I/S: A JOURNAL OF LAW AND POLICY FOR THE INFORMATION SOCIETY

Is Financial Privacy Preventing Legitimate Research?

PETER WILLIAMS*

Abstract: The United States may be facing its greatest economic crisis since the Great Depression. Understandably, society wants to learn from this crisis to ensure that the United States will not repeat the same mistakes that lead to the current situation. Obtaining the answers undoubtedly starts with research. But, what if the researcher's methods tread on financial privacy? Society will have solved one problem only to create another. In this note, I use three steps to show that financial research can be expanded without subtracting from financial privacy. First, I assert that current financial privacy regulations have unnecessarily burdened and impeded academic research. To reach this assertion I use a real world example of researchers studying the mortgage meltdown. Second, I demonstrate that it is possible to uphold financial privacy rights and advance research. I do this by examining research and privacy in other contexts, specifically the HIPAA research exception. Finally, by learning from HIPAA and by utilizing current data collection mechanisms already in place, I suggest there are policy approaches to allow for both a high level of financial privacy and access to valuable data.

I. INTRODUCTION

In an economic downturn, there is a tendency to explain the causes and potential implications of the financial crisis. This research is important as it could lead to a better understanding of what went wrong, and therefore provide policy suggestions to avoid another

* Peter Williams is a J.D. candidate at The Ohio State University Moritz College of Law and is a member of *I/S: A Journal of Law and Policy for the Information Society.*

crisis. However, this research is time consuming and often lacks data because of certain regulatory barriers that limit researchers' access to data.

This note documents the difficulty in gathering financial mortgage data. Part II examines the National Consumer Law Center's attempt to gather financial data, showing that there is a gap in the market between data that needs to be gathered and data that is available. Part III looks at the current regulatory framework and databases in place that researchers use to gather financial data. Part IV examines the Health Insurance and Portability Assurance Act ("HIPPA"), and uses the Act as an example of a research exception to regulatory requirements. Finally, the note focuses on potential policy options that would both protect financial privacy and expand the data that is realistically available to the researcher.

II. FINANCIAL DATA COLLECTION IN PRACTICE

The National Consumer Law Center ("NCLC") advocates for consumer protection laws on behalf of low income and vulnerable individuals.[1] The NCLC monitors mortgages and foreclosures in poor neighborhoods and cities. Specifically, the NCLC is determined to curb predatory lending, which "disproportionately affects minority homeowners and communities."[2]

To examine the issues surrounding mortgages in low-income areas, the NCLC has attempted to gather mortgage data in order to better illustrate and document the adverse effects of predatory lending. From 2006 to 2009, the NCLC gathered home data from roughly three thousand mortgages in the hopes of creating a dynamic and extensive data bank. The NCLC maintains a project called the National Mortgage Data Repository ("NMDR") and expects the NMDR to be the authoritative source on predatory lending and sub-prime loan terms.[3]

While the goals of the NMDR are laudable, this note is concerned with how the NCLC gathers the data necessary to complete the

---

[1] National Consumer Law Center, About Us, http://www.consumerlaw.org/about/index.shtml (last visited April 8, 2010).

[2] National Consumer Law Center, Predatory Mortgage Lending, http://www.consumerlaw.org/issues/predatory_mortgage/index.shtml (last visited April 8, 2010).

[3] National Consumer Law Center, National Mortgage Data Repository, http://www.consumerlaw.org/issues/repository/index.shtml (last visited April 8, 2010).

NMDR. This note will use the NCLC as a case study to illustrate the regulatory framework associated with financial research. To understand and determine how the NCLC obtains its data, I interviewed Elizabeth Renuart, an attorney working in the NCLC Boston office.[4]

According to Ms. Renuart, the NCLC wanted to analyze mortgage information by zip code. The majority of the data gathered independently by the NCLC came from attorneys who represented consumers in predatory lending class actions or related suits. With their clients' consent, attorneys would redact information and provide it to the NCLC. Other sources of data include housing councils, state attorney generals, and consumers themselves. Notably absent from the source list are the financial institutions that issue the loans, because certain regulations restrict them from releasing financial mortgage data.

In addition to the direct gathering of data, the NCLC may also rely on data gathered under the Home Mortgage Disclosure Act ("HDMA"). This database provides a wide range of information with a specific focus on monitoring low-income lending patterns. This information is then combined with the directly gathered information, and, if enough data can be obtained, the NCLC will have a comprehensive database.

In over three years of gathering data, the NCLC was only able to collect roughly 3,000 mortgage applications out of the hundreds of thousands in existence. This figure is low considering that the NCLC has multiple researchers working on the project, and while their budget is not immense, it is likely larger than a single academic working on a project. This low figure indicates that there could possibly be a policy problem that is stymieing research, and without financial institution cooperation, the NCLC is restricted in what it can reliably retrieve. Using the NCLC example as a backdrop, this note examines the financial data regulations and the financial privacy restrictions in place to protect consumers and tries to determine if policy changes could aid research. This note concludes that the regulations in place are weighted too much in favor of financial privacy and unnecessarily burden legitimate research.

---

[4] Telephone Interview with Elizabeth Renuart, Attorney, National Consumer Law Center (Mar. 11, 2009) (No longer with the NCLC, Miss Renuart is currently an assistant professor of law at Albany Law School).

III. COLLECTING FINANCIAL DATA

This section will detail financial data collection and the impediments researchers face in gathering data. The first part will examine what data is publicly available to researchers under the current regulatory framework and the potential problems with this publicly available data. Then, having examined how data is gathered through public databases, the second part will examine how researchers can gather data privately and the subsequent regulatory barriers that can limit collection of data.

When researching financial data, the researcher would first consult the publicly available HMDA data, and then if unsuccessful, the researcher may privately collect the data. Private data collection could involve using a private company if the researcher has funds, going directly to the institutions and asking for the data, or contacting other sources such as plaintiffs' lawyers or individual homeowners.

A. PUBLICLY AVAILABLE DATA

The government collects a large amount and a wide range of mortgage data. In 1975 Congress enacted the Home Mortgage Disclosure Act to combat the perceived lack of lending to less fortunate geographic areas.[5] By accumulating data, Congress can monitor trends and determine if lenders are unfairly under-serving certain areas.[6] Specifically, Congress wanted to ensure that minorities were not being denied access to loans.[7] These priorities are set out in Regulation C and largely determine how the government proceeds in collection data.

In the HMDA, Congress gave The Federal Reserve Board of Governors the authority to issue regulations concerning what data lenders must report to the federal government. Regulation C contains the guidelines.[8] Under Regulation C, the lender is required to submit a loan application register ("LAR") once a year, and for each loan application include: (a) the loan type and amount, (b) the property,

---

[5] 12 U.S.C. § 2801(a) (2006).

[6] *Id.* at § 2801(b).

[7] Alicia H. Munnell, Geoffrey M. B. Tootell, Lynn E. Browne & James McEneaney, *Mortgage Lending in Boston: Interpreting HMDA Data*, 86 AM. ECON. REV. 25 (1996).

[8] 12 C.F.R. § 203.1(a) (2009).

such as location and type, (c) the disposition of the application, denied or approved, and (d) the applicant's ethnicity, race, sex, and income[9]

The information is then compiled by the Federal Financial Institutions Examination Council ("FFIEC") and released to the public in an aggregate form by institution, county, or metropolitan statistical area.[10] While the data is rich, and getting richer after Regulation C[11] was implemented in 2004, it is not sufficiently broad to allow for a wide range of economic studies.

B. INDIVIDUAL AND PRIVATE RESEARCH EFFORTS

While the LARs can provide a wide range of data, they do not provide every piece of information that researchers need. While this note does not specifically address what data researchers would want, it is assumed researchers would like to gather more individualized and specific data not found in LARs. Essentially research is necessarily bounded by whatever is in the LARs, and this reduces the range of studies and data collection that can be completed. Moreover, researchers cannot collect data from the largest holder of financial data: financial institutions. As a practical matter the banks control the data and, due to privacy laws, cannot release the data, even to a well-intentioned researcher. Further, the banks view mortgage information as proprietary and even without privacy laws would be unlikely to part with the data. Therefore, the most direct method a researcher can employ is to directly contact individual mortgage holders and attempt to gather information that way.

Financial institutions can use the privacy restrictions enacted in the Gramm Leach Bliley Act as a reason to deny information. Congress passed the Gramm Leach Bliley Act ("GLB") in 1999, "[t]o enhance competition in the financial services industry by providing a prudential framework for the affiliation of banks, securities firms, insurance companies, and other financial service providers."[12] The

---

[9] FED. FIN. INSTITUTIONS EXAMINATION COUNCIL, A GUIDE TO HDMA REPORTING, GETTING IT RIGHT! p. 8 (June 2008), available at http://www.ffiec.gov/hmda/pdf/2008guide.pdf.

[10] *See* Federal Financial Institutions Examination Council, Reports, http://www.ffiec.gov/reports.htm.

[11] For more information about regulation C., *see* FED. FIN. INSTITUTIONS EXAMINATION COUNCIL REGULATION C-HOME MORTGAGE DISCLOSURE (2004), *available at* http://www.ffiec.gov/hmda/pdf/regulationc2004.pdf.

[12] Pub. L. No. 106-102, 113 Stat 1338 (1999).

Act created strict and precise financial privacy regulations, and placed rulemaking authority with a number of federal regulatory bodies.[13]

The GLB Act requires "financial institutions" to adequately safeguard the nonpublic information of their customers.[14] "Financial institutions" is an inclusive term, and if an institution engages in commercial and consumer lending, it is almost certainly a financial institution, and thus under the purview of the GLB Act.[15] Additionally, the GLB Act distinguishes between customers and consumers. A consumer has a one-time relationship with the financial institution, whereas a customer is a consumer who has repeated long-term business transactions with the institution.[16] A mortgage qualifies the consumer as a customer, and the mortgagee is therefore governed by the financial privacy protections of the GLB Act.[17]

As a customer of a financial institution, the mortgage holder must first be directly provided a privacy notice, which details the institution's general privacy policies.[18] This notice informs the customer of the manner in which the institution will use the customer's non-public information.[19] If the institution plans to distribute non-public information to non-affiliated third parties (i.e. researchers), the customer must be provided with express notice of the disclosure[20] and the opportunity to opt-out of the disclosure.[21]

A number of exceptions allow financial institutions the ability to disclose non-public information without facing repercussions. The Act applies only to "non-public" information, not to information that is available through public records such as mortgage data that is

---

[13] 15 U.S.C. §§ 6801-09; 15 U.S.C. § 6804(a)(1); *See* Peter Swire, *The Surprising Virtues of the New Financial Privacy Law*, 86 MINN. L. REV. 1263 (2002).

[14] 15 U.S.C. § 6801(a)(2006).

[15] 12 U.S.C. § 1843(k)(4)(A)(2006).

[16] FEDERAL TRADE COMMISSION, IN BRIEF: THE FINANCIAL PRIVACY REQUIREMENTS OF THE GRAMM LEACH BLILEY ACT (2009), *available at* http://www.ftc.gov/bcp/edu/pubs/business/idtheft/bus53.pdf.

[17] *Id.*

[18] 15 U.S.C. § 6803(a)(2006).

[19] *Id.*

[20] *Id.* at § 6802(b)(1)(A)(2006).

[21] *Id.* at § 6802(b)(1)(B)(2006).

included in tax records.[22] Also, the customer has no opt-out provision for certain types of data such as data collected by regulatory agencies within their powers.[23]

However, there is no research exception. If the financial institution agrees to provide information, it must first notify customers and allow them to opt out. The costs associated with notification, mainly postage fees, are a disincentive for companies to agree to the research, and even if the company agrees to disclose information, potential subjects may opt out.[24] Therefore, by conveniently claiming GLB Act compliance costs, an institution can deny information requests.

Because researchers cannot use the institutions, they must gather any information not contained in the LARs from individual homeowners or other people who are permitted to release the information on their behalf. This includes obtaining information from attorneys who represent homeowners in class action suits or attorney generals who are filing suits against lenders. While this direct method of gathering data can produce results, the researcher is not likely to be able to gather a broad and comprehensive group of data. A researcher wishing to explore a new area of research or a new line of data indicators will be hard pressed to gather a large sample size because, if HMDA does not have the necessary data, the researcher must then devote his energy to data collection, which may or may not actually produce results.

IV. THE HEALTH INSURANCE AND PORTABILITY ASSURANCE ACT

When discussing financial research, it can be helpful to analyze research in other academic fields. Medical research is similar to financial research in that gathering data often conflicts with personal privacy concerns. However, in the medical context the regulatory framework clearly creates a niche for research, while in the financial context research does not enjoy special privacy considerations. This section examines the medical research privacy exceptions and outlines

---

[22] FEDERAL TRADE COMMISSION, *supra* note 16.

[23] *Id.*

[24] For example, if a bank allowed a researcher access to 10,000 mortgages the bank would have to send a notice to each of the 10,000 customers affected. At forty-four cents postage per notification it would cost the bank $4,400 to send out the notices. The bank, which already wants to keep the information proprietarily, likely would not want to incur this cost, and the researcher may not have the funds to pay the cost.

the different regulatory hurdles a researcher must go through to complete his or her research, because these medical research procedures could provide a template for developing financial research procedures.

A. THE PRIVACY RULE AND RESEARCH: RESEARCH PROCEDURES

As technology advanced in the period leading up to the mid-1990's, the ability of the medical profession to digitally collect and more efficiently manage patient records dramatically improved.[25] These increased abilities created greater privacy concerns for sensitive patient medical records.[26] Responding to these concerns, Congress enacted the HIPAA in 1996.[27] Congress granted the Department of Health and Human Services ("HHS") the right to "promulgate final regulations containing standards [governing privacy of individually identifiable health information] not later than the date that is 42 months after the date of the enactment of this Act," so long as Congress did not first legislate the standards within thirty-six months of enacting HIPAA.[28] When Congress failed to act, the Department of Health and Human Services issued an initial "Privacy Rule" in 2000, and updated the rule in 2002 to clarify ambiguous wording and correct unintended administrative burdens.[29]

The Privacy Rule is designed to provide federal standards to protect personally identifiable health information.[30] Personally identifiable health information is classified by the following:

> (1) Is created or received by a health care provider, health plan, employer, or health care clearinghouse; and

---

[25] Randolph C. Barrows Jr. & Paul D. Clayton, *Privacy, Confidentiality, and Electronic Medical Records*, 3 J. AM. MED INFORMATICS ASS'N 139 (1996).

[26] *Id.*

[27] Health Insurance Portability and Accountability Act of 1996, Pub. L. 104-191, 110 Stat. 1936 (2006).

[28] *Id.* at § 264(c)(1).

[29] *Protecting Personal Health Information in Research: Understanding the HIPAA Privacy Rule*, National Institutes of Health, *available at* http://privacyruleandresearch.nih.gov/pr_02.asp (Last visited April 8, 2010).

[30] *Id.*

> (2) Relates to the past, present, or future physical or mental health or condition of an individual; the provision of health care to an individual; or the past, present, or future payment for the provision of health care to an individual; and
>
> (i) That identifies the individual; or
>
> (ii) With respect to which there is a reasonable basis to believe the information can be used to identify the individual. [31]

A subset and more heavily regulated type of personal information is protected health information ("PHI"). PHI includes only the personally identifiable health information that has been transmitted electronically, maintained electronically, or transmitted or maintained by any other medium.[32] As a general rule, PHI may only be disclosed or released by the express authorization of the patient.[33] HHS has enacted regulatory "core elements," which every authorization must include, and these elements must be in plain language.[34]

In general, patient notice and consent must be obtained before PHI may be disclosed, but there are specific circumstances where PHI may be disclosed without a waiver. It may be impractical for researchers conducting certain types of research to obtain consent from every patient involved. In these circumstances the researcher may submit a proposed project to an Institutional Review Board ("IRB").[35] An IRB is a committee designated by an institution to review proposed research on human subjects.[36] The IRB is the institutional gatekeeper that must balance the need for research with the privacy of the patients. The IRB is the institutional gatekeeper

---

[31] 45 C.F.R. § 160.103 (2009).

[32] *Id.*

[33] 45 C.F.R. § 164.508(a)(1) (2009).

[34] *Id.* at § 164.508(c)(1), (3).

[35] *Institutional Review Boards and the HIPAA Privacy Rule*, DEPARTMENT OF HEALTH AND HUMAN SERVICES, *available at* http://privacyruleandresearch.nih.gov/pdf/IRB_Factsheet.pdf.

[36] *Id.*

that must balance the need for research with the privacy of the patients. Before research is approved, HIPAA regulations require the IRB to determine that numerous requirements have been met.[37] Specifically, the IRB must determine that there are proper safeguards for protecting PHI and that the research would not be possible without the PHI.[38] A complete waiver is granted when locating the patients and gaining their consent would be impractical.[39] A partial waiver is granted when a researcher needs PHI for one specific aspect of research, for example to locate possible subjects.[40] Once the waiver is granted, the covered entity may then disclose the PHI to the researcher.

B. EFFECTS OF THE ACT

Medical researchers have attacked the HIPAA Privacy Rule as an unnecessary measure that impedes research. One critique is that the rule is rather complicated, and therefore research institutions have to devote large amounts of time and resources to ensure they are within the guidelines.[41] For example, one recent study conducted by Dr. Robert Ness questioned 1,527 epidemiologists about the effect HIPAA was having on their research.[42] Roughly sixty-eight percent of the respondents claimed the Privacy Rule made research harder, and forty percent indicated that the Privacy Rule made research more expensive.[43] Additionally, half of the respondents claimed the Privacy

---

[37] *See* NIH Pub. No. 03-5448, Privacy Boards and the HIPAA Privacy Rule, DEPARTMENT OF HEALTH AND HUMAN SERVICES (Sep. 2003) *available at* http://privacyruleandresearch.nih.gov/pdf/privacy_boards_hipaa_privacy_rule.pdf.

[38] *Id.* at 3.

[39] 45 C.F.R. § 164.512(i). *See* NIH Pub. No. 03-5448, Privacy Boards and the HIPAA Privacy Rule, DEPARTMENT OF HEALTH AND HUMAN SERVICES (SEP. 2003), 3 *available at* http://privacyruleandresearch.nih.gov/pdf/privacy_boards_hipaa_privacy_rule.pdf.

[40] 45 C.F.R. § 164.512(i).

[41] Jennifer Kulynych & David Korn, *The New HIPAA (Health Insurance Portability and Accountability Act of 1996) Medical Privacy Rule*, 108 CIRCULATION: J. OF THE AM. HEART ASS'N 912, 914 (2003).

[42] Robert Ness, *Influence of the HIPAA Privacy Rule on Health Research*, 298 JAMA 2164-70 (2007), *available at* http://jama.ama-assn.org/cgi/reprint/298/18/2164.

[43] *Id.* at 2166.

Rule added to the time needed to complete research.[44] Perhaps the most alarming finding was that only twenty-five percent of the respondents felt the Privacy Rule had enhanced patient privacy.[45]

V. POLICY IMPLICATIONS

The most obvious policy option for expanding research data would be to expand the information collected under the HMDA. This would make a broad range of data available to all researchers and the public at large. Lending institutions already collect and send copious amounts of data, and the institutions have collected more data than they disclose. Further, the institutions already are familiar with redacting and protecting personal information, and therefore the misuse of personal financial data would be unlikely.

While expanding the range of data collected under the HMDA would expand the information available to researchers, there is an implicit line that needs to be respected as expanding data collection would give the government large amounts of personal financial information, creating data security concerns. Additionally, the data collected under the HMDA cannot possibly suit every researcher's goals. Different researchers will inevitably want to gather information that is either different or independent of the HMDA data. Therefore, while expanding the data collection powers of the HMDA will be beneficial, it alone is not sufficient to expand financial research.

To complement the expansion of the data collected under the HMDA, I would propose creating a research exception to the GLB Act, modeled after the research exception in HIPAA. HIPAA attempts to balance the need for research with the need for privacy. However, medical researchers have lamented that the balance is uneven and has adversely affected their research. Building on this criticism and learning from it, the financial privacy exception could implement some aspects of HIPAA while still maintaining efficient research.

Congress should enact federal regulations that financial researchers would have to abide by, essentially making the researchers accountable in the event that they misuse financial data. This would include making researchers become certified or trained to understand the importance of financial privacy. Secondly, institutions would disclose in their privacy notices that information could be released to third party researchers who agree to abide by the federal regulations

---

[44] *Id.*

[45] *Id.*

or have been certified. Then the "accredited" researcher would submit his research proposal to the institution. Having determined the researcher is accredited the financial institution could release the information. Finally, any information that is released would have to be sufficiently redacted to ensure the data could not be traced back to the customer.

This policy suggestion assumes that financial institutions will give their financial data to a researcher. The medical research exception in HIPAA exists because covered entities wanted the exception to allow researchers to develop new medical technologies. The exception benefits the entities and the researchers; however, that is not the case with financial institutions. The goals of the researchers may diverge from the goals of the financial institutions. Most institutions have their own in-house research departments, and academic research is not likely to benefit those institutions. A research exception might exist, but would be useless if institutions do not want to disclose the data. Therefore, regulations would need to provide a minimum amount of information that financial institutions would be required to disclose in order to allow proper research.

Also, in the medical context there already was cooperation between researchers and entities. The HIPAA Privacy Rules added complexity to an already existing relationship. One benefit of a financial privacy exception is that a privacy research exception cannot impede and frustrate researchers, because there was no relationship to begin with.

In order for the exception to operate, institutions will have to disclose data. This could be accomplished through regulations. Congress could use the leverage created by the recent financial crisis and bailouts to implement requirements that financial institutions release data to certain researchers. The regulations could require institutions to release information only to not-for-profit researchers, or only to certified researchers. Further, the regulations could limit the range of data the researchers can ask for to avoid bogging the institution down with limitless requests.

Perhaps instead of creating a research exception in the GLB Act, researchers could go through the office that compiles the HMDA data, and ask them to collect more specific data from certain institutions. The office could evaluate the merits of the research and the risks to financial privacy, and then decide whether to collect the data or deny the request. In a sense, the office would be acting as an IRB through which financial research could be conducted and monitored to ensure privacy compliance.

A final policy suggestion is to create a standard contract with which researchers could enter into agreements with institutions. If

the government was to approve a standard contract that provides for data security and financial privacy, researchers could enter an agreement, gather data, and, in the event of the misuse of data, be contractually liable to the financial institution or customers.

All of these policy suggestions attempt to expand research while at the same time limiting the potential for financial privacy misuse. In limiting the financial privacy misuse possibilities, it is important not to over-regulate, making research needlessly complicated and time consuming.