordination among participants in the market.

6. Registration Process

To prevent disruption to market participants, the assignment process will need to work within the timelines associated with the mortgage application process.

7. Quality Assurance

Errors are often introduced in data in the normal course of business. To protect the integrity of a universal mortgage identifier, quality control practices, including best practices such as checksums and good governance practices, must be adopted, and clear responsibility for acquiring each new identifier must be established.

5. Conclusion

The U.S. mortgage finance system is a critical part of our nation's financial system, representing 70 percent of U.S. household liabilities.³² It is also highly complex, with many finance channels, participants, and regulators. The data produced by this system reflect that complexity; unfortunately, no single identifier exists to link the major loan-level mortgage datasets. The

³¹ Office of Management and Budget. Circular No. A-119 Revised. February 10, 1998.

³² Board of Governors of the Federal Reserve System, Z.1 Flow of Funds Accounts, June 6, 2013.

establishment of a single, cradle-to-grave, universal mortgage identifier designed to prevent identification of individuals would significantly benefit regulators and researchers. Such an identifier could also serve as the foundation of a system that could benefit private market participants, as long as such a system protected individual privacy. The establishment of such an identifier will be difficult. It will not be easy to balance competing demands for protecting privacy, while enabling better management of data and documents without generating unnecessary costs and burdens. However, it is a challenge worth meeting.

Appendix: Federal Datasets and Existing Loan Identifier Regimes

Federal Datasets

The major ongoing loan-level mortgage datasets maintained by federal agencies are listed below.

Agency	Dataset	Source	Description
CFPB (FFIEC)	HMDA	Originators	The HMDA dataset contains information on mortgage applicants' race, ethnicity, gender, and income; loan purpose, rate, lien status, and acceptance or denial; and property type, location, and occupancy.
FHA	FHA-Guaranteed Mortgage Portfolio		Includes property address, type, and value; loan purpose, UPB and terms; borrower income, DTI, and credit score; mortgage insurance status; and other data.
FHFA	Enterprise Affordable Housing Goals Dataset	Enterprise administrative data	This dataset contains loan numbers and census tract geographies for loans financed by Fannie Mae and Freddie Mac.
FHFA	Data for HPI Calculation	Administrative data from the Enterprises, FHA, and the Federal Home Loan Bank of New York	Data from all loans held in portfolio or securitized by Fannie Mae or Freddie Mac, including property-level address information, loan purpose, loan amount, loan-to-value ratio, and appraisal value. Data from FHA including property street address, loan type, and sales price information. Data from the Federal Home Loan Bank of New York including property street address, appraisal values, and sales price information for all active loans and certain historical loans that have collateralized FHL Bank of NY advances.
FHFA	Enterprise-Owned and Guaranteed Mortgage Portfolios	Administrative data from the Enterprises	Origination and servicing data for loans owned or guaranteed by Fannie Mae and Freddie Mac. Contains credit score, HMDA, and performance data, with origination data linked to ongoing performance data by loan number; does not contain address or other identifying information.
FHFA	Federal Home Loan Bank Acquired Member Assets	Servicer reports to Federal Home Loan Banks	Files include data from each loan purchased under the FHLBank AMA program along with semiannual progress reports. Data elements include loan characteristics at time of acquisition and updated loan status information.
FHFA	Monthly Survey of Rates and Terms of Conventional 1-Family Nonfarm Mortgage Loans (MIRS)	Sample of loan originators	Loans reported to the FHFA for the purpose of conducting the Monthly Interest Rate Survey (MIRS). This dataset contains data on loans closed during the last five business days of each month, including: property geography and purchase price, and loan terms. Lender ID and loan ID are reported to the FHFA.
FHFA/CFPB	National Mortgage Database	Enterprise administrative data, credit bureau data, and HMDA data, merged by Experian	A five percent nationally representative sample of first-lien single-family mortgages, based on credit report data and supplemented with information from other datasets. This dataset contains detailed information on property characteristics, geography, borrower demographics, and all of a borrower's credit lines.
FRB	FRY-14M	BHCs with \$50B or more in total consolidated assets, submitted through LPS	This monthly data collection includes origination and servicing data for loans owned by respondent banks. The data collection uses addresses for matching purposes.

Agency	Dataset	Source	Description
GNMA	GNMA MBS Pools		Origination and servicing data for loans securitized in GNMA MBS pools. These data include property street address, borrower information including credit score, and loan characteristics.
OCC	Large Bank Mortgage Metrics (MM)	Servicer reports to large banks	Retail mortgage-servicing information collected from large institutions, which includes 103 data elements for each mortgage loan; 76 data elements for each second-lien residential real estate loan; 12 data elements pertaining to property information (matching of first-lien and second-lien residential real estate loans); and 30 home equity portfolio (profit & loss) data elements.
SEC	MBS offering materials and periodic reports on EDGAR	MBS issuer reports	Issuers file offering materials that may contain loan-level information. Ongoing reports may include loan-level information about prepayments, defaults or modifications.
Treasury	HAMP	Servicers, submitted to Fannie Mae	In Supplemental Directive 09-01, the Treasury Department announced the eligibility, underwriting, and servicing requirements for HAMP. Under HAMP, servicers apply a uniform loan modification process to provide eligible borrowers with sustainable monthly payments for their first-lien mortgage loans. Pursuant to Supplemental Directive 09-01, servicers are required to periodically provide HAMP loan level data to Fannie Mae, as HAMP program administrator. The HAMP files contain data on modifications made and their subsequent performance, as well as data collected from borrowers for NPV calculations and whether or not a loan was approved for a modification.
VA	VA-Guaranteed Mortgage Portfolio		Includes property address, type, and value; loan purpose, UPB and terms; borrower income, DTI, and credit score; mortgage insurance status; and other data.

Existing Loan Identifier Regimes

Several loan-level unique identifier regimes currently exist. They are assigned or adopted by government agencies and mortgage lenders with varying degrees of market coverage and for varying purposes. These identifiers can be placed into three categories: regulatory reporting IDs, government business purpose IDs, and proprietary business purpose IDs.

Regulatory reporting IDs

HMDA requires the collection of loan-level application and origination data. Covered institutions annually report to their federal regulator several required fields, including an identifier up to 25 characters long of letters and numbers, for each loan or application. The identifier must be unique within the institution, and it is “strongly recommended” that names and social security numbers be excluded from the identifier.³³ Institutions have relative latitude to assign and report loan identifiers they may already be using for business purposes under HMDA. However, not all lenders comply with the recommendation, and some include borrower names or social security numbers in their IDs.

In 2010, the SEC proposed to require issuers of asset-backed securities to provide asset-level disclosures about the underlying loans backing securities. In providing the asset-level data, a unique number must be assigned to each asset in the underlying pool. As with HMDA, the proposed rules do not specify a numbering convention. The SEC noted that acceptable identifiers could have been generated at origination or at different times during the securitization process.³⁴ However, the proposed rules would require that the asset number used for offering disclosures should also be the same numbers used to identify the asset for all reports required of an issuer under the Exchange Act.

The Office of the Comptroller of the Currency (OCC) collects unique loan identifiers from the largest mortgage servicers for its quarterly Mortgage Metrics Report. The OCC defines the field as a “unique identifier for the loan record that will be the same month to month. Reference numbers may be used in lieu of actual loan numbers as long as it meets this criteria.”³⁵

The Federal Reserve Board collects unique loan identifiers as part of the FR Y-14M Capital Assessments and Stress Testing report, which must be filed by bank holding companies with more than \$50 billion in consolidated assets. The identifier collected by the FRB is an alpha-numeric

³³ See HMDA 2010 Guide to Getting it Right. <http://www.ffiec.gov/hmda/guide.htm>.

³⁴ See Asset-Backed Securities, SEC Release No. 33-9117 (April 7, 2010) [75 FR 23328].

³⁵ <http://www.occ.treas.gov/publications/publications-by-type/other-publications-reports/mortgage-metrics-q4-2008/loan-level-data-field-defin-q4-2008.pdf>

code of up to 32 characters.³⁶ The FRB provides the following instructions in the FR Y-14M for the loan number field:³⁷

An identifier for a loan that will be the same from month to month. Reference numbers may be used in lieu of actual loan numbers as long as it meets these criteria. This loan identifier must uniquely identify any loan in the file. It must identify the loan for its entire life and must be unique (piggy-backs should be separated).

Government business purpose IDs

Several other government agencies with missions relating to housing finance use identifiers for internal data collection and maintenance. The Federal Housing Administration (FHA), the U.S. Department of Veterans Affairs (VA), and the U.S. Department of Agriculture (USDA) have loan-level identifiers for loans guaranteed or insured by the agencies. The FHA case number is generated on the FHA Connection website after input and verification of data fields containing borrower and property information. It is a 10-digit identifier assigned to a mortgage as the first step toward FHA endorsement of mortgage insurance. The VA assigns a 12-digit loan identifier number at the time the appraisal is requested, and it is used in subsequent entries into VA systems and on various documents. Ginnie Mae assigns a 9-digit loan ID.

Proprietary business purpose IDs

In addition to unique IDs that individual entities or loan origination systems may use, there are two broadly available proprietary loan identifiers: (1) the Mortgage Electronic Registration System (MERS) Mortgage Identification Number (MIN) and (2) the American Securitization Forum (ASF) LINC (loan identification number code). Individual business needs drive their use in mortgage origination, servicing, and securitization. The MERS MIN is an 18-digit loan identifier required for loans registered in the proprietary MERS system. The primary purpose of the system is to track ownership interest in registered mortgages. The number itself may be auto-generated by software or assigned directly by MERS. A portion of the number can be either randomly generated or created from an existing loan number created by the lending institution; the string would then be wrapped by a unique MERS residential organizational ID and a check digit. Once the MIN is assigned, it does not change for the life of the loan. Currently, Fannie Mae and Freddie Mac require all e-mortgages to be delivered registered with a MERS MIN, contributing to adoption rates among originators in the market. There are membership fees associated with registering loans on the MERS system.

The LINC is a 16-digit loan identifier for mortgages, auto loans, credit cards, and student loans, developed by the ASF and Standard & Poor's Fixed Income Risk Management Services (FIRMS). The LINC was developed in coordination with ASF's Project RESTART, a loan-level disclosure

³⁶ If the bank holding company (BHC) is already submitting data to the OCC as part of the OCC Mortgage Metrics Data or OCC Home Equity Data, it is required that the BHC use the same loan number for the FR Y-14M data schedules. The FR Y-14M data population may include additional loans, which may not be part of the OCC data sample, and for such loans the general requirements listed in the schedule instructions will be applicable. Overall, the entire data file sent by a BHC should have unique loan numbers across the entire submission.

³⁷ Report forms can be found at: http://www.federalreserve.gov/reportforms/forms/FR_Y-14M20130331_f.zip

initiative for the residential mortgage-backed securities industry. ASF assigns a unique number after receiving 30 data fields from the originator or servicer. LINC includes embedded data, such as underlying loan type, loan origination date, and country code.³⁸ To promote integration of the assignment process, S&P FIRMS developed an automated request module to fit into existing platforms. The LINC is linked to loan-level information in the centralized data repository run by S&P FIRMS. LINC is available to originators and securitizers free of charge, but adoption seems to be dependent upon new regulations and its use is not currently general practice.

In September 2012, MISMO, a not-for-profit subsidiary of the Mortgage Bankers Association focused on technology standards development for the mortgage industry, created a Unique Loan Identification Data Working Group (DWG) with the following statement:³⁹

There is not a single Unique Loan Identification Number (ULIN) of a mortgage loan in use uniformly throughout the industry over the life of the loan, and many in the industry use individual methods of identifying loans at various stages in the loan's life cycle. This proposed DWG will explore the development of a universally accepted ULIN. The group will examine current mortgage identifier systems and the potential for standardization, and will also review other non-mortgage banking industry efforts such as the definition of the ISO [International Organization for Standardization] "bank card" standard for credit and debit cards.⁴⁰

³⁸ http://www.americansecuritization.com/uploadedFiles/ASF_LINC.pdf

³⁹ See discussion on page 2.

⁴⁰ http://www.mismo.org/files/PressReleases/NewDWGPressRelease9_14_2012.pdf

References

American Securitization Forum. "ASF Loan Identification Number Code (ASF LINC™) for Securitized Loans Developed by ASF and Standard & Poor's FIRMS." Press release. September 24, 2009. <http://www.sifma.org/news/news.aspx?id=13216>

Bottega, John, and Linda Powell. "Creating a Linchpin for Financial Data: Toward a Universal Legal Entity Identifier." Federal Reserve Board Finance and Economics Discussion Series, 2011-07.

Mortgage Industry Standards Maintenance Organization (MISMO). "Unique Loan Identifier Development Workgroup White Paper." September 25, 2013.

Panchuk, Kerri Ann. "MISMO developing universal tracking system for mortgages, properties." HousingWire. Published 9/12/2012. Accessed 3/28/2012. <http://www.housingwire.com/fastnews/2012/09/14/mismo-developing-universal-tracking-system-mortgages-properties>.

Sokolowski, Rachael. "Unique Loan Identifiers." Presentation at MISMO forum on 9/26/2012.

U.S. Department of Housing and Urban Development. "Evolution of the U.S. Housing Finance System: A Historical Survey and Lessons for Emerging Mortgage Markets." April, 2006.

Wachter, Susan M. Written testimony for hearing titled "Regulation C: Home Mortgage Disclosure Act (HMDA) Public Hearing." July 15, 2010. Atlanta, Georgia. <http://www.federalreserve.gov/communitydev/files/wachter.pdf>

Warren, Samuel and Louis Brandeis. "The Right to Privacy." Harvard Law Review, Vol. 4, No. 5, 1890, pp. 193-220.

Westin, Alan. "Social and Political Dimensions of Privacy." Journal of Social Issues, Vol. 59, No. 2, 2003, pp. 431-453.