



# TECHNICAL REPORT 1

## National Mortgage Database Technical Documentation

2020  
March 10



## Table of Contents

1.	Introduction.....	1
2.	The Experian Contract .....	3
3.	Selecting the Initial Sample .....	4
4.	Processing the Initial Sample.....	6
5.	Updating the Sample.....	8
6.	Merging with other Data Sources .....	10
7.	Production.....	13
8.	Evaluating the NMDB Sample Frame .....	15
	Appendix A. Origins of NMDB.....	A-1
	Appendix B. Background on Mortgage Delinquency Reporting.....	B-1



## 1. Introduction

The National Mortgage Database (NMDB<sup>®</sup>) program is jointly funded and managed by the Federal Housing Finance Agency (FHFA) and the Consumer Financial Protection Bureau (CFPB). The program is designed to provide a rich source of information about the U.S. mortgage market based on a five percent sample of residential mortgages. It has three primary components:

- (1) the National Mortgage Database (NMDB);
- (2) the National Survey of Mortgage Originations (NSMO); and
- (3) the American Survey of Mortgage Borrowers (ASMB).

The NMDB program enables FHFA to meet the statutory requirements of section 1324(c) of the Federal Housing Enterprises Financial Safety and Soundness Act of 1992, as amended by the Housing and Economic Recovery Act of 2008 (HERA).<sup>1</sup> Specifically, FHFA must, through a monthly survey of the mortgage market, collect data on the characteristics of individual mortgages including those eligible for purchase by Fannie Mae and Freddie Mac and those that are not, and including subprime and nontraditional mortgages. In addition, FHFA must collect information on the creditworthiness of borrowers, including a determination of whether subprime and nontraditional borrowers would have qualified for prime lending.<sup>2</sup>

For CFPB, the NMDB program supports policymaking and research efforts, and helps identify and understand emerging mortgage and housing market trends. CFPB uses NMDB, among other purposes, in support of the market monitoring called for by the Dodd-Frank Wall Street Reform and Consumer Protection Act (Dodd-Frank Act), including understanding how mortgage debt affects consumers and for retrospective rule review required by this statute.<sup>3</sup>

In seeking to meet these objectives, FHFA and CFPB considered using existing databases but determined that none was sufficient and that a new database, NMDB, had to be created.<sup>4</sup> NMDB is a de-identified loan-level database of closed-end first-lien residential mortgages. NMDB has the following features:

- (1) it is representative of the market as a whole;
- (2) it contains detailed, loan-level information on the terms and performance of mortgages, as well as characteristics of the associated borrowers and properties;
- (3) it is continually updated;
- (4) it has a historical component dating back before the financial crisis of 2008; and
- (5) it provides a sampling frame for NSMO and ASMB.<sup>5</sup>

The core data in NMDB represent a statistically valid 1-in-20 random sample of all closed-end first-lien mortgages in the files of Experian, one of the three national credit bureaus.<sup>6</sup> When the

---

<sup>1</sup> Housing and Economic Recovery Act of 2008, Pub. L. 110–289, 122 Stat. 2654 (2008).

<sup>2</sup> FHFA interprets the NMDB program, including NSMO, as the “survey” required by the Safety and Soundness Act. The statutory requirement is for a monthly survey. Core inputs to NMDB, such as a regular refresh of credit-repository data, occur monthly, though NSMO is conducted quarterly.

<sup>3</sup> Dodd-Frank Wall Street Reform and Consumer Protection Act, Pub. L. 111-203, 124 Stat. 1376 (2010).

<sup>4</sup> Please see the Appendix A for a discussion of sources available at the genesis of the program and their limitations.

<sup>5</sup> See NSMO Technical Documentation at <http://www.fhfa.gov/nmdb>.

<sup>6</sup> Experian was chosen through a competitive procurement process to assist in creating NMDB.

NMDB program began, an initial sample was drawn from all mortgage files outstanding at any point from January 1998 through June 2012. Since then the sample has been updated on a quarterly basis with mortgages newly reported to Experian. Mortgages are tracked in NMDB from at least one year prior to origination to one year after termination of the mortgage, whether that termination is through prepayment, adverse termination, or maturity.

The use of a sampling frame substantially reduces the privacy risk associated with any data collection. By contrast, a universal registry can present challenges for privacy since it is known that a particular loan must be in the dataset. However, for a 1-in-20 sample, the odds are 95 out of 100 that a particular loan is not in the database. In addition, the sample used is large enough to support almost all types of statistically valid analyses but small enough to manage logistically, thus dramatically reducing both contract and personnel costs. The restriction of the NMDB frame to closed-end loans was made for two reasons. First, it mimics the reporting requirement of the Home Mortgage Disclosure Act (HMDA) and second, it reflects the practical fact that administrative data, which is a critical input for the NMDB data, is available for very few open-ended loans.

A random 1-in-20 sample of mortgages newly reported to Experian is added each quarter. Information from credit repository files on each borrower associated with the mortgages in the NMDB sample is collected from at least one year prior to origination to one year after termination of the mortgage. The information on borrowers and loans available to the FHFA, CFPB, or any other authorized user of the NMDB data is de-identified and does not include any directly identifying information such as borrower name, address, or Social Security number.

This technical report is designed to provide users of the NMDB data with background on the development of the database, as well as an assessment of the quality of its data. The remaining sections of this report discuss the development of the contract with Experian, outline the process of selecting the initial historical sample, describe how the initial sample data were processed, discuss how the data are being updated, how administrative data are being merged into the NMDB, and the details of the current production version of the database, NMDB 9.0. The final section then evaluates the NMDB sample frame.

## 2. The Experian Contract

By interagency agreement between FHFA and CFPB, FHFA leads the production of the NMDB. Following a competitive procurement process, a five-year contract for the core data of the NMDB was signed between FHFA and Experian in September 2012.<sup>7</sup> Simultaneously, FHFA and CFPB signed an interagency agreement that codified the cost-sharing (shared equally) and administrative arrangement.

The Experian contract has several key elements designed to ensure compliance with the Fair Credit Reporting Act (FCRA) and to protect the privacy of both borrowers and lenders.<sup>8</sup> First, while Experian uses name, address and Social Security number for matching purposes only, this information will not be transmitted to FHFA or CFPB when constructing the NMDB. Second, any user of the database must sign a terms of use agreement that states that they will not attempt to learn the identity of any borrower.<sup>9</sup> Third, all access to the NMDB must be through a server at FHFA or CFPB and strictly controlled. Fourth, the NMDB – which is a sample and designed to describe the market as a whole – cannot be used for enforcement against any specific servicer or lender.

---

<sup>7</sup> A 10-year extension of this contract was signed in September 2017.

<sup>8</sup> The Fair Credit Reporting Act (FCRA), Public Law No. 91-508, was enacted in 1970, and substantially amended since, to promote accuracy, fairness, and the privacy of personal information assembled by credit reporting agencies (CRAs). The Act's primary protection requires that CRAs follow "reasonable procedures" to protect the confidentiality, accuracy, and relevance of credit information. To do so, the FCRA establishes a framework of requirements for credit report information that include rights of data quality (right to access and correct), data security, use limitations, requirements for data destruction, notice, user participation (consent), and accountability.

<sup>9</sup> The Experian contract allows access to the NMDB to be extended to employees of other federal agencies, the Federal Reserve System, Fannie Mae, Freddie Mac, and Federal Home Loan Banks, provided the employee has signed the terms of use agreement. At present, employees from Fannie Mae, Freddie Mac and nine Federal agencies have been granted access.

### 3. Selecting the Initial Sample

The credit repository core of the NMDB was developed in two phases: (1) an initial 1-in-20 random sample of closed-end first-lien mortgages active at any time from January 1998 to June 2012 (January 1998 was the earliest available date given Experian's archive policies); and (2) quarterly updates that add a 1-in-20 random sample of mortgages newly reported to Experian and updated information on existing loans still active in the database.

One of the virtues of the credit repository sampling frame is that the repositories maintain records in a credit report not only of mortgages (and other credit obligations) that are currently active, but also of those that are closed. However, because of FCRA, records with derogatory information are purged from the current credit report after seven years from their point of first continual delinquency, and Experian's policies dictate a purge of all closed accounts 10 years after their closing.

However, since Experian retains archives of their data for 10 years or longer, data on mortgages that have been purged from Experian's current files can be recovered. These archives, which are not used for credit granting decisions, contain snapshots of each credit record as it existed at the close of business on a given day of each month, except that personal information (such as name, address, and Social Security number) is suppressed.

The bulk of the initial sample for the NMDB was drawn from the June 2012 archive. This was supplemented by samples from the December 2005 and July 2001 archives that captured loans that may have been purged from the current files by June 2012.

Trade lines, which are records that contain information about specific loans or debt obligations that are reported by loan servicers, account for most of the information contained in credit records. Loan servicers typically update trade line information on a monthly basis using a standardized format agreed upon by the servicers and the credit repositories (the Metro 2<sup>®</sup> format, introduced in 1997 and made mandatory in 2018). The updates include information on the opening date of the loan, the current and original loan balance, the type of servicer, loan term and type, payment amount, and loan repayment performance.

However, the format agreed upon by loan servicers and the credit repositories does not perfectly identify closed-end first-lien mortgages. Recognizing that some second liens would be sampled and have to be removed later, trade lines falling under the following categories were deemed eligible for the NMDB:

- any trade line with a Metro 2 “Enhanced Account Type Code” of: 08 (Real estate loan, specific type unknown), 19 (FHA real estate mortgage), 2C (FMHA real estate mortgage), 25 (VA real estate mortgage), 26 (Conventional real estate mortgage), 27 (Real estate mortgage, with or without collateral, usually second mortgage), 85 (Bi-monthly mortgage payment), 87 (Semi-monthly mortgage payment), 5A (Real estate – junior liens and non-purchase money first), 17 (Manufactured home loan), and 05 (FHA home-improvement loan); or
- trade lines reported by servicers with “Kind of Business Codes” of: FB (Mortgage



Brokers), FM (Mortgage Companies), FR (Mortgage Reporters), RE (Real Estate Sales and Rentals), BM (Bank-mortgage only), FL (Savings and loan – mortgage department) **and** Metro 2 “Enhanced Account Type Codes” of: 02 (Secured loan), 04 (Home improvement loan), 66 (Government- secured guaranteed loan), 7B (Agriculture), 9A (Secured home improvement) **or** a “Secondary Agency Code” of: 01 (Fannie Mae) or 02 (Freddie Mac).

Trade lines in the June 2012 archive that met either of the above criteria were included in the population from which the initial NMDB 1-in-20 random sample of mortgages was drawn. Any open-ended or revolving loans otherwise meeting one of the criteria were excluded from the sampling universe. No other restrictions were imposed.

The first supplemental sample was a 1-in-20 random sample of trade lines drawn from the December 2005 archive that met the criteria for the June 2012 archive. It had information reported for some period in the past 7 years (indicated by an “Account Balance Date” of January 1998 or later) and contained loans that were opened in September 2005 or earlier. In order to exclude loans from the 2005 sample that should be present in the June 2012 archive, loans were excluded if they were last reported after July 2002 with a reported account status of “current.”

The second supplemental sample, drawn from the July 2001 archive, was a random 1-in-20 sample of trade lines that met the criteria used for the June 2012 archive and that had “Account Balance Dates” of January 1998 or later and “Account Open Dates” of April 1999 or earlier. Any trade line with an “Enhanced Status Code” of “current” was excluded from the sample. Again, these additional conditions were designed to exclude all trade lines from the 2001 sample that should be present in the 2005 archives.

## 4. Processing the Initial Sample

For each archival pull, all available individual depersonalized credit records, including trade lines, inquiries, and public records (collectively, TIPs) associated with all borrowers accompanying any initial sample trade line were provided regardless of the archive from which it was sampled. The data provided by Experian are de-identified and contain no directly identifying personal information such as name, address, or Social Security number. The credit records were tagged with de-identified borrower (PINs), servicer and loan numbers (all in encrypted form).<sup>10</sup> These could be used (imperfectly) to link TIP files to other account-level files both within an archive and over time.

One major problem encountered with the NMDB sample frame is that a single mortgage can be associated with multiple trade lines. This can arise when the servicing of the loan is sold or transferred, and the trade line reported by the original servicer is not properly linked to the trade line reported by the new servicer. In such cases, borrowers may appear to have multiple mortgages, when, in fact, they have only one. Because of these duplicates, randomly sampling trade lines will result in mortgages with multiple records being overrepresented in the data. To correct for this, a processing methodology was developed to identify and combine multiple records that contain information about the same mortgage into one record.

The first step in the process of eliminating duplicate mortgage records (“de-duping”) was to find multiple trade lines for the same mortgage in the same archive. From these duplicates, sample loans were removed when the selected trade line was not the one with the latest “Account Balance Date” (this corrects for the problem of having mortgages associated with multiple trade lines over-represented in the sample). The second step was de-duping across archives. The June 2012, December 2005, and July 2001 samples were treated as sequential NMDB sample frames (in that order) whereby mortgages selected from a NMDB sample frame later in the order (*e.g.*, July 2001) that can be found in a NMDB sample frame earlier in the order (June 2012 or December 2005) would be removed from the sample (again, this corrects for the fact that such mortgages are over-sampled in the raw frame).

The de-duping process also dealt with the problem of ambiguous lien status for the “Enhanced Account Type Codes” of 08 (Real estate, specific type unknown), 27 (Real estate mortgage, with or without collateral, usually second mortgage), and 5A (real estate – junior liens and non-purchase money first). Sample trade lines associated with these codes were removed from the sample when they subsequently could be linked with trade lines that were unambiguously second liens.

Once the initial samples were de-duped, it was necessary to link archival records over time to create a composite picture of each sample loan (this is particularly important for loan

---

<sup>10</sup> The encrypted servicer identification and loan numbers are unique to the NMDB and are used by the NMDB development team primarily to update the database each quarter. They are not available to dataset users even in encrypted form. This is done to ensure compliance with the contract restriction that the database not be used for enforcement against servicers. The borrower PINs are also unique to the NMDB and are randomized. Experian, however, maintains the mapping between the borrower (and servicer and loan) identification numbers used in their system and the PINs supplied to the NMDB team so that records in the NMDB associated with the same PIN will be associated with the same borrower ID in the Experian records.

performance as described in Section 7 and Appendix B). Semi-annual archives were drawn for the period December 2001 to December 2011 for borrowers associated with the initial sample loans. Data from these archives were patched together to create a temporal picture of each loan. One issue that needed to be dealt with was that PINs for a given borrower can change over time. There are times when a loan is first reported to the credit repositories and cannot be connected with existing credit records for the borrower(s). This can happen because lenders make errors in reporting names and addresses or because of changes to a borrower’s addresses or names. In this instance Experian treats the loan as associated with a new borrower. In most of these instances the records are ultimately reconciled with the correct existing borrower and a “PIN-merge” occurs. However, historical archives are stored with the PINs at the time of the archive. Thus, to properly connect borrowers (and mortgages) over time, it was necessary for Experian to provide a PIN-merge transformation table to map historical to current PINs.

As shown in Table 1, the de-duping process substantially reduced the size of the original NMDB sample. About 16 percent of the mortgage trade lines originally sampled from the June 2012 archive, 30 percent of the selections from the 2005 archive, and almost three-quarters of the selections from the 2001 archive were dropped.<sup>11</sup> The percentages were higher for the older archives since many of the loans selected from them were selected because they were not current at the date of the archive and thus subject to FCRA purge rules. However, many of these loans subsequently became current and could be found in later archives.

<b>Archive Date</b>	<b>Sample Tradelines</b>	<b>Final Loans</b>	<b>Final Borrowers</b>	<b>Percentage of Tradelines Dropped</b>
Jul 2001	302,398	79,764	123,289	73.6
Dec 2005	2,955,675	2,059,855	3,365,504	30.3
Jun 2012	9,225,304	7,703,538	12,028,808	16.5

---

<sup>11</sup> A small percentage of the trade lines (1-2 percent) were dropped for other reasons. This included: (1) loans with no balance or terms information; (2) those in American territories other than Puerto Rico, Guam or the Virgin Islands; (3) “frag files” missing information on all other consumer obligations; and (4) those with no information about the borrowers.

## 5. Updating the Sample

Under the NMDB sample design, credit records for borrowers associated with sampled mortgages are to be collected quarterly until one year after the mortgage is reported as closed.<sup>12</sup> As of June 2012, approximately 3 million loans from the initial sample were still active or had been closed less than a year. In addition, to keep the NMDB up-to-date, it was necessary to add a representative sample of the new mortgages reported to Experian each quarter to the database from June 2012 onward.

The initial update of the NMDB from the June 2013 archive covered a full year of newly-reported mortgages since June of 2012. Since that date, updates have taken place quarterly drawing from the last archive of the quarter (March, June, September or December). Each quarterly update follows the same pattern. A 1-in-20 random sample of closed-end first-lien mortgage trade lines is drawn. These loans, which are identified using the same criteria as was used for the June 2012 archive, are selected from among the loans that were newly reported to Experian since the date of the previous quarterly update archive. The new sample is de-duped using the same methodology as used for the initial sample. If multiple trade lines are identified for the mortgage and the selected mortgage is not the one with the latest “Account Balance Date” or the mortgage is deemed to be a second lien or to unacceptable for the other reasons cited earlier then it is dropped. In addition, checks are run to determine if the mortgage was already reported in an earlier archive period (perhaps as a different trade line). If so, the loan is dropped.

Existing sample loans are also updated each quarter. Prior to the update, the PIN-merge transformation table is updated to account for “newly merged” PINs. To ensure that lagged information for all PINs newly added to the dataset is collected, the year-old archive is drawn each quarter for all active PINs for which this archive had not previously been collected. At present, an average of 85,000 new loans are added to the NMDB each quarter (see Table 2). The number of mortgages added to the database is only about two-thirds of the raw trade lines originally selected for the update sample.

---

<sup>12</sup> A partial update is done monthly collecting only limited performance data for active sample mortgages. This allows the database to provide high-frequency information on mortgage delinquency rates.

<b>Table 2</b>				
<b>Archive Date</b>	<b>Sample Tradelines</b>	<b>Final Loans</b>	<b>Final Borrowers</b>	<b>Percentage of Tradelines Dropped</b>
Jun 2013	648,224	496,592	771,531	23.4
Sep 2013	240,001	131,169	199,887	45.3
Dec 2013	174,404	109,254	162,283	37.4
Mar 2014	111,928	53,778	79,809	52.0
Jun 2014	146,406	77,726	114,939	46.9
Sep 2014	124,389	75,592	112,283	39.2
Dec 2014	124,323	76,291	112,935	38.6
Mar 2015	104,613	71,262	105,991	31.9
Jun 2015	129,737	91,819	137,092	29.2
Sep 2015	150,399	98,347	146,370	34.6
Dec 2015	124,413	88,098	130,534	29.2
Mar 2016	123,438	74,718	110,052	39.5
Jun 2016	111,797	83,282	123,079	25.5
Sep 2016	135,699	103,698	152,773	23.6
Dec 2016	177,386	108,554	160,915	38.8
Mar 2017	137,917	96,324	142,712	24.7
Jun 2017	129,953	81,602	119,490	37.2
Sep 2017	125,278	93,272	137,435	25.5
Dec 2017	154,468	92,089	135,066	40.4
Mar 2018	109,340	80,150	116,754	26.7
Jun 2018	117,964	80,896	117,712	31.4
Sep 2018	126,255	86,277	126,224	31.7
Dec 2018	134,045	79,973	116,023	40.3
Mar 2019	147,057	66,152	95,548	55.0
Jun 2019	189,499	82,600	119,988	56.4
Sep 2019	170,330	106,758	156,596	37.3
Dec 2019	168,428	127,679	187,799	24.2

## 6. Merging with other Data Sources

Although extensive, Experian's archive files do not contain information on a number of key mortgage features, such as the loan's purpose (home purchase or refinance), whether it had an adjustable or fixed rate, its securitization status, its origination channel (broker or retail lender), or whether it was for an owner-occupied property, vacation home, or investor property. Moreover, Experian's archives contain no information on the property backing the mortgage, such as its location, purchase price, characteristics, or current value. Finally, key information on borrowers associated with the loan including income is also missing. Consequently, values of these key variables need to be inferred indirectly or acquired from other data sources if they are to be included in the NMDB.

The NMDB obtains much of the missing information from matches to administrative file records. The core administrative files come from Fannie Mae and Freddie Mac (the Enterprises), the Federal Housing Administration (FHA), the U.S. Department of Veterans Affairs (VA), and the Rural Housing Authority (RHS). Collectively, loans associated with these programs comprise about three-quarters of the loans in the NMDB.

The most accurate means of merging information from outside sources into the NMDB is to use information about the borrowers, such as their names, Social Security numbers, addresses, and dates of birth. Using such directly identifying information (PII), however, heightens concerns about data security and borrower privacy. Consequently, FHFA contracted with an outside consultant to conduct a study of how such concerns might be mitigated. The third-party-blind matching process that FHFA used is consistent with the "best practices" and recommendations from that study.

The third-party-blind matching process adheres to three guiding principles. First, neither FHFA, FHA, VA, RHS, nor the Enterprises can receive PII from Experian. Second, Experian cannot access FHA, VA, RHS or Enterprise administrative data and borrower PII in the same place. Third, FHFA must not be able to match loans in the NMDB records to the specific administrative records from FHA, VA, RHS, or the Enterprises.

In December 2014, a process was initiated to supplement the NMDB data with administrative data from Fannie Mae and Freddie Mac. Subsequently, the same process was repeated with FHA, VA, and RHS. The process for matching the data follows seven steps:

- (1) The data partner (e.g. Fannie Mae) creates a unique anonymized identifier (AID) for each loan. This identifier, along with the borrower-level PII associated with each loan (including name, address, Social Security number, and date of birth), is transmitted directly to Experian using a secure portal. FHFA does not receive this information. Other administrative data on these loans is not sent to Experian.
- (2) The data partner simultaneously sends the AID, along with administrative data for each loan, to an FHFA data processing unit that is separate from the NMDB development team. No borrower-level PII is included in the information sent to the FHFA data processing unit. The FHFA data processing unit transmits the administrative data along

with the associated AID to another unit within Experian which is separate from the unit that receives the PII in step (1).

- (3) Behind a secure firewall to protect FCRA-regulated data, Experian matches the PII it receives from the data partner to the PII maintained in its own files on the borrowers in the NMDB to determine potential matches. When a potential match is identified, Experian compiles the PIN for each matched borrower.
- (4) For all potential matches, Experian transfers the partner-supplied AID and the matched NMDB borrower PINs to a separate unit within Experian that has no access to the credit repository data or any PII. This is the same unit that received the administrative data from FHFA in step (2).
- (5) The second Experian unit matches the AIDs received from the first Experian unit in step (4) with the AIDs sent by FHFA in step (2). For all matches, the second Experian unit forwards the administrative data they received from the data processing unit at FHFA, plus the matched borrower PIN that they received from the first Experian unit, to the NMDB development team at FHFA. The information sent to the NMDB development team includes neither the Enterprise-created AID nor any PII.
- (6) The NMDB team compares the characteristics of the loans associated with the PINs received from the second Experian unit to the administrative information on the loans. If the information from both sources was consistent, the match is confirmed. A list of confirmed matches is sent to Experian. Upon confirmation, Experian stores the property address supplied as part of the PII file from the Enterprises but otherwise permanently destroys all PII used in the match.

As of this writing, the file matching process has been completed for all historical loans and is repeated quarterly, reflecting loans which have been newly added to the NMDB.<sup>13</sup> Generally, the matching process operates with a lag where it takes about 10 weeks after the end of a quarter, for loans reported to Experian during that quarter to be flagged as confirmed matches. As is noted below, the matching process has proved to be very accurate. Match rates for all five administrative partners are in the high 90 percent levels. The only potential significant shortfall is for loans on non-owner-occupied properties and Fannie Mae and Freddie Mac where match rates are only about 90 percent. One explanation is that borrowers on some of these loans are partnerships or LLCs and thus not in the Experian files.

Efforts are currently underway to merge property record information into the NMDB, using similar third-party blind matching techniques. At present, all loans in the NMDB database added

---

<sup>13</sup> The administrative file matching process is actually done with all mortgages tracked in the NMDB. Because the database contains the complete credit bureau records of all sample loan borrowers performe it will contain records on any other mortgage loans owed by sample loan borrowers as long as they were closed within ten years of the opening of the sample loan or opened within a year of the sample loan being closed. This includes HELOCS and second liens as well as other closed-end first liens. Information on these loans is subject to the same NMDB cleaning and de-duping process used for sample loans. At present there are 55 million total mortgages tracked in the database of which only 12.7 million are sample loans.

through mid-2019 have been matched to property records with plans to match the remaining ones shortly.<sup>14</sup> For those loans where property matching is complete, data from servicing and private-label databases have also been matched which provides missing data elements for most of the non-government-affiliated loans in the NMDB.

In addition, loans in the NMDB have been matched to loan-level information on loans reported under the Home Mortgage Disclosure Act (HMDA) from 1990 to 2018, loans included in the McDash servicing database, and to data files on the 2.1 million loans purchased by the Federal Home Loan Banks.<sup>15</sup> These matches do not involve PII and thus rely on less accurate matching techniques.

Now in full production, the NMDB combines data from all of these sources into a common file with one record per sampled loan. The record contains variables reflecting all the static characteristics of the loan, culled from multiple sources, as well as vectors of dynamic data, such as the monthly performance of the loan from origination to termination, changes to its interest rate in each month (if a variable rate loan), and the associated loan balances. It should be noted that information from external databases is only used to supplement information about sample loans, not to add new loans to the sample. The NMDB sample frame will continue to be that established in the Experian data files. All information on mortgage performance likewise comes from Experian.

---

<sup>14</sup> To facilitate the property matching, the entire property database of one of the two largest U.S. property data vendors has been placed behind the secure firewall at Experian. This allows information on borrower name and address to be used in the matching process. Again, any PII used in the match is discarded once the matching process is completed.

<sup>15</sup> Such merges use information common to the NMDB and the external dataset to perform a match but not PII. Primarily the matches rely on the original loan balance, the opening date of the mortgage and the general location of the property (census tract, ZIP Code or state/county). Unfortunately, mortgage servicers report the billing address of the mortgage borrowers to Experian, but this is not necessarily the property address, particularly for mortgages on non-owner occupied properties. Additional address information maintained within Experian's databases sometimes proves useful in supplementing the repository addresses, as does historical information on borrower location. Nevertheless, such merges are less accurate than those employing PII because the later are less reliant on address.



## 7. Production

On October 30, 2017, NMDB 1.0 was certified and released for internal production use at FHFA and CFPB. The NMDB 1.0 dataset featured payment performance data for the 11.9 million NMDB sample mortgage loans and 18.6 million borrowers associated with those mortgages as of June 2017.<sup>16</sup> The initial release was used by the CFPB to support a public release of monthly county-level mortgage performance statistics. More detailed background information is provided on this component of NMDB 1.0 in Appendix B.

An updated production database, NMDB 2.0, was certified and released for internal production use at FHFA and CFPB on May 31, 2018. This process has been repeated quarterly through the present (NMDB 9.0 was released on March 10, 2020).

NMDB 9.0 is a fully developed database consisting of information on 12.7 million sample mortgage loans and 19.7 million associated borrowers. The database is cumulative, in that potential updated information is given on all sample loans for each quarterly update. Only about 1/4th of the database sample loans are currently active, but information on closed loans can still be updated due to new administrative file matching.

The database fully reflects the results of the NMDB matching process. The database is “complete” with no missing variable values for any sample loans. Most values reflect “hard” information from administrative or Experian files. In some instances though, particularly with older loans, no administrative matches could be made. In these instances, information that would have been available from an administrative match, such as loan-to-value ratio (LTV) or borrower income, are imputed using statistical techniques. These techniques rely heavily on census-tract-level variables created from administrative files that reflect other loans made in the borrower’s neighborhood during the same year that the loan was originated.

The dataset includes the state, county and census tract that the property associated with the loan is estimated to be in 2019 and in the year the loan was originated. For loans with an administrative match, the address supplied by the administrative partner is geocoded by Experian and the tracts supplied to the NMDB. For other loans, Experian maintains addresses for each borrower compiled from the billing addresses supplied by loan servicers augmented by addresses from marketing sources. FHFA is supplied with the 2000 and 2010 census tracts for each address by Experian. From these data and the dates they were first reported to Experian, the tract for each mortgage was inferred. For example, for all active mortgages the tract of the most recently reported address is used. For older loans, addresses are given preference when the borrower moves around the time the loan is originated.

The dataset also includes “static” loan variables based on origination values contained in the

---

<sup>16</sup> The production database excludes a small number of loan/borrower records because the borrower was not originally associated with the loan when it was sampled but added by Experian (or the servicer) later. This can only be picked up if the borrower happened to be in the NMDB because of another loan and thus will not be representative of all new borrowers added to loans. Consequently they are excluded and the production database should be considered a dataset of all loan/borrower combinations where the borrower was associated with the loan from the beginning.

Experian files including: original term (months), loan amount, balloon status, origination date, and number of borrowers. Administrative matches are used to populate a number of other origination variables including property (collateral) value (used to compute loan-to-value (LTV)), income relied upon in underwriting, back-end debt-to-income ratio (DTI), occupancy status, ARM status, loan type (FHA, VA, RHS), purchased by GSE (Fannie Mae, Freddie Mac, FHLB), in private security (PLMBS), property units, manufactured housing and HARP loan flags, and initial contract interest rate. The previous loan payoff balance (refinance only) and simultaneous new second lien (piggyback or HELOC) data are used to derive combined-loan-to-value (CLTV) and “cash-out” status.

Demographic data on loan borrowers derived from variety of sources are used to construct variables for age at loan origination, gender, race, ethnicity and VantageScore 3.0 at origination. Information on the date of the borrower’s first mortgage in the credit bureau and whether they have ever had a VA loan are used to infer veterans and first-time homebuyer status.

The dataset also contains a number of “dynamic” variables which vary by time. These include the monthly and quarterly unpaid loan balance and loan performance metric (e.g. 30 days past due), quarterly credit scores for each borrower, the quarterly FHFA house price index applicable to the each loan’s county (can be used to estimate the loans current loan-to-value ratio), quarterly loan inquiry counts for borrowers associated with the loan (can be used to forecast prepayment), variables reflecting loan modifications, and the closing date (if applicable).

The exact variables included in each new production release are described in the NMDB codebook which is updated quarterly.

## 8. Evaluating the NMDB Sample Frame

The NMDB is intended to be representative of the mortgage market as a whole. One way of testing whether this goal was achieved is to compare loan totals implied by the NMDB with control totals obtained elsewhere. Comparisons were made of implied national quarterly mortgage originations derived from the NMDB with control totals obtained from the five NMDB partner agencies (FHA, VA, RHS, Fannie Mae and Freddie Mac) for the period 1998 to 2017. The comparisons for Fannie Mae and Freddie Mac are done separately for loans backed by owner-occupied properties and those that are not.

The NMDB tracks loan originations for the five partners remarkably well. Almost all quarterly ratios are over 90 percent with the vast majority over 95 percent (the average quarterly match rates are 98.3% and 99.1% for Fannie Mae and Freddie Mac owner-occupied properties, 99.5% for FHA and 98.0% for VA). The average RHS ratio is somewhat higher than the others (101.1%) likely because the control totals are only for the RHS-guaranteed program whereas the NMDB includes some RHS-direct loans. There are issues with timing where a shortfall in one quarter is often offset by overshooting in the next. Missing information on opening date is common, particularly in the control files which generally reflect the current servicer not necessarily the originator (NMDB cleaning rules use the original opening date when servicing is transferred). When this happens servicers often assume that the loan is opened two months prior to the date of the first payment. Reflecting this, when the Experian date differs from the servicing file date for matched loans, 47 percent of the servicing file loans have an opening date of the 17<sup>th</sup> of the month—an implausible result. Also the reported opening date falls on a Saturday or Sunday for 12.4 percent of the matched servicing files versus only 3.7 percent for Experian (loans almost always close on a business day).

NMDB also compares well with HMDA and the FHLB program. There is somewhat of a shortfall for PLMBS securities loans with an overall coverage rate of only 82.9 percent. This is likely caused by the fact that these loans require a double match – NMDB to a property and the CoreLogic property address to an address in the securities servicing file. Only about one-half of the loans in the securities servicing file have a complete address, inhibiting the match process. To supplement the Experian match, additional direct non-PII matching between NMDB and the CoreLogic securities servicing file was conducted but with less than full success.

HMDA loan totals for owner-occupied properties fall slightly below those estimated from the NMDB. This may stem from known gaps in HMDA's coverage. Loan originators that are very small or that operate exclusively in rural areas are exempt from HMDA reporting requirements, so their lending activity is not included in the HMDA data. Additionally, HMDA data excludes commercial loans and (non-purchase) loans backed by properties that were previously mortgage-free. Many of these loans, however, may not be reported to the credit repositories either. For example, loans to corporations, loans made as part of a seller-financed property sale, and loans made by non-traditional lenders are unlikely to be in either database. Moreover, some types of loans may be missed by the NMDB though they are captured in the HMDA data. Some lenders that retain all of their loans in portfolio, particularly credit unions, are known not to report their loans to the credit repositories, but are nevertheless still subject to HMDA reporting requirements.

Loans backed by owner-occupied properties compare more favorably with control totals and HMDA (an average quarterly origination ratio of 104.3%) than those backed by investor and secondary homes (82.2%). Some of this difference may be accounted for by the fact that some investor loans are done through partnerships and LLCs and thus may not end up being reported to the consumer credit bureaus which form the basis of the NMDB. Another explanation of the lower match rate for PLMBS loans is that these loans are disproportionately made for secondary or investment properties.

The success of the NMDB coverage is a reflection of the quality of the matching process. For every product-type except RHS Direct, 98 percent or more of the matches have a dollar-perfect match between the “original loan amount” reported in Experian and the amount shown in administrative records. As discussed above, however, “date opened” matches at a much lower rate (except for RHS Direct). Fannie Mae loans, in particular, have very low match rates. This may occur, in part, due to a relatively high rate of servicing transfer for loans in the five partner programs. In contrast, for loans “held in portfolio” and not placed in PLMBS securities or sold to the FHLB banks, only 18 percent have a servicing transfer.

Despite the fact the Experian maintains the billing address for the mortgage and the administrative records maintain the property address, over 90 percent of the matches found an exact address match for all but the non-owner-occupied properties. For non-owner-occupied properties, address matches were only about 60 percent, potentially contributing to lower overall match rates for these types of loan.

Another validation of the NMDB matching process is the high correlation between credit bureau codes indicating FHA, VA, RHS, Freddie Mac and Fannie Mae loans and the match results (which do not depend upon these codes). Over 98 percent of the FHA and VA loans matched to administrative files were correctly marked by bureau account type codes. Only two-thirds of the Freddie Mac and Fannie Mae loans were correctly identified despite rules that servicers are supposed to report these loans to the bureaus. Some of this may be due to timing, as the reporting rules were not put into effect until 2010. More than 80 percent of Freddie Mac and Fannie Mae loans originated since 2008, for example, are correctly flagged.

Overall, these results suggest that the NMDB should provide a very accurate representation of the mortgage market as a whole and for the market subprograms represented by FHA, VA, RHS, FHLB, and loans backed by owner-occupied properties sold to Fannie Mae and Freddie MAC and in PLMBS securities. There does not appear to be a materially different degree of accuracy between earlier years and post-crisis originations. Loans backed by investor and secondary properties do appear to have somewhat lower levels of coverage but coverage levels still should be above 80 percent even in earlier years.

## Appendix A. Origins of NMDB

Prior to deciding to develop the NMDB, FHFA and CFPB considered a number of alternatives to meet their data requirements. The primary alternatives explored were the Home Mortgage Disclosure Act (HMDA) data, the Federal Reserve Bank of New York's Equifax Consumer Credit Panel, the CoreLogic property database, the servicing databases owned by CoreLogic and Black Knight Financial Services, and data available from the three national credit repositories—Experian, Equifax, and TransUnion. Public survey databases, particularly the American Housing Survey (AHS), were also considered. All of these sources share several desirable features such as: (1) the databases are de-identified containing no direct-identifying information such as borrower name, address, or Social Security number; (2) they are collected for other purposes, thus their use entails no new data collection from lenders, servicers or borrowers; and (3) all of them have been collected for a period of time and are expected to continue into the future.

However, each was also found to be deficient in significant ways.

The HMDA data include loan applications and underwriting outcomes for most mortgages with selected information about the loan, property, and borrower. The data are arguably the most representative publicly available existing data source about the mortgage market. However, the HMDA data contain no information on loan performance, little information on borrower credit-worthiness, and have up to a 21-month delay in release. The CoreLogic property database suffers from similar deficiencies. Although it has widespread coverage, the database contains very limited information on mortgage characteristics or performance and nothing on the borrower.

The Federal Reserve Bank of New York's Equifax Consumer Credit Panel provides a nationally representative 1-in-20 sample of individuals with credit records, observed quarterly from 1999 onward. However, mortgage loans are often represented by duplicate trade lines and important information is missing, such as loan purpose, owner-occupancy, pricing, loan-to-value ratio, income, and borrower demographics. Finally, these data are accessible at present only to the Federal Reserve System.

CoreLogic and Black Knight Financial Services produce loan-level databases with performance information collected from mortgage servicers. The servicing fields available from CoreLogic and Black Knight are relatively comprehensive in both variables and coverage: the CoreLogic database claims about 32 million active mortgage loans, while the Black Knight database claims about 31 million active mortgage loans. However, these data offer no assurance of being representative, as data are only collected (currently) from about 55 servicers each. Moreover, mortgages cannot be tracked if servicing is transferred. Other drawbacks include minimal borrower demographics and no information on the borrower's other obligations.

The semi-annual AHS contains comprehensive information on a nationally representative 1-in-2,000 sample of mortgages of owner-occupied properties with very good information about the property and borrower demographics. However, the AHS has only limited information about the mortgage itself. As with the other nationally representative consumer survey data sources, AHS contains no information on mortgage performance, provides only a small number of

observations, and is released with a significant lag.

The credit repository data from Equifax, Experian, and TransUnion are rich in credit information. By construction they incorporate data on credit card debt, installment loans, credit inquiries, and public records for the consumers they have in their respective databases. Their data can be linked to marketing datasets that provide borrower characteristics including age, gender, and marital status which, if validated, could be of potential use in a dataset. The credit repositories also maintain data on borrowers' changes of address and broader geographic classifications, such as the census tract. However, there are important areas that are not covered. They lack some information on borrowers (*e.g.*, income), mortgages (*e.g.*, loan product and contract rate), and the underlying property (*e.g.*, location and value).

Given the foregoing, FHFA and CFPB, along with HUD, the Federal Reserve Board, Freddie Mac, and others, decided that a modified derivative of the credit repository data offered the best source from which to construct a nationally representative comprehensive mortgage database. The three credit repositories all actively pursue loan servicers as data providers. As a result, they obtain information on almost the entire population of non-private mortgage loans made in the United States. Furthermore, they archive their data, making it possible to “jump start” the data collection process by going back in time, collecting data in almost the same fashion as if it had taken place in real time.

As part of the exploratory process, using a competitive procurement process, Experian was engaged by Freddie Mac to construct a prototype to confirm the appropriateness of using credit repository data for the database. This effort confirmed the concept but suggested that a number of steps needed to be taken in order to meet the design objectives.

First, it was recommended that the database should be a sample rather than a universal registry of loans. Second, that the database be restricted to closed-end first lien mortgages, mimicking the coverage of HMDA and the availability of matching information. Third, while these data contain detailed information on loan performance and other borrower credit obligations, they are missing critical data items needed for the database such as the location and features of the property, demographics, and loans characteristics such as whether the loan had an adjustable- or fixed-rate mortgage and whether the loan was a refinance or for a home purchase. Thus, it would be necessary to access other data sources and merge information gleaned from them with the repository data in order to make the database comprehensive. Pilot testing also confirmed that the best method of merging data would rely on third-party blind matching conducted behind a firewall at the credit repositories.



## Appendix B. Background on Mortgage Delinquency Reporting

Almost all closed-end first-lien mortgages, such as those in the NMDB, have a payment-in-arrears structure. That is, the mortgage payment for a month (*e.g.*, January) is generally due on the first day of the next month (*e.g.*, February 1). Moreover, the first “ever” payment on a mortgage is generally due on the first day of the second full month after the mortgage closing date. For example, borrowers who close on their mortgage on January 15 will have their first payment due on March 1. These borrowers would have prepaid the interest for the period covering January 15 through January 31 at their mortgage closing. One component of each monthly mortgage payment is the interest of the previous month based on the balance at the beginning of the month. While the monthly mortgage payment is generally due on the first of the month, most lenders allow a 15-day grace period for borrowers to pay. However, if the payment is not received by the 15th, the mortgage loan is considered past due 15 days. Thereafter, a loan not paid by the X-th date after the due date is considered X days past due.

Each month mortgage servicers report the performance information to the credit bureaus for each mortgage loan they service as of a snapshot date (balance date). Generally, the bureaus will accept only one report per loan per month. Servicers report three measures of performance: (1) the account condition code which describes the condition of the mortgage, *e.g.*, whether it is open, paid in full, closed, transferred, or inactive; (2) a special comment code which provides special information on the mortgage such as a loan modification, location in a disaster area county, or dispute by the borrower; and (3) a loan status code which provides information on how many days “past due” a loan is as determined by the oldest non-paid payment (loan payments are generally applied against the oldest non-paid payment). Industry and Metro 2 credit bureau reporting guidelines (available since 1997) differentiate between loans that are current or past due 29 or fewer days; 30 to 59 days past due; 60 to 89 days past due; 90 to 119 days past due; 120 to 149 days past due; and 150 to 179 days past due; and 180 or more days past due. This is the classification used for most loans. If a loan becomes 90 days past due under many mortgage contracts the lender can declare the loan “in default.” The borrower then typically has 90 days to become current. If not, the lender can file a foreclosure action in which case, the loan status is changed from “days past due” to some form of foreclosure or collection. If the borrower files for bankruptcy, the loan may be assigned a “bankruptcy” status even if the payments are current.

A status code may be suppressed or not reported for some loans. This can occur for a variety of reasons—a borrower’s payments may have been suspended because of a natural disaster; reports may not be supplied for the period between a loan’s closing or the first due date; status is often not continually reported when loan servicing is sold from one lender to another when it takes time for the acquiring lender to set up reporting; status updates are often not reported for loans in foreclosure or other forms of serious delinquency (there are spikes in missing values for August 2015 and April 2016 because several larger servicers had problems with their servicing systems). There are also loans which do not fit these circumstances—some loans do not have due dates of the first; others have bi-weekly or quarterly payment requirements; some borrowers make partial payments (often to what is called a suspense account) which can leave them in a perpetually past-due status; others can make extra payments (curtailment) to reduce their loan balance more rapidly.

Servicers reporting to the credit bureaus using Metro 2 guidelines are supposed to follow the guidelines described above. Thus, a loan with a due date of January 1 will be considered 30 days past due on January 31 if the payment has not been received by that date (the “days past due” standard). In the past, however, many lenders used a “billing cycle month” standard. Under the “billing cycle month” standard, a loan was not considered “30 days” past due until the due date of the next month (e.g. February 1 for a January 1 payment). The “billing cycle month” standard, associated with the older Metro 1 reporting format, was phased out over the 2000s for bureau reporting. However, this has not necessarily happened for other regulatory reporting. Mortgage delinquency metrics reported to the Federal Financial Institutions Examination Council (FFIEC) for banking institutions can be based on either “days past due” or “billing cycle month” standards at the reporter’s discretion. Credit unions used the “billing cycle month” standard until 2013 when they were required to report using both methods. Freddie Mac and Fannie Mae report delinquency statistics for loans in their security pools using the “billing cycle month” standard.

Within the Metro 2 reporting guidelines and for other delinquency reporting there is also variation based on precisely when a loan’s status is measured. Under the Mortgage Bankers Association’s “MBA” method, a loan is considered past due X days if a payment is not received by close of business (COB) on the X-th day following its due date. That is, a loan with a due date of March 1 is considered 30 days past due at COB on March 31. Under the Office of Thrift Supervision’s “OTS” method, a loan is considered past due X days if the payment has not been received by COB on the X+1-th day (e.g., April 1 for a March 1 due date).

The credit bureaus allow loan servicers to choose whichever reporting day within the month that they wish to use, and either the MBA or OTS method. Currently, about 90 percent of reporters use the same day of the month every month and the same day for all of their loans. For NMDB loans active in 2014 and later, the modal report day (31 percent) was the last day of the month; 16 percent were on the 5th; 12 percent on the 7th, and 8 percent on the 21st. For “prime” first lien closed-end mortgages, which dominate the NMDB, lenders generally use the MBA method. Subprime servicers, however, who played a significant role in the 2003 to 2007 period, typically used the OTS method.

These differences in reporting day and method can lead to significant variation in the incidence of delinquency for loans with identical payment patterns when comparisons are made month-to-month or between lenders with different reporting patterns. This is shown in Table B-1. Servicers who report at the end of the month using the MBA method will maximize the 30-day delinquent count in the seven months with 31 days because the reporting day is the first day a loan can be 30 days delinquent. Lenders reporting in the latter half of the month—but not on the last day—will tend to systematically show lower delinquency rates.

These distinctions will matter when aggregated measures of delinquency are computed, particularly those for 30-days past due. For example, in 2016 lenders in the NMDB reporting on the last day of the month showed an average 30-day delinquency rate 0.76 percentage points higher in the seven months with 31 days than they did for the five months with 30 days or less. For servicers reporting between the 16th and the second-to-last day of the month there is only a



0.03 percentage point difference.

These reporting differences can cause systematic differences across states as well. For example, 47 percent of the 2016 reporters in Mississippi were end-of-the-month reporters versus 17 percent in Alaska, almost surely influencing the number of 30-day delinquencies. Because reporters can only report loan status once a month and it is impossible to know when a loan payment was received for many loans, this bias is difficult to correct for. If no adjustment is made, users of a bureau-based delinquency series need to be cautious in making comparisons across geographic units. Also, as just noted, numbers for the NMDB are likely to show persistent monthly patterns if results are not seasonally-adjusted.

The mixture of reporting patterns in the credit bureau data is likely to lead to systematic differences in aggregate delinquency metrics constructed from the NMDB data when compared to other delinquency measures. For example, The MBA National Delinquency Survey asks respondents to classify loans by their status at COB on the last day of the quarter using the MBA method although it appears that the lender can use either the “billing cycle month” or “days past due” standard. If the “days past due” standard is used it means that reports for March and December (which each have 31 days) will show persistently higher 30-day delinquency rates than those of June and September (which each have 30 days). Similarly, 90-day delinquency rates will be lower in the first quarter except for leap years. The degree of seasonality will depend on what percentage of the reporters use the “billing cycle month” versus “days past due” standard. FFIEC call report statistics, which are also reported COB on the last day of the quarter, will also exhibit seasonality depending on the mix of lenders using different methods. FFIEC statistics are further clouded by the fact that lenders can use either the MBA or OTS accounting method.

Delinquency statistics reported by Freddie Mac and Fannie Mae for loans in their security pools should show the least month-to-month distortions. Both companies use an end-of-month measure computed using the MBA and “billing cycle month” standard which should lead to stable monthly patterns. Given that the standard mortgage contract is based on a monthly payment standard there is a compelling argument that the “billing cycle month” standard for measuring delinquency is the most appropriate. Nevertheless that is not the standard used by most credit bureau reporters and thus is not the standard reflected in the NMDB data.

The delinquency data in the NMDB are built from the lender reports supplied by Experian but with some additional processing. The performance information supplied by lenders for loan status, account condition, and special comments is static; that is, each month when the servicers update the performance data for a loan, the previous values for these variables are overwritten with new information. The values supplied in the previous months can only be recovered from archives. However, all the credit bureaus maintain an abbreviated record of historical performance, known as a payment grid, which is not overwritten, but can be (and is) updated. Under FCRA rules the payment grid can only go back 84 months. When an initial report is supplied for a month (say June 2016) the “June 2016” element of the payment grid is initially populated. However, in subsequent filings the servicer can change the “June 2016” value. This can happen for a variety of reasons—the lender can catch an error, they may have inadvertently failed to report performance in the first filing, the consumer could dispute the report and get the

record changed, or the report could subsequently be suppressed, for example, because the borrower was impacted by a natural disaster.

In general, the monthly performance measure in the NMDB is constructed from the payment grid, using the most recently reported information for a given month. Payment grids retrieved from archival data—which were collected quarterly from June 2012 on and semi-annually before that—were used to piece together a full measure of performance and to get around the 84 month limitation on current data. An additional problem is created when loans are transferred from one servicer to another. Here, payment grids need to be combined for two different reporters to create a continuous measure of performance. Often when this happens, the transferring servicer will initially report the loan as delinquent but then correct it when they receive the transfer notice. Transfers often create gaps in the payment grid when the new servicer is slow to report the loan. FCRA rules also place restrictions on how the new servicer reports performance under the assumption that some borrowers may have sent payments to the wrong place.

The effect of this process is that the initial performance report for a loan is often subsequently changed. On net, this tends to improve the overall measure of performance, but in recent years the change is small. For example, the initial NMDB report for June 2016 differed from “final” report in June 2017 for 1.3 percent of the cases. The majority of these were blanks in the initial report but there were some real changes. Changes went both ways—2.3 percent of the loans originally reported as 30-days past due were corrected to current. But an almost equivalent number were changed from current to delinquent.

On balance, the updating process reflected in the NMDB is likely to mean that in recent years delinquency measures in the NMDB will be slightly more positive than other indices, such as the Equifax index, which are compiled only from the initial report. However, during the mid-2000s when the private label subprime market was a significant part of the mortgage market, sale of servicing was more prevalent and more likely to have led to initially inaccurate delinquency reports. Here, the NMDB data show noticeable differences from indices based on initial reports.

Another difference arises when seriously delinquent loans are transferred within an organization (say from normal servicing to “work out” departments). It is not unusual for the loan to be reported as open and delinquent by both departments creating a double counting if not corrected. In constructing the NMDB these reports are combined, but may not be in other indices which are based on open accounts with positive balances. Consequently, indices of serious delinquency constructed from the NMDB will likely be lower than those constructed from other sources.

Finally, there is some ambiguity as to how to define an open account. It is not unusual for lenders to initiate foreclosure actions on small mortgage loans but never complete the process, perhaps because they decide the property isn’t worth acquiring. In other cases, state law allows lenders to maintain a claim on the borrower, termed a deficiency judgement, after a foreclosure. These loans can remain on the Experian files as open, with positive balances, for a long time until they are purged by FCRA rules. However, the borrower may well have lost title to the house or moved out much earlier in the process.

Table B-1														
Impact of Reporting Cycle Standard and Reporting Method on Delinquency Measurement														
Reporting of Days Past Due for a Mortgage where Payments were Stopped*														
Month Payments Stopped	MBA Method**							OTS Method**						
	Current Month	Month Plus 1	Month Plus 2	Month Plus 3	Month Plus 4	Month Plus 5	Month Plus 6	Current Month	Month Plus 1	Month Plus 2	Month Plus 3	Month Plus 4	Month Plus 5	Month Plus 6
<b>End of Month Reporters - Days Past Due Standard</b>														
January	<b>D30</b>	<b>D30</b>	D60	D90	D150	D180	D180	C	D30	D60	D90	D120	<i>D150</i>	<i>D180</i>
February	C	D30	D60	D90	<i>D120</i>	<i>D180</i>	D180	C	D30	D60	D90	D120	<i>D150</i>	<i>D180</i>
March	D30	D60	D90	D120	D150	D180	D180	C	<i>D30</i>	<i>D90</i>	D120	D150	D180	D180
April	C	<i>D60</i>	D90	D120	D150	D180	D180	C	D30	<i>D60</i>	<i>D120</i>	D150	D180	D180
May	D30	D60	D90	D120	D150	D180	D180	C	<i>D30</i>	<i>D90</i>	D120	D150	D180	D180
June	C	<i>D60</i>	D90	D120	D150	D180	D180	C	<i>D30</i>	<i>D90</i>	D120	D150	D180	D180
July	D30	D60	D90	D120	D150	D180	D180	C	<i>D60</i>	D90	D120	D150	D180	D180
August	D30	D60	D90	D120	D150	D180	D180	C	<i>D30</i>	<i>D90</i>	D120	D150	D180	D180
September	C	<i>D60</i>	D90	D120	D150	D180	D180	C	D30	<i>D60</i>	<i>D120</i>	<b>D150</b>	<b>D150</b>	<i>D180</i>
October	D30	D60	D90	D120	D150	D180	D180	C	<i>D30</i>	<i>D90</i>	<b>D120</b>	<b>D120</b>	D180	D180
November	C	<i>D60</i>	<b>D90</b>	<b>D90</b>	<i>D150</i>	D180	D180	C	<i>D30</i>	<b>D90</b>	<b>D90</b>	D120	<i>D150</i>	<i>D180</i>
December	D30	<b>D60</b>	<b>D60</b>	<i>D120</i>	D150	D180	D180	C	<b>D60</b>	<b>D60</b>	D90	D120	D180	D180
<b>Middle of Month Reporters - Days Past Due Standard</b>														
All Months	C	D30	D60	D90	D120	D150	D180	C	D30	D60	D90	D120	D150	D180
<b>Billing Cycle Month Standard Reporters**</b>														
All Months	D30	D60	D90	D120	D150	D180	D180	C	D30	D60	D90	D120	D150	D180

Impact by Number of Days in the Month and Timing of Reporting Date													
Reporting of Missed Payments and Subsequent Cure Using the Days Past Due Standard and MBA Method													
March (31 Day Month)							April (30 Day Month)						
Date Cured	Performance Month						Date Cured	Performance Month					
	March	April	May	June	July	August		April	May	June	July	August	Sept.
<b>End of Month Reporters</b>													
15-Apr	D30	C	C	C	C	C	15-May	C	C	C	C	C	C
15-May	D30	D60	C	C	C	C	15-Jun	C	D60	C	C	C	C
15-Jun	D30	D60	D90	C	C	C	15-Jul	C	D60	D90	C	C	C
15-Jul	D30	D60	D90	D120	C	C	15-Aug	C	D60	D90	D120	C	C
<b>Report on 22nd of Month</b>													
15-Apr	C	C	C	C	C	C	15-May	C	C	C	C	C	C
15-May	C	D30	C	C	C	C	15-Jun	C	D30	C	C	C	C
15-Jun	C	D30	D60	C	C	C	15-Jul	C	D30	D60	C	C	C
15-Jul	C	D30	D60	D90	C	C	15-Aug	C	D30	D60	D90	C	C
<b>Report on 7th of Month</b>													
15-Apr	C	D30	C	C	C	C	15-May	C	D30	C	C	C	C
15-May	C	D30	D60	C	C	C	15-Jun	C	D30	D60	C	C	C
15-Jun	C	D30	D60	D90	C	C	15-Jul	C	D30	D60	D90	C	C
15-Jul	C	D30	D60	D90	D120	C	15-Aug	C	D30	D60	D90	D120	C

Note: This illustration is for loans where the payment due date is on the first of the month. C = Current. D30 = 30-59 days past due. D60 = 60-89 days past due. D90 = 90-119 days past due. D120 = 120-149 days past due. D150 = 150-179 days past due. D180 = 180 or more days past due.

\***Bold** indicates where performance reporting stays the same and *italics* indicates where reporting skips a reporting cycle.

\*\*See text for explanation of "MBA" and "OTS" methods; and "Days Past Due" and "Billing Cycle Month" standards.





