WORLD **PRIVACY** FORUM

Comments of the World Privacy Forum
To
The Consumer Financial Protection Bureau
Regarding Docket No. CFPB-2014-0019, Home Mortgage Disclosure (Regulation C)

Sent via email to FederalRegisterComments@cfpb.gov.

Monica Jackson,
Office of the Executive Secretary,
Consumer Financial Protection Bureau,
1700 G Street NW.,
Washington, DC 20552.

Re: Docket No. CFPB-2014-0019

October 29, 2014

The World Privacy Forum appreciates the opportunity to submit comments on the Consumer Financial Protection Bureau's proposed rule on the Home Mortgage Disclosure Act (HMDA). The proposed rule appeared at 79 Federal Register 71732 (Aug. 29, 2014), https://www.federalregister.gov/articles/2014/08/29/2014-18353/home-mortgage-disclosure-regulation-c.

The World Privacy Forum is a non-profit, non-partisan public interest research group. Our activities focus on research and analysis of privacy issues, along with consumer education.[1]

I. Securities and Exchange Commission Asset Backed Securities Proposals

We begin our comments by pointing to 2010[2] and 2014[3] proposals by the Securities and Exchange Commission to make public the personal information about individuals who have mortgages (and other types of loans) as part of an SEC effort to make more data about asset backed securities available to purchasers of those securities. While the SEC's effort to address a different aspect of home mortgage disclosures appears on the surface to be unrelated to the CFPB's HMDA proposals, we think that the two efforts overlap and may threaten privacy in

[1] For more information, see http://www.worldprivacyforum.org.
[2] The first proposed SEC rule appeared at 75 Federal Register 23328 (May 3, 2010), http://edocket.access.gpo.gov/2010/pdf/2010-8282.pdf.
[3] The second SEC action was a request for comment on the Memorandum from the Division of Corporation Finance regarding Disclosure of Asset-Level Data, http://www.sec.gov/comments/s7-08-10/s70810-258.pdf.

similar ways. Both involve making information about mortgages and mortgage applicants publicly available.

In 2010, we said this about the original SEC proposal:

> While we understand the purpose of and motivation for the Commission's proposed rule, the proposal is a direct and substantial threat to the privacy of every individual who obtains a mortgage. If adopted, the proposal would be an unprecedented release of individual-level financial data and would greatly increase borrowers' risk for identity theft and other problems related to the public release of detailed financial information. Specifically, the SEC's proposed rule would:
>
> - Expand the risk of identity theft for every borrower whose information would be disclosed because of the rule;
> - Place on the public record the largest amount of personal financial information about borrowers ever disclosed, including information never before made public;
> - Circumvent or undermine privacy protections in other laws, including the Fair Credit Reporting Act, Health Insurance Portability and Accountability Act, and the Privacy Act of 1974;
> - Weaken the utility and security of knowledge-based authentication techniques and activities by exposing details of mortgages to more people throughout the world and would undermine NIST electronic authentication recommendations;
> - Undermine the financial stability of households;
> - Threaten the stability of the asset-backed instruments that the proposed rule seeks to protect by placing all borrowers at greater risk to be victims of criminal activity and thereby lowering the value of the asset-backed instruments.

This is just a summary of our full comments to the SEC.[4]

We found the SEC's 2014 proposal on the same subject to be better but still not sufficient to meet the privacy needs of borrowers. Here is a summary of what we said to the SEC earlier this year:

---

[4] Comments of the World Privacy Forum to the SEC Regarding the SEC Notice of Proposed Rulemaking on Asset-Backed Securities (Aug. 2, 2010), http://www.worldprivacyforum.org/2010/08/comments-wpf-files-comments-on-deeply-flawed-sec-plan-2.

> The new proposal, while far from perfect, does a better job of balancing the Commission's interest in disclosure with the privacy interests of borrowers. We do not endorse the proposal wholeheartedly, but we think it offers an outline of an approach that will do a better (but not perfect) job of protecting the privacy and other interest of consumers *with some adjustments*.
>
> Key elements of the Division of Corporations Finance's new proposal are:
>
> - Relying on issuers to provide information to investors and potential investors through a Web site without disseminating the potentially sensitive information on EDGAR
>
> - Use of an issuer or issuer-sponsored Web site that would permit issuers to implement their own procedures that would provide both flexibility and the possibility of innovation in the delivery of information and the design of privacy controls.
>
> - Issuers would provide access to potentially sensitive asset-level data to investors and potential investors and would be permitted to restrict access as necessary to comply with privacy laws.
>
> We would be happier with this proposal if it were practical to keep all sensitive asset-level data under the direct control of the Commission or, perhaps, the Consumer Financial Protection Bureau. Direct involvement by a federal agency, while no guarantee of a better outcome for data subjects, would provide better and clearer accountability for maintenance of the data as well as the possibility of meaningful enforcement.[5]

We hope and assume that the CFPB is aware of the ongoing SEC activities regarding mortgage information disclosure. We believe strongly that the proposals of the SEC and the CFPB overlap as far as both affect privacy and both involve public disclosure of personal information about borrowers. We recognize that the two efforts involve different laws and have different objectives.

Without meaningful coordination by the two agencies, we fear that independent decisions could have a multiplier effect on privacy by releasing to the public different views of the same consumer transactions that will result in unplanned data linkage and leakage. Each agency might make a separate and perhaps reasonable judgment, but together the two activities could make consumers substantially worse off than either activity on its own.

We cannot foresee what the SEC will do in the asset backed security domain. Similarly, we do not know how the CFPB will change its HMDA proposed regulations. Our first request is that the two agencies actively talk to each other to make sure that the consequences of their separate

[5] Comments of World Privacy Forum to the Securities and Exchange Commission regarding Disclosure of Asset-Level Data (April 18, 2014), http://www.worldprivacyforum.org/2014/04/public-comments-wpf-files-comments-urging-the-sec-to-protect-asset-level-data-privacy-of-consumers/.

actions on the privacy of consumers be jointly considered, that any negative synergies of the separate proposals be identified in advance, and that the actions of each agency are coordinated and suitably protective of privacy. We recognize and are encouraged that both agencies have expressed similar concerns about consumer privacy. We are confident that both agencies can produce reasonable results that fulfill their separate objectives while still protecting privacy interests.

II. Balancing Test

The CFPB proposal says that it reads the law "to call for the use of a balancing test to determine whether and how HMDA data should be modified prior to its public release in order to protect applicant and borrower privacy while also fulfilling the public disclosure purposes of the statute." We fundamentally agree with the need to balance different objectives. We do want, however, to make sure that the CFPB balance test considers all important factors. We have several thoughts to share regarding this matter.

A. Capabilities of the Consumer Data Industry

In order to appreciate the stakes in any disclosures, it is important to understand the existing marketplace in consumer data. Data brokers, consumer profilers, advertisers, identity thieves and other crooks, and other consumer data users have a vast and seemingly limitless appetite for consumer data.[6] Much of the market for consumer data is wholly unregulated for privacy. We can only describe a tiny fraction of the market here, but the point is that there is a vibrant market for unregulated consumer information, there is a demand for more information, and that any new personal information disclosed can be readily linked with data already in private hands. Data that you might think cannot be tied to a particular consumer or a particular household for a lack of identifiers can nevertheless be connected to that consumer or household by virtue of the other data already available.

WPF issued a report recently on consumer scoring, which is one aspect of the market for consumer data that we just referenced.[7] The report – *The Scoring of America* -- documents the range of personal information already being used. As we discuss consumer scores, keep in mind that we are not talking about credit scores or scores subject to regulation under the Fair Credit Reporting Act. The scores and the data we reference here are not derived from regulated credit bureaus, and they are not subject to any regulation for privacy under any law. Nevertheless, these scores have substantial marketplace impacts. Scores such as the consumer prominence score, loyalty score, churn score, fraud score, even proxy (unregulated) scores using alternative data exist that stand in for the regulated credit score.

These data elements are available for and used in hundreds or thousands of consumer scores:

---

[6] See generally, Testimony of Pam Dixon, Executive Director, World Privacy Forum, before the Senate Committee on Commerce, Science, and Transportation (Dec. 13, 2013), http://www.worldprivacyforum.org/wp-content/uploads/2013/12/WPF_PamDixon_CongressionalTestimony_DataBrokers_2013_fs.pdf.

[7] World Privacy Forum, *The Scoring of America: How Secret Consumer Scores Threaten Your Privacy and Your Future* (April 2014), http://www.worldprivacyforum.org/2014/04/wpf-report-the-scoring-of-america-how-secret-consumer-scores-threaten-your-privacy-and-your-future/.

Demographic Information:

- Age
- Age range
- Date of birth
- Education
- Exact date of birth
- Gender
- Marital status
- Home ownership
- Own or rent
- Estimated income
- Exact income
- Ethnicity
- Presence of children
- Number of children
- Age range of children
- Age of children
- Gender of children
- Language preference
- Religion
- Occupation - category of occupation
- Examples: Beauty (cosmetologists, barbers, manicurists) civil servants, clergy, clerical/office workers, doctors/physicians/surgeons, executives/administrators, farming/agriculture, health services, middle management, nurses, professional/technical, retail service, retired, sales, marketing, self-employed, skilled/trade/machine operator/laborer, teacher/educator.
- Occupation - title of occupation
- Military history
- Veteran in household
- Voter party
- Professional certificates (teacher, etc.)
- Education level reached or median education

Contact Information:

- Full name
- Email address
- City
- State
- ZIP
- ZIP + 4
- Home Address
- Land-line phone
- Social IDs / social media handles and aliases

- Mobile phone number
- Carrier
- Device type
- Email address

Vehicles:

- Vehicle make, model and year
- VIN
- Estimated vehicle value
- Vehicle lifestyle indicator
- Model and brand affinity
- Used vehicle preference indicator

Lifestyle, Interests and Activities data (including medical):

- Antiques
- Apparel (women, men & child)
- Art
- Average direct mail purchase amounts
- Museums
- Audio books
- Auto parts, auto accessories
- Beauty and cosmetics
- Bible purchaser
- Bird owner
- Books
- Book purchases - plus types. (Mystery, romance, religious, etc.)
- Book clubs
- Career
- Career improvement
- Cat owner
- Charitable giving indicators:
- Charitable donor by type of donation (religious, health, social justice)
- Charitable donor by ethnicity or religion (Jewish donors, Christian donors, Hispanic donors)
- Charitable donor by financial status (wealthy donors)
- Children or teen interests
- Fashion and clothing (Multiple: sports, high fashion, shoes, accessories, etc.)
- Collectibles
- Collector
- Christian families
- Computer games
- Computers
- Consumer electronics (Many categories, including electronic fitness devices)
- Dieting and weight loss

- Telecommunications and mobile
- Dog owner
- Investing
- DVD purchasers
- Electronics - home, computing, office, other
- Empty nester
- Expectant parents
- Frequent mail order buyer
- Frequency of purchase indicator
- Getting married
- Getting divorced
- Gun ownership
- Health and beauty
- Health and medical: for example, Allergies, Alzheimer's disease, angina, arthritis/rheumatism, asthma, back pain, cancer, clinical depression, diabetes, emphysema, erectile dysfunction, epilepsy, frequent heartburn, gum problems, hearing difficulty, high blood pressure, high cholesterol, irritable bowel syndrome, lactose intolerant, ulcer, menopause, migraines/frequent headaches, multiple sclerosis, osteoporosis, Parkinson's disease, prostate problems, psoriasis/eczema, sinusitis/sinuses.
- High-end appliances
- Home improvement
- Household consumer expenditures — many categories.
- Jewelry
- Magazine subscriptions
- Mail order buyer
- Mobile location data (some analytics companies)
- Movies - attendance / collector
- Musical instruments
- Music
- New mover
- New parent
- Online and continuing education
- Online purchasing - many categories
- Parenting
- Pets - other
- Plus size clothing purchase
- Political affiliation
- Recent home buyer
- Recent mortgage borrower
- Retail purchasing - many categories.
- Science-related
- Sexual orientation
- Social media sites likely to be used by an individual or household, heavy or light users
- Spa
- Sports interests: (large category, these are examples)
- Birdwatching

- Equestrian
- Exercise and fitness
- Gardening
- Golf
- Fishing
- Outdoor interests - hiking, camping, climbing
- Swimming, diving, snorkeling
- Spectator Sports
- Stamps/coins
- Yoga
- Television, Cable, Satellite/Dish
- Travel: Vacations, domestic and/or international
- Purchase of international hotel or air flights
- Frequent flyer
- Types of purchases indicator
- Veteran in household
- Vitamins
- Volunteering

Financial and Economic – Property and Assets data:

- Estimated income
- Estimated household income
- Home value
- Length of residence
- Payment data: 30, 60, 90-day mortgage lates
- Purchase date
- Purchase price
- Purchase amount
- Most recent interest rate type
- Most recent loan type code
- Sales transaction code
- Most recent lender code
- Purchase lender code
- Most recent lender name
- Purchase lender name
- Fuel source
- Loan to value
- Purchase interest rate type
- Most recent interest rate
- Purchase interest rate
- Pool or spa
- Home - year built
- Air conditioning
- Boat ownership
- Plane ownership

- Motorcycle ownership
- Commercial assets or business ownership

Financial and Credit data:

- Bankruptcy
- Beacon score
- Credit score - actual
- Certificates of deposit/ money market funds
- Estimated household income ranges
- Income producing assets indicator
- Estimated net worth ranges
- IRAs
- Life insurance
- Low-end credit scores
- Mutual funds/annuities
- Summarized credit score or modeled credit score by neighborhood
- Payday loan purchaser
- Number of credit lines
- Tax liens
- Card data:
- Card holder - single card holder
- Range of new credit
- Debit or credit card present in household
- Card holder - brand (Discover, Visa, Mastercard, etc.)
- Card holder - type (Gas, bank, premium, luxury, prepaid, etc.)
- Frequent credit card user
- New retail card holders
- Underbanked or "thin file"
- Stocks or bonds
- Average online purchase
- Average offline purchase[8]

We reproduced this long (and yet still incomplete) list of data categories to illustrate our point about the range of consumer data bought and sold every day in the commercial marketplace. The WPF scoring report focuses on the use of consumer data for consumer scoring, but lists reflecting each specific category of data are also available for purchase in many other ways. For example, list brokers offer to break down their lists by offering "selections" from the entire list. You can buy a list of individuals by category but limited by age, ethnicity, zip code, and many more characteristics. Thus, you can buy a list like this:

[8] Id. at 33-38.

Hispanic Pay Day Loan Responders Mailing List

These Hispanic Cash Advance/Payday loan seekers need cash quickly to pay monthly bills such as car payments, utility bills, credit card payments, etc., before they receive penalties for late payment.

Get Count | Get Pricing | Get More Information

SEGMENTS		*COUNTS THROUGH 12/10/2013*
750,900	TOTAL UNIVERSE / BASE RATE	$95.00/M
	PHONE NUMBERS	$200.00/M

DESCRIPTION

These Hispanic consumers reached out to secure a loan quickly and easily in order to :

- Avoid bounced check fees.
- Pay monthly bills before they become penalized.

These Hispanic Cash Advance/Payday loan seekers need cash quickly to pay monthly bills such as car payments, utility bills, credit card payments, etc., before they receive penalties for late payment.

In order to qualify, a consumer needs to have a checking account, be currently employed, have a working telephone number, be at least 18 years old, and a U.S. resident. A poor credit rating will not necessarily inhibit consumers from qualifying for a loan.

These proven direct response applicants are great prospects for secured and unsecured credit offers, pay day loan offers, instant financing programs, debt consolidation, seminars, money making opportunities, sweepstakes, at home educational programs, distance learning, and career oriented offers.

MARKET:	CONSUMER
CHANNELS:	
SOURCE:	COMPILED LISTS, DIRECT RESPONSE
PRIVACY:	UNKNOWN
DMA?:	YES - MEMBER
STATUS:	PREFERRED PROVIDER
GEO:	USA
GENDER:	45% FEMALE 45% MALE

SELECTS	
AGE	$10.00/M
GEO/GEOGRAPHICAL	$7.00/M
SCF	$7.00/M

ADDRESSING	
KEY CODING	$3.00/M
E-MAIL/FTP	$50.00/F

RELATED LISTS

- DOWN ON THEIR LUCK
- CREDIT CARD TRANSACTION SOURCE
- DEBT CRUNCH FROM NVP
- HOT TREND DRTV PRODUCT BUYERS
- HISPANIC CONSUMERS WHO ARE UNDERBANKED
- SEASONS GREETINGS CHRISTMAS SHOPPERS
- CREDIT REPAIR SOLUTIONS
- ELECTRONICS - NEW GADGET AND ELECTRONICS BUYERS
- MANAGE MY DEBT
- AMERICAN CONSUMERS SHOPPING FOR COOKING PRODUCTS

Let's take this example and apply it to the possible disclosure of HMDA data. Data fields that are disclosed include age (but not birth date), ethnicity, and geography. If the data shows that a 35 -year -old Hispanic male and a 33 –year- old White female living in a particular zip code (or other defined geographic area) applied for and received a mortgage, it would not be hard for a data broker with a list of all households in the United States to identify the specific individuals in a particular household given a few data points. It's simply a matter of matching the available information to the databases already in private hands to find the two individuals. In some cases, there might be more than one household meeting the initial search criteria. Adding a few additional data points from publicly available real estate records will likely identify the specific household with precision.

The capability of the consumer data industry to compile and use data is vast, and it would be a mistake to underestimate this capacity. The WPF consumer scoring report discusses the capabilities of one company as an illustration:

> Versium is an exemplar of the trend of smaller analytics companies competing in a space that used to be reserved for much larger companies. Versium uses 300 billion attributes from over 8,000 compiled lists covering up to, for example, 410 million unique emails, 240 million records of demographic data, 90 million mobile phone numbers, 120 million land line numbers, and 1.6 billion records of address history trail, among others.[9]

This brief description of the consumer data industry illustrates the background into which HMDA data is disclosed. Data brokers and resellers gobble up every consumer database placed in the public domain by local, state, and federal agencies. They sort information and categorize people into individual and household profiles. Their files are supplemented by data from merchants. Data also comes from consumers, sometimes through deceptive practices that induce consumers to disclose sensitive personal information in exchange for vague promises of "coupons" and without revealing that data disclosed will be compiled and sold. This is how consumer scores can be compiled using hundreds or thousands of data elements.

Our point is simple. Given the significant, multi-billion dollar consumer data industry already in place, it will be challenging to release more data publicly without having the data re-identified and added to existing profiles of individuals and households. Even when individual matching is not possible, consumer data companies will happily use the data to classify neighborhoods. Remember that for many marketing and increasingly other uses, precision is not important. A merchant can spam a hundred households and be profitable with a two or three percent response rate. Beyond merchants, we are uncertain how else this data could be used. Would it be used for determining someone's authentication or identification in other settings? If so, then the data would become quite important to individual consumers at the ground level of where they live their lives.

We applaud CFPB's recognition that "any privacy risks created by the disclosure of loan-level HMDA data may evolve over time." The volume of data held by the consumer data industry has expanded greatly over the past several decades. The availability of additional identifiable

[9] Id. at 52 (footnote omitted).

information from social networks, behavioral tracking on websites, cell phone tracking, surveillance cameras, facial recognition, and other commercial and technological developments has contributed to private held and unregulated consumer databases. These trends will obviously continue. Technology always outpaces the ability of regulators to adjust so it may be necessary for CFPB to anticipate developments to some extent. In factoring in legal developments, we suggest that CFPB consider that states have been much more active in enacting privacy legislation than the Congress, and we expect this trend to continue.

B. Permissible Purpose

WPF believes that any data released about mortgages will find its way into the files of the consumer data industry. It is likely to be true that HMDA data available is already being obtained and used by the consumer data industry, but the sources and methods of the consumer data industry are typically secret. Matching records without identifiers to identifiable records is, as we suggest above, not difficult. Expansion of electronic reporting, central availability for HMDA data, and other required and proposed data enhancements will make HMDA data more useful to everyone, including those who want the data to fulfill the purpose of the Act and those who want to exploit the data for commercial purposes.

We gently suggest that CFPB conduct an investigation into HMDA data usage before making a decision about what data to release so that we know just how the consumer data industry uses the data. The first step is to trace any data made public today to learn who acquires the data (directly and indirectly), and the next step is to learn who uses the data for non-HMDA purposes and how they use it.

We do not, however, oppose making data more useful or restricting its available to legitimate users. Instead, we think another approach may allow those who want the data to oversee the mortgage lending process to have all the data they need but to deny the data to other users.

The model for this is the Fair Credit Reporting Act. The FCRA limits disclosure of a credit report to those who intend to use the report for a permissible purpose.[10] This provision allows for appropriate use of credit reports. Credit grantors, insurers, and employers can seek and use credit reports. Marketers, list compilers, and data brokers cannot access or use credit reports.

HMDA should operate under a similar rule. Rather than make HMDA data publicly available in a totally unrestricted way, we propose that CFPB restrict access to those who state that: 1) they want the data for appropriate oversight of the mortgage lending process; 2) they will safeguard the data properly; and 3) they will not share the data with anyone who does not have the same purpose and who does not agree to the same terms.

A simple way to control data access is a **data use agreement**. A model for a data use agreement can be found in the privacy rule issued under the Health Information Portability and Accountability Act. Under the rule, certain disclosures of protected health information are permissible to users who have specific purposes (research, public health, or health care

---

[10] 15 U.S. Code § 1681b, http://www.law.cornell.edu/uscode/text/15/1681b. The FCRA allows disclosure under other circumstances not material here. Id.

operations) and who agree to limits on use and disclosure, and on safeguards.[11] The HIPAA requirements are essentially what we proposed in the last paragraph.

Limiting disclosure to those who have the requisite purpose and who agree to appropriate restrictions will give legitimate users the data they want in the same way that they have in the past. We do not seek to limit access or use by legitimate users in any way, other than to prevent them from sharing the data inappropriately with those who are not legitimate users. Further, when striking a privacy balance under a permissible purposes regime, the balance may change in favor of data users. It may mean that more data can be disclosed because its use will be limited.

Limitations reduce the need to worry that the data will leak into the marketing system because marketing uses will be prohibited. Also reduced is the worry that the data will end up in the hands of identity thieves and other crooks who might re-identify the data (perhaps using databases bought from commercial data brokers) and use the information for nefarious purposes that will harm individuals as well as mortgage lending institutions and purchasers of mortgage securities.

The administrative cost of permissible purposes controls will be small. A standard data use agreement covering all users will take an hour or two to draft. Users will have to check a few boxes and sign the form before they obtain access to the data they need. Some agency oversight will be appropriate, but likely very little. The simplest penalty for violating a data use agreement can be termination of access to the data. That is likely to be enough of a deterrent in most cases. Stronger penalties can be provided for egregious cases involving commercial exploitation.

If CFPB feels obliged to make some HMDA data fully public, we suggest in the alternative that CFPB apply the strictest pro-privacy balancing test that it can, modifying and withholding data to the maximum extent. Those who want more data than CFPB releases in this way should be required to have a permissible purpose and to sign a data use agreement as outlined above. This dual approach allows CFPB to satisfy any obligation for publication while still making more useful data available to public sector and private users who want the data for HMDA and HMDA-related activities.

III. Age vs. Date of Birth and other Data Adjustments

The Bureau proposes that mortgage applications show the **age of the borrower** rather than the **date of birth**. Date of birth is a powerful non-unique identifier that sorts individuals into a large number of bins. When date of birth is used in conjunction with five digit zip code and gender, 87% of Americans may be uniquely identified.[12] If age rather than birth data is combined with five digit zip code and gender, uniqueness drops to below 1 percent.[13]

The WPF agrees strongly that not disclosing data of birth will help to protect the privacy of applicants. We think, however, that if the information is to be publicly released, age should be

[11] 45 C.F.R. § 164.514(e).
[12] See Latanya Sweeney, Simple Demographics Often Identify People Uniquely (2000), (Carnegie Mellon University, Data Privacy Working Paper 3), http://dataprivacylab.org/projects/identifiability/paper1.pdf.
[13] See Id.

further restricted. An age range for applicants – e.g., age 21 to 25, 26 to 30 – would serve the privacy protection objective better. Age ranges would not diminish the value of the data for mortgage lending oversight. We suspect that even broader ranges – e.g., age 21 to 30, 31 to 40 – might be adequate for mortgage oversight purposes. However, we cannot support that suspicion with hard data, and we leave the suggestion for CFPB to consider.

We found amusing the reported comments by some in the mortgage lending business that calculating age from a date of birth was difficult. We point to http://www-users.med.cornell.edu/~spon/picu/calc/agecalc.htm, a website that does the calculation online at no charge. It takes only a few seconds to enter the birth date into the calculator and obtain the answer. Further, a spreadsheet formula that accomplishes the purpose is widely available on the Internet, including here: http://www.extendoffice.com/documents/excel/853-excel-convert-birthdate-to-age.html. Any financial institution that lends money with 30-year amortization tables must have employees capable of the simple function of translating a date of birth into a current age.

We have a further point. In the previous section of our comments, we recommended limiting the disclosure of HMDA data to those who have a permissible purpose relating to mortgage lending oversight. If that suggestion is adopted, then concerns about HMDA data leaking into consumer data and marketing files will diminish, if not disappear altogether. The balance between privacy and utility may tip in the utility direction because the privacy concerns that come with unrestricted publication diminish. If so, then age in narrower rather than broader ranges might be included in HMDA data, and if particularly useful to legitimate users, year of birth.

The same modification of itemized information that the statute[14] allows may be needed for other data fields. Race and ethnicity data may allow identification of individuals in cases where combinations are less common or when two individuals are co-borrowers, each with a distinct and somewhat less common racial or ethnic background. When a Caucasian individual of Anglo-Canadian-Native American ancestry is the co-borrower of an African-American of Kenyan-Scottish-Irish ancestry, the pair's combination of racial and ethnic characteristics may well be as unique as a Social Security Number.

We are aware, of course, that racial and ethnic data are vitally important to mortgage lending oversight, and we only recommend modification of data in cases where the combinations of racial and ethnic data by borrowers are unusual <u>and</u> if the data will be made public in an unrestricted way. If CFPB chooses to follow our suggestion of limiting disclosures to those with permissible purposes who agree to restrict other uses and disclosures, then we are less concerned about the possibility of identification of individuals and privacy consequences. Under these circumstances, the balance between privacy and utility will be different.

IV. Identifiers

CFPB's proposal discusses the possible creation of several new identifiers for use in connection with mortgages. One is a universal mortgage identifier (UMID), and the other is a parcel identifier. We understand that these identifiers may be useful in various ways, and we have no

[14] 12 U.S.C. § 2803(h), http://www.law.cornell.edu/uscode/text/12/2803.

comments about the ultimate value of the identifiers. We do not know enough specifics about the use of existing identifiers in the mortgage loan process, but our concerns about privacy grow as identifiers expand in use, become part of a unitary identification program, or spread to other activities beyond mortgage loans.

Our concerns about the creation of new identifiers – even ones only partially connected to individuals – have to do with secondary consequences for privacy. Everyone knows how the Social Security Number – originally expressly limited to Social Security purposes – eventually became a universal identifier for multiple purposes, with many negative consequences for individuals and for the economy at large. For decades, new uses of the SSN proliferated without restriction. It was only in recent years that governments began to understand the consequences of SSN misuse and enacted some restrictions. Private uses of SSN remain largely unrestricted today. We observe that the massive losses due to or aided by SSN misuse do not just fall on individuals but harm banks, credit grantors, and many other institutions as well.

We reviewed the Treasury study on the universal loan identifier.[15] We appreciate that the author of the paper considered the privacy consequences of a new identifier, although we would have been happier had the paper addressed current privacy standards, such as Fair Information Practices,[16] in addition to its discussion of older notions about privacy. We agree generally with this conclusion from the paper:

> [A]ny potential UMID [universal mortgage identifier] 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.
>
> 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.[17]

Of course, we agree with the paper's recommendation that no information should be encoded in the identifier.[18] We see that as essential. However, we are not convinced that the paper's proposal that a UMID be publicly available with no restrictions on the use of the identifier.[19] If

[15] http://www.treasury.gov/initiatives/ofr/research/Documents/OFRwp0012_McCormickCalahan_CommonGroundNeedforUniversalMortgageIdentifier.pdf.
[16] See generally, Robert Gellman, Fair Information Practices: A Basic History, http://bobgellman.com/rg-docs/rg-FIPshistory.pdf.
[17] Id. at 12-13.
[18] Id. at 13.
[19] Id. at 14.

the UMID is freely available for unrestricted use, then we fear that shortly after a universal loan identifier comes into common use, a data broker will find a way to add the number to the profile maintained on individuals and households. The next step will be a data file that provides a crosswalk between a loan identifier and the Social Security Number of each borrower. Once the UMID becomes available in this way, other users and other uses will develop. Individuals will likely become unable to participate in any credit activity without having to disclose his or her UMID. This would be a very poor privacy outcome.

The second identifier discussed by CFPB is a parcel identifier. We see the privacy issues here as possibly less consequential for individuals, although it is true that an individual is associated with any real property that the individual owns. We also recognize that precise identification of real estate by traditional methods as well as geocoding will continue. Individual are traditionally associated with a street address, and we are not sure that a formal parcel identifier will be meaningfully different from a street address. However, we observe that there are some laws protecting victims of domestic violence that allow individuals to shield their street addresses in driver's licenses, voting records, and other court and government documents. Some consideration should be given to how a parcel identifier might affect these programs. Our concern is the same as with the UMID. Any identifier that is associated with an individual will attract other uses and users. In the case of domestic violence victims, protections built into other programs may disappear if individuals are readily linkable with parcel identifiers. Even those who do not own real property may be associated with a parcel identifier for their rented home or apartment. This aspect of any identifier requires much more attention.

We offer three suggestions for any new or expanded identifiers:

1. No linking

First, there should be nothing in the actual identification number that can be linked in any way to an identifiable individual. That means no names, SSN, or other identifiers. We support CFPB's proposal to prohibit using information that could be used to identify directly the borrower or applicant.

2. Controls on collection

Second, there must be controls on the collection, use, and disclosure of any new identifiers. If a new identifier will be readily available for private use unrelated to the specific purposes that justify creation of the identifier, then we oppose the creation of the identifier. We recognize that the standard we suggest will not be easy to satisfy. It may take legislation, contracts, data use agreements, or all of the above to control unwanted use of new identifiers. Without some restrictions, a new identifier – and particularly a UMID – will just become another field in the individual or household profiles maintained by an untold number of data brokers.

3. Address vulnerable population issues before use – including victims of domestic violence

Third, the consequences of a new identifier for victims of domestic violence and others who have special confidentiality needs must be considered and address before a new identifier comes into use.

Of our three standards, we believe that the first will be easy to satisfy. The third will take time and effort, but it should be possible to accomplish the goals in the same way the other identification programs have adjusted to the needs. Our second standard – preventing secondary use – may be much harder to accomplish without legislation. That is another reason why secondary use restrictions must be considered at the start of any new identification program.

We are uncertain about the value of hashing identifiers. We think it would be much better if an identifier included no personally identifying information that needs to be masked by hashing. Hashing might have some value in allowing the hashed number to be used for some purposes and the unaltered number for other purposes. But if there is a possibility that someone can produce a crosswalk between the two numbers, there is nothing to be gained.

V. Cost to Consumers

In various places, CFPB discusses the possible cost to consumers of various proposals. While we do not take issue with any of the narrow conclusions drawn, we observe that had anyone considered the cost to consumers of issuing SSNs in 1935, the conclusion would most likely have been that consumers would bear no costs. Yet over time, choices made about the use of SSNs and the secondary purposes to which SSNs have been put suggest that the cost to consumers was enormous. Identity theft losses alone are measured in the billions. We caution against casual judgments when the creation of identifiers is at issue and when the public disclosure of personal data is under consideration. CFPB must anticipate how data might be used by the consumer data industry and by crooks worldwide for purposes that will impose financial costs, privacy harms, or other costs on consumers.

VI. Conclusion

Thank you for the opportunity to comment on a proposal that is important to privacy and stands to impact millions of people over time. We urge CFPB to strongly consider limiting the release of data to the public as described in our comments, and if release of the data is contemplated, then we urge CFPB to put substantive and meaningful controls in place. We also strongly recommend that unique mortgage identifiers are not linked, have meaningful controls on collection, and that the CFPB carefully assesses how the identifiers can be risk factors for vulnerable populations such as victims of domestic violence or crime. Thank you for the thoughtful approach to the NPRM generally. WPF stands ready to answer any questions you may have regarding our comments, or the ideas we have discussed here.

Respectfully submitted,

Pam Dixon

Executive Director, World Privacy Forum
www.worldprivacyforum.org