ARTICLE

THE IMPACT OF DATA MINING ON INFORMATION DISCLOSURE BY REGULATORY AGENCIES: WITH AN APPLICATION TO REDLINING

W.C. BUNTING, J.D., PH.D.*

Data mining techniques can be used to locate statistical outliers that are improperly characterized as evidence of unlawful conduct. Using home mortgage loan data made publicly available by financial regulators, a simple data mining exercise finds that approximately three percent of all lender-MSA pairs (or approximately seven to nine percent of all lending institutions) flagged as having redlined minority neighborhoods is attributable to a failure to correct for the multiple hypothesis testing problem. The false positive rate does not fully explain, however, the high estimated frequency of statistical redlining. Three possible models of information disclosure by regulatory agencies are considered: (1) full information, (2) no information, and (3) limited information. Under a limited information model, litigation serves to correctly implement statistical hypothesis testing: A plaintiff must formulate a hypothesis prior to examining the data and can only obtain the information necessary to test this hypothesis through discovery.

I. INTRODUCTION

Data mining is the process of discovering reoccurring events or patterns in large datasets.[1] The objective of the data mining process is to transform complex data into an understandable structure for future applications.[2] Data mining poses problems, however, when used to locate statistical outliers in the data that are characterized as evidence of unlawful conduct. This process can lead to violations of the key statistical principle that a hypothesis should not be tested on the same data that was used to generate the hypothesis. To illuminate this principle, suppose that twenty researchers each run an experiment to test if an after-school program improves test scores, and nineteen find no significant differences in test scores between the students who participated in the program and those who did not. As a result of random noise, however, suppose that one study does find a significant correlation between the after-school program and test scores. Although these twenty studies, in the aggregate, do not provide strong statistical evidence that the program has

* The author is an Assistant Professor in the Department of Business Law of the Fox School of Business and Management at Temple University. Email: william.bunting@temple.edu. All errors should be attributed to the author alone.

[1] *See, e.g.*, Usama Fayyad et al., *From Data Mining to Knowledge Discovery: An Overview*, AI MAG., Fall 1996, at 39.

[2] *See id.*

a positive impact on test scores, the researcher of this one outlier study might, nonetheless, formulate a hypothesis suggested by the data by isolating a unique aspect of the study—for example, that the after-school program in this study was the only one to offer dance instruction—and claiming that this particular aspect explains the differing result. Using a statistical outlier in this manner to generate a hypothesis that is tested on the same dataset containing this outlier is not a proper application of the scientific method or the statistical hypothesis testing in particular—the theory should precede the empirical analysis.[3]

Given full access to an administrative dataset, it is possible to employ data mining techniques to locate a statistical outlier that can be improperly characterized as evidence in support of the hypothesis that the outlier signifies a violation of the law. For example, that a local law enforcement agency has unusually large racial disparities in arrest rates compared to all other such agencies may provide criminal justice activists opportunity to argue that this particular local law enforcement agency employs unlawful discriminatory policing practices. Likewise, that an industrial manufacturer has a statistically significant rate of workplace accidents compared to the mean rate in the industry may lead complainants to argue that the manufacturer is in violation of certain workplace safety regulations.

The example considered in this study—of a statistical outlier in the data being improperly used to generate a hypothesis regarding a potential violation of the law—is redlining by lending institutions. Redlining occurs when a lender provides unequal access to credit, or unequal terms of credit, on the basis of the race, color, national origin, or another protected characteristic of the residents of the area in which applicants reside, or will reside, or in which the residential properties to be mortgaged are located.[4] Neither the Equal Credit Opportunity Act ("ECOA") nor the Fair Housing Act ("FHA") expressly uses the term "redlining" in its text.[5] However, courts—as well as regulatory agencies that have enforcement responsibilities with respect to these statutes—have consistently interpreted them as prohibiting lenders from having different lending practices for geographic areas, relative to other comparable lending institutions in these areas, where the purpose or

---

[3] *See, e.g.*, Alan Bryman, Quantity and Quality in Social Research 20 (1988).

[4] *See* Fed. Fin. Insts. Examination Council, Interagency Fair Lending Examination Procedures iv (2009). Redlining may also include "reverse redlining," the practice of targeting certain borrowers or areas for credit on unfair terms. *Id.* at 30; *see also* Williams v. Gelt Fin. Corp., 237 B.R. 590, 594 (E.D. Penn. 1999) (stating that "reverse redlining" is the practice of "targeting of persons for 'credit on unfair terms' based on their income, race, or ethnicity") (quoting S. Rep. No. 103-169, at 21 (1993)). *See generally* Dan Immergluck, Credit to the Community: Community Reinvestment and Fair Lending Policy in the United States (2004).

[5] 15 U.S.C. § 1691 (2012); 42 U.S.C. §§ 3601–3619 (2012). The term "redlining" derives from the historical practice of some lenders evaluating loan applications by relying upon a residential map where integrated and minority neighborhoods were marked off in red as risk areas. *See* Robert G. Schwemm, Housing Discrimination, Law, and Litigation § 13:15 (2009).

effect of such difference is to discriminate on the basis of a legally protected characteristic.[6]

Pursuant to the applicable case law discussed in Part IV, this study assumes that there is statistical evidence sufficient to allege a prima facie case of redlining if the difference between the percentage of loans made in majority-minority census tracts in a given metropolitan statistical area ("MSA") and time period by a targeted lender and the percentage of such loans made by its peer lending institutions is statistically significant at a five percent confidence level or less. The dataset in this study is a panel dataset of annual, census-tract level observations starting in 2012 and ending in 2016. Each tract-level observation contains Home Mortgage Disclosure Act ("HMDA") home mortgage loan data and demographic census-tract level characteristics.[7] Simple data mining techniques are used to compare, for each lender-MSA pair in the HMDA data, the proportion of loan applications received by a targeted lender in a given MSA and time period from majority-minority census tracts to the proportion of such applications received by its peer group in that same MSA and year. This screening exercise generates a set of hypotheses (regarding potential violations of the law) that will naturally be confirmed if subsequently tested on the HMDA data used to generate this set of hypotheses.

In addition to the HMDA data, this screen is also run on a randomly generated dataset. The goal is to determine if, and how many, lending institutions can be classified as having engaged in redlining when there are no true violators in the data, *by construction*, and to compare the magnitude of the resulting percentages to those derived using actual HMDA data. To the extent that the screen identifies a subset of lenders as having engaged in statistical redlining when run on this randomly generated negative data, this subset of alleged violators is interpreted as the product of random variability, representing statistical outliers, and would not indicate true discrimination. The goal is to approximately establish what proportion of the set of flagged lenders using actual HMDA data is purely the result of improper "*p*-hacking," which can be defined as the misuse of data analysis to find patterns in data that are spuriously presented as statistically significant when, in fact, no true underlying effect exists.

Using nationwide home mortgage loan data made publicly available by financial regulators, this Article finds that approximately three percent of all lender-MSA pairs flagged by the data mining exercise, or approximately seven to nine percent of all lending institutions flagged, can be attributed to a failure to correct for the multiple hypothesis testing problem described in Part II. In addition, the findings suggest a surprisingly high degree of unlawful conduct on the part of lending institutions. For a large number of lenders, the percent of home mortgage loan applications received from majority-mi-

---

[6] *See* Robert G. Schwemm, *supra* note 5, at § 13:15.

[7] *See infra* Part III.

nority census tracts is significantly different from peer institutions—a finding shown to be highly robust to model specification in Part V.

A better understanding of why this finding is true and the extent to which these differences in lending rates are truly a product of discrimination—or some other feature of the residential mortgage market—is an important topic for future research that is not resolved here. Rather, the purpose of this study is simply to point out that, with full information disclosure by regulatory agencies, the number of targets flagged using actual data corresponds in some ratio to the number of targets flagged using randomly generated negative data. In certain situations, this ratio might be close to one, implying that random sampling error is a primary explanation for why certain firms have been flagged by a data mining exercise that simultaneously tests numerous statistical hypotheses. The rate of workplace accidents, for example, for all firms in a particular industry may be approximately equal to the significance level applied in the screen, yielding a ratio that is relatively low (and close to one). In other cases, such as with respect to alleged redlining, the estimated ratio may be relatively high (and greater than one), suggesting that a large majority of the potential targets flagged by the screen are the product, not of spurious correlation, but of truly unlawful behavior. In theory, a plaintiff will couple the necessary statistical findings with qualitative evidence, alleging, for example, that the defendant-lender has avoided providing residential mortgage loans to residents of majority-minority neighborhoods (1) by configuring the geography delineated as its Community Reinvestment Act ("CRA") assessment area to include primarily majority-minority residential areas, (2) by systematically locating its branch offices in primarily non-majority-minority areas, and/or (3) by strategically choosing to market its lending products to residents of non-majority-minority neighborhoods. In practice, however, such qualitative evidence has become less probative of redlining as a result of the increasing prevalence of digital technologies that have allowed lenders to conduct an increasingly larger percentage of business online and not at traditional brick-and-mortar locations.

A regulatory agency can choose between three general models regarding the disclosure of private information collected while operating in its supervisory capacity: (1) full information disclosure, (2) no information disclosure, or (3) limited information disclosure. Although full information disclosure tends to promote greater transparency and accountability, this Article identifies a potential problem with this model of disclosure: improper *p*-hacking by means of multiple comparisons. Using redlining as a motivating example, this study demonstrates how it is possible to use simple data mining techniques to uncover patterns in publicly available administrative data that can be spuriously presented as statistically significant. As a potential remedy, a regulatory agency might choose not to publicly disclose any (or very little) private information obtained while operating in its supervisory capacity, thereby depriving the public of the data from which the very problem of multiple comparisons arises. The primary disadvantage of this disclosure model is that institutions become severely restricted in their capacity to implement internal corporate compliance programs. Indeed, this no information disclosure model appears to represent an overly far-reaching response to a technical problem that is fixable through other methods.

Under a limited information disclosure model, regulatory agencies continue to collect data in their supervisory capacities but substantially limit the disclosure of this data to the general public. The primary advantage of this model is that lenders can still conduct internal corporate compliance, but other private actors can no longer use improper data mining techniques to uncover spurious correlations in publicly disclosed data. Interestingly, litigation plays an important role under this model, serving as a way to correctly implement statistical hypothesis testing—and the scientific method more generally. Under limited information disclosure, a potential litigant cannot p-hack publicly available data to generate a set of hypotheses that will be subsequently confirmed if tested on the same data during the course of a lawsuit.[8] Rather, a plaintiff must formulate a hypothesis (e.g., lender i has engaged in redlining in MSA j during time period t) prior to examination of the data and obtain the data necessary to test this hypothesis only through discovery. Under this disclosure model, a plaintiff cannot simply download the data and run the relevant statistical analysis because data is not freely disclosed by the applicable regulatory agencies to the public. Instead, the plaintiff must state a legal claim against a particular party and subsequently acquire the data to prove this claim by incurring the costs of discovery, which are often large. In this way, limited regulatory information disclosure, coupled with costly discovery, ensures that hypotheses (concerning violations of the law) are tested by properly applying the scientific method, with the formulation of a theory of liability preceding empirical analysis of the data ultimately used to test that theory.

The Article proceeds as follows: Part II explains the multiple hypothesis testing problem. The home mortgage lending data used in this study is described in Part III. Part IV sets forth the empirical strategy and details the statistical analysis used to isolate and identify evidence of redlining by lending institutions. Part V presents my main findings and provides a discussion of the results. Part VI considers the policy implications of these results for regulatory agencies. Part VII briefly concludes.

II. The Multiple Hypothesis Testing Problem

This Part briefly describes the concept of statistical significance testing and explains the multiple hypothesis testing problem that arises when multiple hypotheses are tested simultaneously on the same dataset.

---

[8] *See, e.g.*, Samuel Issacharoff & Geoffrey Miller, *An Information-Forcing Approach to the Motion to Dismiss*, 5 J. Legal Analysis 437, 447–48 (2013).

A. Statistical Significance Testing

Statistical significance testing is the foundation of experimental science and helps determine whether an observed result is the product of mere chance.[9] The idea is to formulate a theory and then collect data to test if this theory is true.[10] Statistical tests are structured in the form of hypothesis tests.[11] There are two types of hypotheses: (1) the null hypothesis and (2) the alternative hypothesis.[12] The null hypothesis represents the case where the theory does not hold and is the assumed outcome of the experiment if the tested theory is false.[13] In other words, if the null hypothesis is true, then no new information has been obtained other than the negative result that the theory cannot be rejected as false. The alternative hypothesis is the complement of the null hypothesis.[14] A statistical test is designed to determine if the null hypothesis can be rejected.[15]

To illustrate the concept, suppose one wishes to know whether a coin is fair. In this case, the theory to be tested is that the coin is biased. Thus, the null hypothesis corresponds to the case where both outcomes of the coin toss are equally likely (i.e., unbiased), with the coin landing heads or tails with equal probability. Defined as the complement of the null hypothesis, the alternative hypothesis corresponds to the case where one of the two outcomes (either heads or tails) is more likely.[16] To distinguish between these two competing hypotheses, statistical evidence is collected and a test statistic is defined. Here, a sensible test statistic can be defined as the number of heads observed, z, after k coin tosses. Calculating the test statistic by tossing the coin k times and summing up the total number of heads observed, z, the statistical evidence is represented by a p-value that expresses the probability that the test statistic would be equal to, or greater than, the actual observed results, *assuming that the null hypothesis is true*.[17] The smaller the p-value, the more likely it is that the null hypothesis under consideration may not satisfactorily explain the observation.[18] In particular, the null hypothesis is rejected if the p-value is less than, or equal to, a small, fixed, but arbitrarily pre-defined, threshold value α, which is referred to as the level of significance.[19] If the null hypothesis is rejected, then the observed result is said to

[9] *See* David H. Kaye, *Is Proof of Statistical Significance Relevant?*, 61 Wash. L. Rev. 1333, 1337 (1986).
[10] *See id.*
[11] *See id.*
[12] *See id.* at 1337–39.
[13] *See id.*
[14] *See id.*
[15] *See id.*
[16] Note that the alternative hypothesis does not state which side of the coin is favored, only that the coin is biased in one direction or the other.
[17] *See also* Kaye, *supra* note 9, at 1342. *See generally* Daniel L. Rubinfeld, *Econometrics in the Courtroom*, 85 Colum. L. Rev. 1048 (1985).
[18] Kaye, *supra* note 9, at 1342.
[19] *See id.* at 1345-49.

be statistically significant at an α-confidence level.[20] Unlike the *p*-value, the α-level is not derived from observational data, nor does the α-level depend upon the underlying hypothesis; rather, the researcher should fix the value of α *prior* to examining the data.[21]

B. Multiple Hypothesis Testing

The problem of multiple hypothesis testing arises when a collection of statistical hypotheses are all tested simultaneously. Traditional statistical methods designed for a single hypothesis cannot be employed in this case because the probability of false positives is likely to increase as the number of hypotheses tested increases.[22] Consider, for example, a case where the researcher has twenty hypotheses to test and has set the confidence level at 0.05 (i.e., $\alpha = 0.05$) prior to examination of any data collected. What is the probability of observing at least one significant result due purely to chance? It is straightforward to show that there exists an approximately sixty-four percent chance of observing at least one significant result, even if it is true that *all the tests are not statistically significant*.[23] Thus, the α-confidence level no longer expresses the maximum allowed probability of a false positive. In particular, α is now strictly less than the true probability of falsely rejecting a true null hypothesis.[24]

Testing multiple (unadjusted) statistical hypotheses in an effort to locate a significant result, whether done unintentionally or deliberately, is commonly referred to as "*p*-hacking."[25] Specifically, *p*-hacking is the use of data mining techniques to discover patterns in the data that are spuriously presented as statistically significant without first devising a specific hypothesis as to the underlying causality.[26] By running multiple statistical tests, and then focusing on only those tests that yield statistically significant results, it is possible to report a false positive (or type I error) as a true positive.[27] In a widely cited paper, for example, the authors demonstrate how it is possible

---

[20] *See id.*

[21] *See* Sander Greenland et al., *Statistical Tests, P-Values, Confidence Intervals, and Power: A Guide to Misinterpretations*, 31 Eur. J. Epidemiology 337, 339–40 (2016).

[22] *See* Hans Zeisel & David Kaye, Prove It With Figures: Empirical Methods in Law and Litigation 93 (1997).

[23] The probability calculation is as follows: P(at least one significant result) = 1 – P(no significant results) = $1 - (1 - 0.05)^{20} \approx 0.64$.

[24] *See, e.g.*, Craig M. Bennett et al., Neural Correlates of Interspecies Perspective Taking in the Post-Mortem Atlantic Salmon: An Argument for Multiple Comparisons Correction 1 (2009), http://prefrontal.org/files/posters/Bennett-Salmon-2009.pdf [https://perma.cc/97BX-JP6B].

[25] *See, e.g.*, Regina Nuzzo, *Scientific Method: Statistical Errors*, 506 Nature 150, 151–52 (2014); *see also* Andrew Gelman & Hal Stern, *The Difference Between "Significant" and "Not Significant" Is Not Itself Statistically Significant*, 60 Am. Sci. 328 (2006).

[26] Regina Nuzzo, *supra* note 25, at 152.

[27] *See, e.g.*, John P. Ioannidis, *Why Most Published Research Findings Are False*, 2 PLOS Med. 696, 696–97 (2005); *see also* Eric Loken & Andrew Gelman, *Measurement Error and the Replication Crisis*, 355 Sci. 584 (2017).

to "consciously or unconsciously" exploit "researcher degrees of freedom" to obtain a *p*-value of 0.05.[28] Using entirely negative data, the authors show that a *p*-value of 0.05 can be achieved sixty percent of the time purely by exploiting researcher degrees of freedom (i.e., by *p*-hacking).[29]

However, the problem of *p*-hacking is solvable. Several statistical methods have been developed to compensate for multiple comparisons, such as the Bonferroni correction or the Benjamini-Hochberg set-up procedure.[30] Generally speaking, these techniques require a stricter significance level for individual comparisons, compensating for the multiple statistical hypotheses tested.[31] A somewhat different approach to the problem of *p*-hacking is to conduct randomized out-of-sample tests. Under this approach, a research program generally proceeds in two distinct steps: (1) a preliminary set of data is collected and (2) if the preliminary data appears promising, then a replication analysis is performed in which all decisions relating to a fresh data collection (e.g., when to stop recording the data, what variables to record, which comparisons to make, and which statistical methods to employ) have been both pre-determined and registered publicly.[32] The key feature of such independent replication is that the research hypothesis is not tested on the same data that was used to construct the research hypothesis.[33] Maintaining the integrity of this aspect of the scientific method is critical because every dataset contains some patterns that are attributable entirely to chance.[34] If the hypothesis is not tested on a different dataset from the same statistical population, then it is not possible to determine the likelihood that chance alone produced the observed patterns.[35] Generating hypotheses based upon observed data, in the absence of subsequent testing on new data, is some-

---

[28] Joseph P. Simmons, Leif D. Nelson & Uri Simonsohn, *False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allow Presenting Anything as Significant*, 22 Psychol. Sci. 1359, 1361 (2011).

[29] *See id.*; *see also* Daniele Fanelli, *How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data*, PLOS One, May 29, 2009, at 1 (using survey data to show that a third of researchers admit to engaging in questionable methods that would exploit degrees of freedom in order to generate positive statistical results).

[30] *See, e.g.*, Jill E. Fisch, Jonah B. Gelbach & Jonathan Klick, *The Logic and Limits of Event Studies in Securities Fraud Litigation*, 96 Tex. L. Rev. 553, 600–01 (2018).

[31] *See id.*

[32] *See, e.g.*, Andrew Gelman & Eric Loken, *The Statistical Crisis in Science*, 102 Am. Sci. 460, 463–65 (2014); Macartan Humphrey, Raul Sanchez de la Sierra & Peter van der Windt, *Fishing, Commitment, and Communication: A Proposal for Comprehensive Nonbinding Research Registration*, 21 Pol. Analysis 1, 1, 3–5 (2013); Uri Simonsohn, Lief D. Nelson & Joseph P. Simmons, *P-Curve: A Key to the File-Drawer*, 143 J. Experimental Psychol. 534, 546 (2013); *see also* Simmons et al., *supra* note 28, at 1362–63 (advocating for greater transparency in data collection processes as a means to prevent *p*-hacking and contending that researchers should establish robustness by disclosing what the statistical results would have been had other data analysis decisions been made).

[33] *See, e.g.*, Richard Berk, Lawrence Brown & Linda Zhao, *Statistical Inference after Model Selection*, 26 J. Quantitative Criminology 217, 218–20 (2009).

[34] *See, e.g.*, Ronald L. Wasserstein & Nicole A. Lazar, *The ASA's Statement on p-Values: Context, Process, and Purpose*, 70 Am. Stat. 129, 131–32 (2016).

[35] *See id.*

times referred to as "post hoc theorizing," and represents a form of circular reasoning. A statistical result appears true in a limited dataset. The result is, therefore, hypothesized to hold true in general. Then, the hypothesis is (improperly) tested on the same limited dataset, confirming, entirely predictably, that the "general" hypothesis is true.[36]

In what follows below, I intentionally use a form of *p*-hacking to improperly isolate and identify a set of lending institutions that, on the basis of the statistical evidence, appear to have engaged in redlining. Rather than starting the analysis with a specific hypothesis as to a particular lender that has allegedly engaged in redlining home mortgage loan applications, simple data mining procedures are used to generate a set of hypotheses as to lenders that might be redlining. Specifically, for each lender-MSA pair in a nationwide dataset, a potential target list of lenders is generated by exhaustively searching for those observations that show a statistically significant difference in the rate of lending in minority areas at a five percent confidence level. If the hypothesis of redlining were tested on this same existing nationwide data for any lender in the generated potential target list, then, of course, the discrimination hypothesis would be confirmed—although this confirmation is innately circular and somewhat meaningless. Nonetheless, a plaintiff can use these spurious data mining techniques, not necessarily maliciously or even intentionally, to allege a prima facie case of disparate impact sufficient to survive a defendant-lender's motion to dismiss.[37]

III. Data

The dataset in this study is a panel dataset of annual, census tract-level observations starting in 2012 and ending in 2016. Each tract-level observation contains HMDA home mortgage loan data and demographic census tract-level characteristics.

A. HMDA Mortgage Data

The home mortgage loan data come from reports submitted by lenders under the Home Mortgage Disclosure Act of 1975.[38] Originators of home

[36] *See* James Jaccard & Jacob Jacoby, Theory Construction and Model-Building Skills: A Practical Guide for Social Scientists 355–56 (2010); *see also* Andrew Gelman & John Carlin, *Some Natural Solutions to the p-Value Communication Problem—and Why They Won't Work*, 112 J. Am. Stat. Ass'n 899, 899 (2017).

[37] *See generally* Mitchel B. Rachlis & Anthony M. Yezer, *Serious Flaws in Statistical Tests for Discrimination in Mortgage Lending*, 4 J. Housing Res. 315 (1993); Anthony M. J. Yezer, Robert F. Phillips & Robert P. Trost, *Bias in Estimates of Discrimination and Default in Mortgage Lending: The Effects of Simultaneity and Self-Selection*, 9 J. Real Est. Fin. & Econ. 197 (1994).

[38] 12 U.S.C. §§ 2801–2809 (2012); *see, e.g.*, Robert P. Voyles, Comment, *Red-lining and the Home Mortgage Disclosure Act of 1975: A Decisive Step Toward Private Urban Redevelopment*, 25 Emory L.J. 667 (1976).

mortgages covered under HMDA are required to annually report the basic attributes of home mortgage loan applications received in MSAs to the Federal Financial Institutions Examination Council ("FFIEC").[39] Since 1990, HMDA has contained information on individual borrower characteristics, such as applicant race, ethnicity, and gender.[40] In 2004, an additional variable was added to HMDA identifying first-lien mortgages with an annual percentage rate ("APR") three percentage points above the applicable Treasury benchmark (and for junior liens with an APR five percentage points above the applicable Treasury benchmark).[41] Mortgages with a reported spread are commonly referred to as "higher-priced" loans. In this study, "higher-priced" is used as a proxy for subprime.[42] Further, the HMDA data contains the census tract of the property subject to each individually reported mortgage.[43] This variable can be used to calculate the total number of loan applications received from a given census tract. The HMDA data also includes additional information on individual loan characteristics, such as the disposition of mortgage loan applications (e.g., loan origination, application denied, and so forth), the loan type (i.e., conventional or government-backed the latter including FHA-insured and VA-guaranteed), and whether the loan was sold in the secondary market and, if so, the type of institution to which the loan was subsequently sold (e.g., Fannie Mae, Freddie Mac, a commercial bank, and so forth).[44] Finally, the HMDA data identifies the supervisory or regulatory agency of each reporting lender.[45]

To avoid the possibility of double counting, the home mortgage loans analyzed in this study are non-purchased home mortgage loans, meaning mortgage loans that are not purchased in the secondary market.[46] Moreover, the set of lenders considered does not include subprime lenders, which are

---

[39] 12 U.S.C. §§ 2801–2809. A financial institution must report HMDA data to its supervisory/regulatory agency if certain criteria are met, such as having assets above a specific threshold. The criteria differ with respect to depository and non-depository institutions and are available on the FFIEC website. *See Home Mortgage Disclosure Act*, Fed. Fin. Insts. Examination Council (Sept. 6, 2018), https://www.ffiec.gov/hmda/reporter.htm [http://perma.cc/D5NJ-4U76].

[40] *See, e.g.*, Robert Avery, Kenneth P. Brevoort & Glenn B. Canner, *Opportunities and Issues in Using HMDA Data*, 29 J. Real Est. Res. 351, 357–62, 373 (2007); *see also* Allen Fishbein, *Fair Lending Conference: Home Mortgage Disclosure Act Report*, 28 J. Marshall L. Rev. 343 (1995).

[41] *See* Robert Avery, Kenneth P. Brevoort & Glenn B. Canner, *Opportunities and Issues in Using HMDA Data*, 29 J. Real Est. Res. 351, 353, 368 (2007).

[42] *See*, *e.g.*, Kevin Park, Subprime Lending and the Community Reinvestment Act (Nov. 2008) (unpublished manuscript) (on file with the Joint Center for Housing Studies of Harvard University).

[43] Census tracts are designed by the U.S. Census Bureau to be relatively homogeneous units with respect to population characteristics, economic status, and living conditions, and typically average about 4000 to 5000 residents. U.S. Dep't of Labor & U.S. Dep't of Com., Census Use Study Handbook for Manpower Planners 67 (1975).

[44] *See* Avery et al., *supra* note 41, at 378 n.27.

[45] *See id.*

[46] *See generally* Robin Paul Malloy, *The Secondary Mortgage Market—A Catalyst for Change in Real Estate Transactions*, 39 Sw. L.J. 991 (1985).

defined as any lender having a percentage of interest rate-reportable loans originated equal to, or greater than, twenty-five percent. Finally, to be included in the redlining analysis described in Part IV.B, a lender must have received at least twenty loan applications from the given MSA in each year of a given lookback period.[47]

B. Census Data

The MSAs analyzed in this study are based on the revised delineations announced by the Office of Management and Budget ("OMB") in February 2013.[48] For each MSA, census tract-level information relating to race and ethnicity were retrieved from the 2010 Decennial Census.[49] This census tract-level information is used to determine if a census tract in a MSA is "majority-minority," meaning that less than fifty percent of the total census tract population can be categorized as non-Hispanic white (i.e., a person who is racially white and is not of Hispanic or Latino origin or ethnicity). A MSA must contain at least one majority-minority census tract to be included in the analysis.[50]

Table 1 provides summary statistics of individual lender-level and MSA-level variables. There is substantial variance in loan application volume across lenders, with the largest lenders receiving over 60,000 loan applications in a given MSA and year. Across these lenders, there is also substantial variance in lending rates in majority-minority census tracts. The mean lending rate in majority-minority census tracts is approximately fifteen percent. At the aggregate level, the corresponding mean rate is approximately sixteen percent. Notably, the percentage of majority-minority tracts is approximately twenty-eight percent, indicating that a disproportionate percentage of home mortgage loan applications are received from *non*-majority-minority census tracts. Finally, there exists substantial variance in the percentage of majority-minority tracts across MSAs.[51]

---

[47] This condition holds true unless expressly stated otherwise (as in Table 5).

[48] *See* OFFICE OF MGMT. & BUDGET, EXEC. OFFICE OF THE PRESIDENT, OMB BULL. NO. 13-01, REVISED DELINEATIONS OF METROPOLITAN STATISTICAL AREAS, MICROPOLITAN STATISTICAL AREAS, AND COMBINED STATISTICAL AREAS, AND GUIDANCE ON USES OF THE DELINEATION OF THESE AREAS (2013), https://obamawhitehouse.archives.gov/sites/default/files/omb/bulletins/2013/b-13-01.pdf [http://perma.cc/PS6C-Z2Q9].

[49] *See* U.S. CENSUS BUREAU, DECENNIAL CENSUS OF POPULATION AND HOUSING (2010), https://www.census.gov/programs-surveys/decennial-census/decade/decennial-publications.2010.html [http://perma.cc/G2VV-ZC5H].

[50] This condition reduces the number of MSAs from 949 to 487.

[51] The MSAs with the highest percentage of majority-minority tracts tend to include large populations of Latino/Hispanic or Asian/Native Hawaiian & Pacific Islander individuals. For example, the top 5 MSAs are (in descending order) as follows: (1) Brownsville-Harlingen, TX; (2) El Paso, TX; (3) El Centro, CA; (4) Gallup, NM; and (5) Hilo, Hawaii.

IV. Empirical Strategy

This Part describes the empirical approach used in this study to determine if a lending institution has engaged in redlining. The discussion begins with a legal analysis of redlining and then describes the type of statistical evidence that is generally relied upon within this legal framework to establish a prima facie case of redlining.

A. Legal Analysis of Redlining

As noted, both the FHA and ECOA prohibit discrimination in residential lending on the basis of race or national origin. Section 805 of the FHA expressly prohibits "any person or other entity whose business includes engaging in residential real-estate-related transactions [from] discriminat[ing] against any person in making available such a transaction, or in the terms or conditions of such transaction, because of race."[52] Section 804(a) of the FHA prohibits actions that "otherwise make unavailable or deny" housing "because of race . . . or national origin,"[53] and section 804(b) prohibits discrimination "in the terms, conditions, or privileges of sale or rental of a dwelling, or in the provision of services or facilities in connection therewith, because of race . . . or national origin."[54] In addition to the FHA's prohibitions, the ECOA makes it "unlawful for any creditor to discriminate against any applicant, with respect to any aspect of a credit transaction on the basis of race [or] national origin . . . ."[55] The ECOA regulations are clear that this prohibits a creditor from "discourag[ing] on a prohibited basis a reasonable person from making or pursuing an application."[56] These statutory protections are "broad and inclusive,"[57] and generally encompass any policy or practice that discriminates in the terms or conditions of residential mortgage loans, or that otherwise makes housing unavailable on the basis of race or national origin, which includes redlining.[58]

---

[52] 42 U.S.C. § 3605(a) (2012). The FHA defines a "residential real estate-related transaction" as, among other things, "[t]he making or purchasing of loans or providing other financial assistance—(A) for purchasing . . . a dwelling; or (B) secured by residential real estate." *Id.* § 3605(b).

[53] *Id.* § 3604(a).

[54] *Id.* § 3604(b).

[55] 15 U.S.C. § 1691(a) (2012).

[56] 12 C.F.R. § 1002.4(b) (2019).

[57] City of Edmonds v. Oxford House, 514 U.S. 725, 731 (1995) (quoting Trafficante v. Metro Life Ins. Co., 409 U.S. 205, 209 (1972)).

[58] *See, e.g.*, Harrison v. Otto G. Heinzeroth Mortg. Co., 430 F. Supp. 893, 896 (N.D. Ohio 1977); *see also* Ring v. First Interstate Mortg., Inc., 984 F.2d 924, 927 (8th Cir. 1993) (denying a motion to dismiss where the complaint alleged a refusal to make available a residential real-estate-related transaction "based on the racial composition of the tenants of the properties or the neighborhoods in question"); United Companies Lending Corp. v. Sargeant, 20 F. Supp. 2d 192, 203 n.5 (D. Mass. 1998) (redlining addresses differential outcomes targeted to "specific geographies due to the income, race, or ethnicity of its resident") (citing S. Rep. No. 103-

To prove a pattern or practice of discrimination, the plaintiff must show that it was a regular, and not an unusual, practice of the defendant to act in a discriminatory manner.[59] A pattern or practice claim of discrimination under either the FHA or ECOA may be brought under a theory of disparate treatment.[60] Disparate treatment may be proven either with direct evidence of intent or with indirect and circumstantial evidence, using a burden-shifting method that relies upon facts which together create an inference of discriminatory intent.[61] Moreover, if similarly situated people are treated differently, then a reasonable factfinder may presume that the defendant "based his decision on an impermissible consideration such as race."[62]

A plaintiff may also bring a pattern or practice claim under the FHA relying solely upon a disparate impact theory.[63] To prove a disparate impact claim, a plaintiff must first establish a prima facie case by showing: (1) a specific policy (2) had a discriminatory impact on a protected class[64] and (3) the facts show a "robust" causal connection between the policy and dispa-

169, at 21 (1993)). Department of Housing and Urban Development ("HUD") regulations also prohibit discrimination in residential lending. *See* 24 C.F.R. §100.120 (2019).

[59] Int'l Bhd. of Teamsters v. United States, 431 U.S. 324, 336 n.16 (1977) (Title VII); *see also* Hohider v. United Parcel Serv., Inc., 574 F.3d 169, 178 (3d Cir. 2009) (employment discrimination context); United States v. Balistrieri, 981 F.2d 916, 929 (7th Cir. 1992); United States v. Pelzer Realty Co., 484 F.2d 438, 445 (5th Cir. 1973).

[60] *See* Betsey v. Turtle Creek Assocs., 736 F.2d 983, 988 (4th Cir. 1984); *see also* Ledbetter v. Goodyear Tire & Rubber Co., 550 U.S. 618, 631 (2007), *superseded by statute on other grounds*; *Teamsters*, 431 U.S. at 335 n.15 (Disparate treatment "is the most easily understood type of discrimination. The [defendant] simply treats some people less favorably than others because of their race, color, religion, sex, or [other protected trait].").

[61] *See Ledbetter*, 550 U.S. at 631; McDonnell Douglas Corp. v. Green, 411 U.S. 792, 801 (1973); *see also* Desert Palace, Inc. v. Costa, 539 U.S. 90, 100 (2003) ("A plaintiff does not need a 'smoking gun' to prove invidious intent, and few plaintiffs will have one. Rather, '[c]ircumstantial evidence is not only sufficient, but may also be more certain, satisfying and persuasive than direct evidence.' "); *Teamsters*, 431 U.S. at 335 n.15 ("Proof of discriminatory motive is critical, although it can in some situations be inferred from the mere fact of differences in treatment."); Vill. of Arlington Heights v. Metro. Hous. Corp., 429 U.S. 252, 266 (1977) ("Determining whether invidious discriminatory purpose was a motivating factor demands a sensitive inquiry into such circumstantial and direct evidence of intent as may be available.").

[62] Clark v. Int'l Bhd. of Elec. Workers, Local No. 98, No. 12-5897, 2013 WL 6284171, at *4 (E.D. Pa. Dec. 4, 2013) (employment discrimination).

[63] *See* Tex. Dep't of Hous. and Cmty. Affairs v. The Inclusive Cmtys. Project, Inc., 135 S. Ct. 2507, 2512 (2015) (affirming the continued viability of disparate impact claims under the FHA); *see also* Mt. Holly Gardens Citizens in Action, Inc. v. Twp. of Mount Holly, 658 F.3d 375, 383 (3d Cir. 2011) (noting that the appropriate inquiry is "whether minorities are disproportionately affected" by the challenged practice). The Supreme Court has not yet ruled on the issue of whether disparate impact claims are legally cognizable under the ECOA. To the extent that the Supreme Court tends to accord some degree of deference to the views of regulatory agencies, it is noteworthy that the Consumer Financial Protection Bureau ("CFPB") has concluded that disparate impact liability is consistent with the ECOA. *See* Consumer Fin. Prot. Bureau, CFPB Bull. No. 2012-14 (Fair Lending (2013)), https://files.consumerfinance.gov/f/201404_cfpb_bulletin_lending_discrimination.pdf [https://perma.cc/9CZH-WAVT] ("Consistent with other federal supervisory and law enforcement agencies, the CFPB reaffirms that the legal doctrine of disparate impact remains applicable . . . to enforce compliance with the ECOA and Regulation B.").

[64] *See* Smith v. City of Jackson, 544 U.S. 228, 241 (2005).

rate impact.[65] If a plaintiff has established a prima facie case, then the burden of proof shifts to the defendant to offer a "business justification" for its policy necessary to achieve a valid interest.[66] Policies are "not contrary to the disparate-impact requirement unless these policies are 'artificial, arbitrary, and unnecessary barriers.'"[67] If a defendant adequately satisfies this burden, then the plaintiff may still establish liability by showing "that there is an available alternative . . . practice that has less disparate impact and serves the [bank's] legitimate needs."[68]

Three relatively recent cases—all settled shortly after the filing of the complaint—illustrate the substantial financial stakes involved as the target of a redlining lawsuit and highlight the type of statistical allegations that are likely to compel a lender to settle rather than enter into costly litigation. In all three cases, the plaintiff sought to establish a prima facie case of redlining, in large part, by showing a statistically significant difference between the bank's lending rate in minority areas and lending rates in minority areas by its peer lenders. In *United States Department of Housing and Urban Development v. Associated Bank, N.A.*,[69] for example, the plaintiff asserted that "*compared to other lenders*, [Associated Bank's] lending in majority-minority census tracts was lower than in other neighborhoods, and the difference was statistically significant."[70] Similarly, in *Consumer Financial Protection Bureau & United States v. Hudson City Savings Bank, F.S.B.*,[71] the plaintiff alleged that "[a]nalysis of Hudson City's mortgage applications . . . *as compared to its peers* showed disparities in lending to majority Black-and-Hispanic neighborhoods between Hudson City and its peers. These disparities [were] statistically significant."[72] Finally, in *United States v. Eagle Bank & Trust Company of Missouri*,[73] the complaint asserted that "statistical analyses of [Eagle Bank's] residential real estate-related loan applications and originations for each year from 2006 to 2012 showed that . . . [d]uring that time, there were statistically significant disparities with respect to the defendant's residential real estate lending activity *when compared with similar lenders*."[74] Specifically, in *Hudson City*, 0.1% of the loan applications came from high Black-and-Hispanic areas in the Camden

[65] *Inclusive Cmtys.*, 135 S. Ct. at 2512.

[66] *See, e.g.*, Resident Advisory Bd. v. Rizzo, 564 F.2d 126, 148 (3d Cir. 1977).

[67] *Inclusive Cmtys.*, 135 S. Ct. at 2524 (quoting Griggs v. Duke Power Co., 401 U.S. 424, 431 (1971)).

[68] *Id.* at 2518 (citing Ricci v. DeStefano, 557 U.S. 557, 578 (2009)).

[69] No. 00-12-0003-8 (May 26, 2015).

[70] *Id.*

[71] No. 15 Civ. 7056 (D.N.J. Sept. 24, 2015).

[72] Complaint at 13, Consumer Fin. Prot. Bureau & United States v. Hudson City Sav. Bank, F.S.B., No. 15 Civ. 7056 (D.N.J. Sept. 24, 2015) (emphasis added), https://www.justice.gov/crt/case-document/consumer-financial-protection-bureau-and-united-states-v-hudson-city-bank-fsb [https://perma.cc/FQ7A-9X5K].

[73] No. 15 Civ. 1492 (E.D. Mo. Sept. 29, 2015).

[74] Complaint at 5, United States v. Eagle Bank & Trust Co. of Mo., No. 15 Civ. 1492 (E.D. Mo. Sept. 29, 2015) (emphasis added).

MSA, compared to 4.4% for Hudson City's peer institutions.[75] Likewise, only 1.9% of Eagle Bank's total loan applications were located in majority-Black census tracts in the St. Louis MSA, whereas Eagle Bank's peer lenders generated 11.1% of their total loan application volume during the same time period from majority-Black census tracts.[76] Although, in each case, there was also qualitative evidence relating to both limited branch locations and marketing efforts in majority-minority areas, as well as improperly delineated CRA assessment areas, it was arguably the statistical allegations that compelled these lending institutions to settle the cases early and pay large monetary damage awards. In particular, Hudson City was required to pay a $5.5 million civil monetary penalty and to provide an additional $25 million in direct loan subsidies to qualified borrowers in majority Black-and-Hispanic neighborhoods (up to $18,750 per individual applicant).[77] Similarly, in its consent order with the U.S. Department of Housing and Urban Development ("HUD"), Associated Bank agreed to provide $9.5 million in subsidies to qualified loan applicants in certain majority-minority census tracts and to originate, fund, or purchase an additional $190.8 million in residential mortgage loans in several previously underserved markets.[78]

B. Statistical Analysis of Redlining

As noted, courts have recognized that plaintiffs can rely upon the appropriate statistics, properly analyzed, to establish proof of disparate treatment and disparate impact on the basis of protected class status. Thus, a statistically significant difference between a lender's rate of loan applications received from majority-minority census tracts and the corresponding rate of its peer institutions can be used to establish a prima facie case of discriminatory intent.[79] Regarding the important question of what constitutes statistical significance, the Supreme Court in *Hazelwood School District v. United*

---

[75] *See* Complaint, Hudson City, *supra* note 72, at 14.

[76] *See* Complaint, Eagle Bank, *supra* note 74, at 6.

[77] *See* Consent Order at 13, *Hudson City*, No. 15 Civ. 7056. In addition, Hudson City was required to pay $750,000 in community development partnership programs, $200,000 per year on targeted advertising and outreach, and $100,000 per year on consumer financial education programs. *See id.* at 14–17.

[78] *See* Conciliation Agreement at 8–13, U.S. Dep't of Hous. and Urban Dev. v. Associated Bank, N.A., No. 00-12-0003-8 (May 26, 2015). Associated Bank was also required to pay $3 million in affordable home repair grants, $1.4 million in affirmative marketing and outreach, and $1.35 million for CRA training and education. *See id.*

[79] McCleskey v. Kemp, 481 U.S. 279, 293 (1987); Hazelwood Sch. Dist. v. United States, 433 U.S. 299, 308–09 n.14 (1977) (finding gross statistical disparities in excess of two or three standard deviations can make out a prima facie case in a Title VII action); Vill. of Arlington Heights v. Metro. Hous. Corp., 429 U.S. 252, 266 (1977) (finding the discriminatory effect of conduct over a prolonged period that is "unexplainable on grounds other than race" may warrant an inference of purposeful discrimination); *see also* Mt. Holly Gardens Citizens in Action, Inc. v. Twp. of Mount Holly, 658 F.3d 375, 382 (3d Cir. 2011) (holding that disparate impact shown by means of statistical evidence was sufficient to establish a prima facie case of discrimination).

States set a two standard deviation threshold to establish a prima facie case of discrimination.[80] This implies that an observed difference in lending rates is statistically significant only if the null hypothesis (of no difference) can be rejected at a five percent confidence level or less.[81] As for statistical methodology, the Supreme Court, in a series of Title VII cases, endorsed a standard methodology for comparing the expected number of events (e.g., the expected number of home mortgage loan applications received by the defendant-lender from majority-minority census tracts assuming no redlining) with the actual number of events (e.g., the actual number of home mortgage loan applications received by the defendant-lender from majority-minority census tracts).[82] Specifically, the Supreme Court held that two parameters must be defined to ensure a legally probative statistical analysis: (1) the relevant peer group for statistical comparison, and (2) the relevant market area.[83]

With respect to peer group comparison, the Supreme Court in *Teamsters* stated that "non-discriminatory hiring practices will in time result in a work force more or less representative of the racial and ethnic composition of the population in the community from which employees are hired."[84] In *Hazelwood*, the Court refined this comparison, holding that, in an employment discrimination context, the employer's work force should be compared to the group of people who possess the necessary qualifications for the job at issue—in that case, the qualified public school teacher population.[85] Applying the holdings of this line of cases to redlining, courts compare the home mortgage loan applications received by a defendant-lender from allegedly redlined geographic areas to home mortgage loan applications received by the group of lenders against which the lender generally competes in its home mortgage lending business: its "peer group."[86] According to guidelines set forth by the Federal Reserve Board, a peer for a defendant-lender can be selected on the basis of loan application volume, and, in particular, can be defined as any lender in a given MSA and time period having between 50%

---

[80] 433 U.S. 299, 308 n.14 (1977).

[81] *See id.* at 309; *see also* Adams v. Ameritech Servs., Inc., 231 F.3d 414, 424 (7th Cir. 2000) (stating that "two standard deviations is normally enough to show that it is extremely unlikely" that a "disparity is due to chance"); Anderson v. Zubieta, 180 F.3d 329, 333–41 (D.C. Cir. 1999) (indicating that "disparities . . . exceed[ing] 1.96 standard deviations under a two-tailed test of statistical significance" are sufficient to establish a prima facie case of disparate impact).

[82] *See id.*; Int'l Bhd. of Teamsters v. United States, 431 U.S. 324, 340 (1977); *see also* Craik v. Minnesota State Univ. Bd., 731 F.2d 465 (8th Cir. 1984) (holding, in a Title VII case, that drawing applications at an uneven rate could weigh in favor of a finding of discrimination). *See generally* Castaneda v. Partida, 430 U.S. 482 (1977).

[83] *See Hazelwood*, 433 U.S. at 308–12.

[84] 431 U.S. at 340 n.20.

[85] *See Hazelwood*, 433 U.S. at 308 n.13.

[86] *See*, *e.g.*, Ring v. First Interstate Mortg., Inc., 984 F.2d 924, 924 (8th Cir. 1993); Gallagher v. Magner, 619 F.3d 823, 831 (8th Cir. 2010).

and 200% of the defendant-lender's total home mortgage loan application volume in that MSA during the same time period.[87]

As for the relevant market area, if a lender is a depository institution subject to the requirements of the CRA and its enabling regulations promulgated by the Consumer Financial Protection Bureau ("CFPB"), then the lender's delineated CRA assessment area might serve reasonably well as the relevant market area for the purposes of a redlining analysis.[88] The CRA requires most banks to delineate one or more assessment areas for federal regulators to consider in evaluating if the institution is meeting the existing credit needs of the entire community. Under these regulations, a bank's CRA assessment area must, in general, consist of one or more MSAs (or one or more contiguous political subdivisions such as a county or city).[89] A bank's assessment area must consist only of whole geographies and may not reflect illegal discrimination, taking into account the bank's size and financial condition.[90]

A bank's delineated CRA assessment area, however, is likely too limited in most cases to serve as the relevant market area in a redlining analysis. A proper redlining analysis should focus on a lender's decisions concerning how much access to credit it should provide to different geographical areas. The areas for which those decisions can best be compared are areas in which the institution actually marketed and provided credit in the past and, also, where the institution *could have reasonably been expected to have marketed and provided credit*.[91] These reasonably expected market areas ("REMAs") are likely to extend beyond, or otherwise differ, from the lender's CRA assessment area.[92] Moreover, a bank's delineated CRA assessment area may itself reflect discriminatory intent: the bank may have intentionally defined its assessment area to exclude certain majority-minority neighborhoods in a

---

[87] *See* Maureen Yap, *Fair Lending Webinar: Questions and Answers,* Consumer Compliance Outlook (Oct. 17, 2012), https://consumercomplianceoutlook.org/2013/second-quarter/fair-lending-webinar/ [https://perma.cc/BW58-XB7U]. If the screening procedure here generates fewer than ten peers for a lender in a given MSA and time period, then the closest lenders in terms of application volume are added to the list of peers until the total number of peer lenders equals ten. So, for example, suppose that a targeted lender receives 1000 home mortgage loan applications in a given MSA and year. According to the Federal Reserve Board guidelines, this targeted lender's peer group is defined as the set of all lenders receiving between 500 and 2000 loan applications in that same MSA and time period. If this set contains less than ten lenders (say, for example, only nine lenders), then the lender's peer group will also include that lender whose total number of loan applications, in that same MSA and time period, is closest to the closed interval, [500, 2000].

[88] 12 U.S.C. § 2901(a)(3) (2012) (stating that a bank has a "continuing and affirmative obligation to help meet the credit needs of the local communities in which [the bank is] chartered").

[89] *See* 12 C.F.R. § 228.41(c)(1) (2019).

[90] *See id.* § 228.41(e).

[91] *See* Fed. Fin. Insts. Examination Council, Interagency Fair Lending Examination Procedures 32 (2009), https://www.ffiec.gov/pdf/fairlend.pdf [https://perma.cc/5HYU-FUDZ].

[92] *See id.*

given MSA.[93] In other words, a bank may have drawn a strangely shaped or gerrymandered CRA assessment area that excludes majority-minority communities, constituting indirect evidence of the bank's intent to avoid lending in these areas because of the corresponding racial or ethnic composition. For these reasons, in the redlining analysis below, the relevant geographic market area considered is not the lender's CRA assessment area, but is, instead, simply defined as the entire MSA in which the lender operates.

Having defined both the relevant group of lenders for statistical comparison, as well as the relevant market area, the redlining screen conducted in this study compares, for each lender in the HMDA data, the proportion of loan applications received by a targeted lender in a given time period from majority-minority census tracts in a MSA to the proportion of such applications received by its peer group in that same time period and MSA.[94] Because redlining is based upon geography, not the race or ethnicity of an individual borrower, a loan application is categorized as "minority" if the property is located in a census tract with a majority-minority population (even if the individual applicant herself does not belong to a protected minority group).[95] For each lender, a pooled chi-squared test statistic is calculated.[96] If the calculated test statistic is statistically significant under a two-tailed test at a five percent confidence level (as set forth in *Hazelwood*), then the redlining screen will flag the lender as having engaged in redlining in the given MSA during the given time period.[97] This Article contends that this statistical significance, unlike other qualitative evidence, is *necessary* to establish a prima facie case of redlining against the lender. In fact, it may be possible to allege such a claim purely on the basis of this statistical evidence alone. Of course, as noted in the Introduction, other qualitative evidence is

---

[93] *See*, *e.g.*, Complaint, Hudson City, *supra* note 72 (exclusion of majority-Black-and-Hispanic neighborhoods from CRA assessment areas); Complaint, Eagle Bank, *supra* note 74 (exclusion of majority-Black census tracts from CRA assessment areas).

[94] The focus is on loan applications, and not originations, because whether or not a lender originates a loan implicates underwriting issues that possibly confound the redlining analysis. Notwithstanding, the results are robust to this difference.

[95] *See* George J. Benston, *Mortgage Redlining Research: A Review and Critical Analysis Discussion*, 12 J. BANK RES. 8, 10 (1981). *See generally* James A. Berkovec et al., *Race, Redlining, and Residential Mortgage Loan Performance*, 9 J. REAL EST. FIN. & ECON. 263 (1994).

[96] "Pooled" here refers to the manner in which the standard error is calculated. In the pooled version of this test, the proportion of loan applications received from majority-minority census tracts by the targeted lender and the proportion of such loan applications received by its peer group are averaged and this one proportion is used to estimate the standard error. In the un-pooled version of the test, the two proportions are used separately to estimate the standard error.

[97] This simple model accords with authors who argue in favor of the greater use of generic models on the theory that statistically controlling for non-racial variables may actually bias the results and mask the existence of unlawful disparate impact. *See*, *e.g.*, STEPHEN L. ROSS & JOHN YINGER, THE COLOR OF CREDIT: MORTGAGE DISCRIMINATION, RESEARCH METHODOLOGY, AND FAIR-LENDING ENFORCEMENT (2002); Ian Ayres, *Three Tests for Measuring Unjustified Disparate Impacts in Organ Transplantation: The Problem of "Included Variable" Bias*, 48 PERSP. BIOLOGY & MED. S68 (2005).

helpful and is likely necessary for the plaintiff to succeed at trial, or even to survive a motion to dismiss. But it is the statistical findings that form the principal basis of the allegation.

In addition to HMDA data, the redlining screen described above is also applied to randomly generated data. The goal is to see if, and how many, lenders can be classified as having engaged in redlining where there are no true violators in the data, *by construction*, and to compare the magnitude of the resulting percentages to those derived using actual HMDA data. To generate this random negative data, the percent of home mortgage loan applications received by the lender from majority-minority census tracks in a given MSA during a given time period is assumed to follow a binomial distribution. Each lender, on average, is constructed to have the same rate of home mortgage lending in majority-minority tracts as all other lenders in the MSA during a given time period, and any deviations from this MSA-wide mean rate are, therefore, purely random. To the extent that the screen identifies a subset of lenders as having engaged in statistical redlining when run on this randomly-generated negative data, this subset of alleged violators is purely the product of random variability, representing statistical outliers, and would not indicate true discrimination.

V. Results and Discussion

This Part presents the results of the redlining screen described in Part IV. As the foregoing analysis shows, the number of lenders flagged by the screen as having engaged in redlining is relatively large. Further, the analysis shows that approximately three percent of all lender-MSA pairs flagged (or approximately seven to nine percent of all lending institutions flagged) can be attributed to a failure to correct for the multiple hypothesis testing problem described in Part II.B.

A. Annual Performance Metrics

Table 2 presents the results of the screen on 2012–2016 HMDA data. The statistical results suggest, in general, a surprisingly high level of unlawful conduct. Using a one-sided test (at a ninety-five percent significance level as set forth in *Hazelwood*) and requiring statistical significance in at least one year, Column (1) suggests that approximately forty-seven percent of all MSA-lender pairs have engaged in redlining between 2012 and 2016 and approximately sixty-two percent of all lenders have engaged in such unlawful activity in at least one MSA. Column (2) provides similar results using a two-sided test. Using this two-sided test as the baseline test, Columns (3) to (6) of Table 2 show that the number of lenders flagged decreases as the required number of years of statistical significance increases. For instance, if a lender's rate of lending must be significantly different from its peer group in all five years, then the results in Table 2 suggest that approximately nine percent of all MSA-lender pairs have redlined between 2012 and 2016 and approximately twenty-three percent of all lenders have engaged in such unlawful conduct in at least one MSA. Although the percentages are now lower, Column (6) still implies a high degree of unlawful behavior. Is the law really being broken by financial institutions this frequently? The claim here is that the results in Table 2 are, in part, explainable by the *p*-hacking inherent when using a screen that is designed to "loop over" a large set of potential targets. In other words, by running multiple tests and then focusing upon only those tests that yield statistically significant results, some number of false positives may be presented as real statistical evidence of redlining.

To help assess this claim, Table 3 examines the impact of applying the redlining screen to randomly generated negative data (as described in Part IV). The numbers in Table 3 are lower compared to Table 2, although these statistical results still imply a high level of unlawful conduct, with Columns (1) and (2) of Table 3 suggesting that approximately eight to eighteen percent of all MSA-lender pairs have redlined between 2012 and 2016 and approximately eighteen to thirty-two percent of all lenders have engaged in such unlawful conduct in at least one MSA. Using a two-sided test as the baseline test, Columns (3) through (6) of Table 3 show that the percent of lenders flagged decreases as the required number of years of statistical significance increases and at a faster rate compared to Table 2. In fact, the number of lenders flagged by the screen is equal to zero if a lender's rate of lending must be significantly different from its peer group in three or more years. The smaller number of lenders flagged is due, in part, to the fact that lending in majority-minority tracts is not correlated within lender-MSA groups. This lack of correlation is an unrealistic feature of the randomly generated data compared to the HMDA dataset in which the intraclass correlation with respect to lender-MSA-level variables is positive over time.[98]

To account for this positive correlation, sets of random data are generated with varying degrees of intraclass correlation. As above, Table 4 demonstrates that the percent of lenders flagged decreases as the required number of years of statistical significance increases. If the intraclass correlation is 0.75, for example, then the number of lenders flagged (using a two-sided test) decreases from 1811 to 82. Unlike in Table 3, however, the total number of lenders flagged does not decrease to zero. The screen still yields potential targets, even if the difference between a lender's rate of loan applications received from majority-minority areas and that of its peer institutions must be statistically significant in all five years. Moreover, as the level of intraclass correlation increases, the number of targets increases. For exam-

[98] Specifically, the pooled chi-squared test statistic for each lender-MSA pair is positively correlated over time, and the intraclass correlation is calculated as 0.755. The intraclass correlation describes how strongly units in the same group resemble each other. Here, the intraclass correlation is used to quantify the degree to which a given lender's rate of loan applications received from majority-minority areas resembles other rates over time.

ple, if the data is perfectly correlated over time within MSA-lender pairs, then, assuming that a lender's rate of loan applications received from majority-minority areas must be significantly different from its peer group in all five years, the screening analysis suggests that approximately two percent of all MSA-lender pairs have redlined and approximately six percent of all lenders have engaged in this conduct in at least one MSA.

Recall that the data to which the screen is applied was constructed such that there are no "true" offenders. Any statistically significant deviation from the mean MSA-wide rate of lending in a given time period is due purely to chance and not to truly discriminatory conduct. In the best-case scenario, a statistical screen would account for this fact and not flag any lenders (or flag only a relatively small number of lenders). Instead, as Table 4 shows, the screen falsely suggests that at least six percent of all lenders in the United States are engaged in unlawful discrimination in the residential mortgage market. This subset of flagged lenders represents the statistical error produced if multiple hypotheses are tested simultaneously on the same dataset. By conducting multiple tests on a single dataset, looping over approximately 17,500 lender-MSA pairs, and then highlighting only those tests that yield statistically significant results at a five percent confidence level, Tables 3 and 4 demonstrate how it is possible to present false positives as real results, for almost any data with any measure of randomness is likely to contain spurious correlations.

B. Aggregated Performance Metrics

Arguably, the use of an annual performance metric is misplaced and should be replaced by a metric that aggregates information over a multiyear time period.[99] For instance, lending activity may be better examined aggregated over a five- or six-year time period. The use of an annual test deprives lenders of a certain measure of flexibility over time. A lender, for example, may choose to initiate and pursue a loan outreach program conducted over a two-year period, expanding its mortgage lending business into previously underserved majority-minority areas. In these years, the proportion of loan

[99] That lender-MSA-level variables in the HMDA dataset are positively correlated over time is not terribly surprising. A lender's business model, which includes strategic decisions about advertising and marketing, branch locations, underwriting and pricing procedures, the availability of loan products, and so forth, is unlikely to change abruptly from year to year. Instead, a lender is more likely to commit to a particular business model, typically making only relatively minor adjustments to its business operations such that overall lending outcomes remain relatively stable over time. Accordingly, if a lender's business decisions result in a distribution of home mortgage loan applications that supports a finding of statistical redlining in one year, then statistically significant results would be expected in the subsequent year (or years) as well. In other words, statistically significant differences are not randomly distributed across all observations. Rather, if the difference between a lender's rate of minority lending and that of its peer institutions is statistically significant in one year, then the probability that this difference is statistically significant in the following year is greater than would be the case if observations were perfectly uncorrelated (as in Table 3).

applications received from majority-minority tracts by this lender might exceed the corresponding percentages of its peer institutions by a considerable margin. And yet, under an annual performance metric, the lender receives no credit for "excess" loans as this lender would under a performance metric that aggregates lending data over time. Suppose that the lender temporarily curtails its outreach efforts for a year or two. The screen might very well flag this lender as a potential target in these years, even if the lender had vastly outperformed its peers in terms of minority lending when total lending activity is measured over a longer multiyear time period (because of the large volume of home mortgage loans received from majority-minority areas during its outreach program).

The use of an annual metric is further flawed insofar as there is relatively more opportunity for a lender to "game" the screen under this type of metric. Assume, for example, that the difference between a lender's rate of loan applications received from majority-minority areas and the corresponding lending rate of its peer institutions must be statistically significant in all five years for that lender to be flagged by the screen. To avoid detection, a lender, therefore, need only ensure its lending in majority-minority communities is sufficiently high in one or two years out of every five. In fact, the lender could receive zero loan applications from majority-minority census tracks in these "off" years for purely discriminatory reasons and still avoid detection. Under an aggregated metric, by contrast, where lending activity is assessed in the aggregate over a multiyear time period, a lender that is truly redlining would not be able to evade detection by sporadically providing residential mortgage credit to majority-minority communities (i.e., by gaming the screen).

Table 5 reports results using one- and two-sided tests (at a ninety-five percent significance level) on HMDA data aggregated over 2012 to 2016. Columns (1) and (2) of Table 5 examine lender-MSA pairs where the minimum total number of applications received from 2012 to 2016 is equal to or greater than 100, and there are no annual thresholds applied. Columns (3) and (4) introduce annual thresholds; specifically, the minimum total number of applications received is equal to or greater than two. Finally, Columns (5) and (6) restore the annual threshold to twenty applications received annually and drops the 100 application threshold for aggregate application totals. The results are consistent across all six columns. Using a two-sided test and HMDA data, Table 5 suggests that approximately thirty-three to thirty-six percent of all MSA-lender pairs have engaged in statistical redlining between 2012 and 2016 and approximately fifty-four to fifty-five percent of all distinct lenders have engaged in such unlawful conduct in at least one MSA over this time period. By contrast, using randomly generated data (as described in Part IV), Table 5 indicates that approximately two to three percent of all MSA-lender pairs have engaged in statistical redlining between 2012 and 2016 and approximately seven to eight percent of all lenders have engaged in such unlawful conduct in at least one MSA. Thus, as the bottom two rows of Table 5 show, the ratio of the number of MSA-lender pairs flagged using HMDA data to the number of MSA-lender pairs flagged using randomly-generated data is approximately seven and the ratio of the number of lenders flagged using HMDA data to the number of lenders flagged using randomly-generated data is approximately twelve. Roughly speaking, these ratios suggest that one out of every thirteen MSA-lender pairs, or one out of every eight lenders, is flagged as a direct consequence of improper statistical techniques and not true discriminatory behavior.

The ratios in Table 5 appear high and continue to imply a surprisingly high degree of unlawful activity on the part of lending institutions. For a large number of lenders, the percentage of home mortgage loan applications received from majority-minority census tracts is significantly different from peer institutions—a finding that, as shown in Table 6, is highly robust to model specification. Gaining a better understanding of why this finding is true and the extent to which these differences in lending rates are truly a product of discrimination—or some other feature of the residential mortgage market—is an important topic for future research not resolved here. Rather, the purpose of this study is simply to point out that, with full information disclosure by regulatory agencies, the number of targets flagged using actual data corresponds in some ratio to the number of targets flagged using random negative data. In certain situations, this ratio may be close to one, implying that random sampling error is a primary explanation for why certain firms have been flagged by a data mining exercise that simultaneously tests thousands of statistical hypotheses. For example, the rate of workplace accidents for all firms in a particular industry may be approximately equal to the significance level applied in the screen, yielding ratios that are relatively low (and close to one). In other cases, such as here, the estimated ratios may be relatively higher (and greater than one), suggesting that a large majority of the potential targets flagged by the screen are the product, not of spurious correlation, but of truly unlawful behavior.

Even if the ratios are large, however, it is still necessary to distinguish true positives from false positives. Table 5 suggests, for example, that between approximately seven to fifteen percent of the lenders flagged by the screen is due entirely to chance. The question is how to determine if a lender targeted by the screen is a true violator of the law or simply an unlucky statistical outlier. Qualitative evidence plays a vital role at this stage of the analysis. Recall that other qualitative evidence is frequently relied upon, in addition to the core statistical findings, in alleging a prima facie case of redlining. Simply relying upon statistical significance, although much more likely than not to reveal a true positive given the relatively high ratios in Table 5, does not squarely address the problem that a lender may have been flagged solely as the result of inadequate attention paid to the problem of multiple simultaneous statistical comparisons. Accordingly, a plaintiff will generally couple the necessary statistical findings with qualitative evidence, alleging that the defendant-lender has avoided providing residential mortgage loans to residents of majority-minority neighborhoods by (1) configuring the geography delineated as its CRA assessment area to include primarily majority-minority residential areas, (2) systematically locating its branch offices in primarily non-majority-minority areas, and/or (3) strategically choosing to market its lending products to residents of non-majority-minority neighborhoods.[100] In theory, this type of additional qualitative evidence can be used to distinguish true offenders among a list of potential targets generated by a statistical screening analysis.

In practice, however, such qualitative evidence has become less probative of redlining as a result of the growing pervasiveness of digital and online technologies. As a greater number of consumers choose to use digital channels to conduct private banking business, avoiding traditional physical channels altogether, the location of branch offices is far less likely to influence the final geographic distribution of loan applications.[101] Moreover, to the extent that digital technology has allowed lenders to market their credit products on a more global scale throughout an entire MSA, advertising and other marketing efforts are less likely to target only specific geographic locations than in the past, when consumers primarily obtained information about credit availability from community newspapers or other local sources.[102] Hence, Table 5 highlights an increasingly important policy problem. Although the proportion of true positives generated by a redlining screening analysis is likely to be high, much of the qualitative evidence used to isolate truly unlawful conduct has no probative value. In addition to improperly drawn CRA assessment areas, regulators and other activists must, therefore, look to new, additional types of qualitative evidence, such as consumer complaints or greater use of paired testing, to isolate true positives.

Table 6 explores the extent to which the baseline specification in Table 5 is sensitive to model specification. Column (1) replaces loan applications with loan originations. Column (2) modifies the peer selection process by no longer requiring that a lender have at least ten peers in a MSA. Instead, peer institutions are selected solely on the basis of the 50% to 200% rule (described in Part IV). Column (3) examines only depository institutions, dropping financial institutions that report either the National Credit Union Administration or HUD as its supervisory or regulatory agency. Column (4) substitutes majority-Black census tracts for majority-minority census tracts. Column (5) considers conventional loans only and excludes all FHA-insured, VA-guaranteed, and FSA/RHS loans.[103] Finally, Column (6) limits the

[100] *See infra* Part IV.A.

[101] *See* Susannah Fox, *51% of U.S. Adults Bank Online*, Pew Research Ctr. (Aug. 7, 2013), http://www.pewinternet.org/2013/08/07/51-of-u-s-adults-bank-online/ [https://perma.cc/6FXA-VZBQ] (finding that fifty-one percent of U.S. adults bank online and thirty-two percent of adults transact bank business on their mobile phones).

[102] *See generally* RonNell Andersen Jones, *Litigation, Legislation, and Democracy in a Post-Newspaper America*, 68 Wash. & Lee L. Rev. 557 (2011).

[103] FSA/RHS denotes loans made and guaranteed by the Department of Agriculture's Farm Service Agency or Rural Housing Service.

scope of the redlining analysis to middle- and upper-income census tracts as determined annually by the FFIEC. The motivation for this specification is that certain "niche" lenders may choose to provide credit based upon income level, which is not a protected category under the FHA or the ECOA, in areas where income is positively correlated with non-majority-minority status.[104] To control for the confounding influence of income, low- and moderate-income census tracts are dropped. The redlining analysis then considers the proportion of lending in majority-minority tracts only across those tracts identified as middle- or high-income. As Table 6 shows, the statistical results are remarkably robust to varying model specification.

VI. Information Disclosure by Regulatory Agencies

A regulatory agency can choose between three possible models regarding the disclosure of private information collected while operating in its supervisory capacity: (1) full information disclosure, (2) no information disclosure, or (3) limited information disclosure. This choice requires an agency to tradeoff the over-enforcement of false positives, in the case of full information disclosure, with the under-enforcement of true positives (and the loss of a data source that may be useful for other purposes), in the cases of limited or no information disclosure.[105] The goal of this Article is to highlight that such a tradeoff exists and to caution that full information disclosure by regulatory agencies does not come without costs. Given full access to an administrative dataset, it is possible to employ data mining techniques to locate a statistical outlier that is then improperly characterized as evidence of a violation of the law. As discussed in Part V, statistical outliers must not automatically be taken as proof of illegality. There will always be statistical outliers in the data; there will not always be statistical evidence of actual unlawful conduct.

A. *Three Models of Information Disclosure*

This Section examines the relative advantages and disadvantages of three possible models of information disclosure by regulatory agencies.

---

[104] *See*, *e.g.*, Rachel Louise Ensign, AnnaMaria Andriotis & Paul Overberg, *First Republic: Is It Wrong to Build a Bank for Wealthy Clients Only?*, Wall St. J. (Aug. 16, 2016), https://www.wsj.com/articles/first-republic-is-it-wrong-to-build-a-bank-for-wealthy-clients-only-1471388308 [https://perma.cc/JG5L-P4ZE].

[105] *See* Talbot Page, *A Generic View of Toxic Chemicals and Similar Risks*, 7 Ecology L.Q. 207, 219–20 (1978) (arguing that a regulatory agency should err on the side of preventing false negatives at the expense of some false positives); *cf.* Sidney A. Shapiro, *Keeping the Baby and Throwing Out the Bathwater: Justice Breyer's Critique of Regulation*, 8 Admin. L.J. 721, 732 (1995) (noting that when a regulator makes a decision under conditions of uncertainty, there are two possible types of error: the regulator can overregulate a risk that proves to be insignificant, or the regulator can under-regulate a risk that proves to be significant).

1. Full Information Disclosure

A model of full information disclosure finds support in the notion that regulatory agencies should tap into the massive parallel-processing power of the public by providing private citizens with the information necessary to track, analyze, and publicize corporate misbehavior. One of the primary rationales for the enactment of HMDA was to enhance the public enforcement of laws prohibiting discrimination in lending by requiring federal regulatory agencies to collect and publicly disclose data on borrower and loan characteristics.[106] Through public disclosure of mortgage data, HMDA envisions a strong role for citizen monitors whose "regulation from below" is expressly intended to augment enforcement efforts by traditional financial regulatory agencies.[107] HMDA arose out of growing public concern in the 1970s over mortgage redlining and the negative effects of disinvestment in older urban neighborhoods.[108] With the deterioration of many minority communities caused, in part, by systematic redlining, the original proponents of HMDA rightly anticipated that a standardized reporting regime, required by federal mandate, would provide direct statistical evidence to support otherwise more anecdotal or qualitative complaints about the prevalence of redlining in minority communities.[109] Providing free, standardized disclosures in the form of mortgage loan information, made publicly available on a disaggregated individual loan level, was intended to provide not only the public with better information on internal corporate practices, but also lenders an incentive to discontinue the redlining of minority neighborhoods—or otherwise risk reputational harm (and litigation).[110]

Although full information disclosure by regulatory agencies tends to promote greater transparency and accountability, this Article has identified a potential problem with this model of information disclosure: improper *p*-hacking by means of multiple comparisons. Using redlining as a motivating example, Part V demonstrated how it is possible to use simple data mining techniques to uncover patterns in publicly available administrative data that can be spuriously presented as statistically significant. This form of data mining need not be malicious, fraudulent, or even intentional. In fact, it is

[106] *See* Allen J. Fishbein, *The Ongoing Experiment with "Regulation from Below": Expanded Reporting Requirements for HMDA and CRA*, 3 Housing Pol'y Debate 601, 615 (1993).

[107] *See id.*; *see also* Patricia McCoy, *The Home Mortgage Disclosure Act: A Synopsis and Recent Legislative History*, 29 J. Real Est. Res. 381 (2007).

[108] *See* Joseph M. Kolar & Jonathan D. Jerison, *The Home Mortgage Disclosure Act: Its History, Evolution, and Limitations*, 59 Consumer Fin. L.Q. Rep. 189, 195 (2005).

[109] *See id.*

[110] *See* Archon Fung, Mary Graham & David Weil, Full Disclosure: The Perils and Promise of Transparency 203–05 (2007). *See generally* Louis D. Brandeis, Other People's Money and How the Bankers Use It (1914).

possible that the same outcome might obtain, in the aggregate, even if no single individual actor considers more than one hypothesis.[111]

Full disclosure may be even more disadvantageous insofar as it pushes activists (or regulators) to find statistical correlations in the data that are largely unrelated to the underlying nature of the offense as originally envisioned by lawmakers. To illustrate, suppose there are ten lenders presently doing business in a MSA, and all ten have purposefully chosen not to receive loan applications from majority-minority areas. Residents in these areas are, therefore, severely limited in their access to home mortgage credit. And yet, the screening procedure set forth above would not identify a single one of these ten lenders as having engaged in redlining. Each lender, when individually compared to its peer lenders—or the rest of the market, for that matter—is generating the same proportion of loan applications from majority-minority areas: zero percent. By contrast, suppose that nine of the ten lenders receive loan applications in majority-minority areas. The remaining lender chooses not to lend in these areas not because of improper discriminatory motives but because the lender rightly perceives that market demand for credit is satisfied in these communities such that it would not be profitable to do business in this part of the MSA. Unlike above, where all ten lenders in the MSA have purposefully chosen to avoid lending in majority-minority areas (perhaps due entirely to racial animus), the redlining screen in this case may very well flag this sole lender as having engaged in redlining, even if the relevant credit markets are not underserved in any meaningful respect. In this way, full information disclosure might lead to an overreliance on quantitative analyses and may cause citizen activists (or regulators) to focus on exactly the wrong factual circumstances, awarding outsized or misplaced focus to statistical outliers and ignoring communities that have been truly fenced-off by a significant majority of existing lenders in the MSA.

---

[111] To amplify, suppose that a single customer of each lender in each lending market in the country uses the HMDA data to test the hypothesis—in the same manner as set forth in Part V—that her lender is redlining. Note that the net result of this aggregate behavior is equivalent to the standard case of multiple simultaneous statistical comparisons where a single individual tests many individual hypotheses at the same time. The impact upon statistical significance is the same whether a different actor tests each individual hypothesis separately or the same actor tests each individual hypothesis simultaneously. The added wrinkle in this hypothetical scenario, however, is that there is no need, in theory, for each individual customer to make a correction for multiple comparisons as only one hypothesis is tested: the individual hypothesis that her lender is engaged in unlawful redlining. And yet, again, just as when a single individual makes multiple comparisons, that some non-trivial subset of customers finds statistical support for the hypothesis of discrimination is not surprising. When enough hypotheses are tested, either by a single actor or individually by many different actors, it is practically certain that some number will prove statistically significant, as almost any dataset with any measure of randomness is likely to contain spurious correlations.

2. *No Information Disclosure*

As one response to the problem of multiple comparisons highlighted in Part V, a regulatory agency might choose not to disclose publicly any (or very little) private information collected while operating in its supervisory capacity, thereby depriving the public of the data from which the very problem of multiple comparisons arises. Indeed, despite the potential benefits described above, the use of public disclosure as a means by which to regulate market failures is not without its critics. The mortgage industry, for instance, fiercely resisted the original enactment of HMDA, believing itself to have been unfairly singled out as the primary culprit responsible for the continuing deterioration of minority neighborhoods, and objected to the increased regulatory burdens created by the new reporting requirements.[112] Echoing the objections of private industry, certain members of Congress likewise argued that HMDA was likely to result in credit misallocation and would create unnecessary compliance costs for lenders.[113] This pushback had a tangible effect—HMDA is today much more limited in scope than would be the case under a truly full disclosure model of mortgage lending activity.[114]

This Article claims that the primary disadvantage of a no information disclosure model is that institutions become severely restricted in their capacity to implement internal corporate compliance programs.[115] A lending institution cannot accurately assess how it is performing relative to its peer institutions, in terms of a potential redlining violation, without access to the lending rates of its peer institutions.[116] It is true that lenders could conceivably police themselves. A private firm could collect information on mortgage lending activity and distribute this data to member institutions in exchange for a monetary fee. In some ways, this national, proprietary, fee-driven database approximates many of the benefits of the limited information disclosure model introduced next. But, there are obvious conflicts of interest here rendering it unlikely that such a proprietary database will succeed. A

---

[112] *See generally* Allen Fishbein & Ren Essene, *Moving Forward: The Home Mortgage Disclosure Act at Thirty-Five: Past History, Current Issues* (Joint Ctr. for Housing Studies, Harvard Univ., Working Paper MF10-7, 2010), *available at* http://www.jchs.harvard.edu/sites/default/files/mf10-7.pdf [https://perma.cc/SEX3-W8FG].

[113] *See id.*

[114] There is no information, for example, relating to small business credit or consumer deposit account data. Also, the HMDA dataset does not include rural loans made by non-urban lenders.

[115] *See, e.g.*, Miriam Hechler Baer, *Governing Corporate Compliance*, 50 B.C. L. Rev. 949, 953 (2009); David Hess & Cristie Ford, *Corporate Corruption and Reform Undertakings: A New Approach to an Old Problem*, 41 Cornell Int'l L.J. 307, 312 (2008).

[116] Recall that a lending institution is flagged as potentially engaging in redlining if there exists a statistically significant difference between the rate of loan applications received from majority-minority census tracts and the corresponding rate of its peer institutions. To assess whether or not it may be flagged, a lender must therefore know the lending rates of its peer institutions.

lender, for example, would presumably provide its lending data to such a private firm in exchange for professional guidance as how to best manage lending risk. However, a lender might be justifiably concerned about sharing its corporate data for the purpose of managing risk with a private firm that also makes this information readily available to litigants as evidence of discrimination in fair lending disputes. Without access to industry-wide data, lenders are, as a consequence, left vulnerable to attack by litigants. A lender might find itself the target of a redlining lawsuit and have no knowledge as to how its lending practices compare to an applicable peer group. Full (or limited) information disclosure by regulatory agencies allows lenders to take focused, proactive corporate compliance measures to avoid unlawful behavior, managing both litigation and reputational risk by closely monitoring the distribution of residential loan applications—both its own as well as that of its peer institutions—particularly in minority communities.

A model of no information disclosure also places tremendous weight on regulatory agencies to screen for the presence of redlining. Regulators cannot always be relied upon, however, to police private market failures for reasons that are well known at this point, including capture by private interest groups or vulnerability to cognitive bias.[117] Further, there is always the possibility a prosecutor will be overly aggressive, misinformed, or even unethical.[118] The broad willingness to experiment with public disclosure and "regulation from below" in the original enactment of HMDA reflected broad "displeasure with the slowness of the regulatory response to unlawful redlining."[119] "The prevailing congressional view was that the traditional regulatory apparatus was insufficiently engaged in efforts to deter redlining, and consequently more vigorous action was needed through the elevation of the role of citizen monitors."[120] The no disclosure paradigm cuts sharply against the positive potential of grassroots activism and, with respect to HMDA specifically, denies citizens the information necessary to assess if lending institutions are failing to supply sufficient home mortgage credit to minority communities.

In the end, the no information disclosure model likely embodies an overly far-reaching response to a technical problem that is fixable through other methods. As discussed in Part II, there exist a number of sensible remedies to the problem of multiple comparisons, including the use of adjusted *p*-values or randomized out-of-sample tests. Abolishing all data collection

---

[117] *See, e.g.*, William N. Eskridge Jr. & John Ferejohn, *Structuring Lawmaking to Reduce Cognitive Bias: A Critical View*, 87 CORNELL L. REV. 616, 620–23 (2002).

[118] David Keenan et al., *The Myth of Prosecutorial Accountability After Connick v. Thompson: Why Existing Professional Responsibility Measures Cannot Protect Against Prosecutorial Misconduct*, 121 YALE L.J. ONLINE 203 (2011), http://yalelawjournal.org/forum/the-myth-of-prosecutorial-accountability-after-connick-v-thompson-why-existing-professional-responsibility-measures-cannot-protect-against-prosecutorial-misconduct [https://perma.cc/PQ9Y-9BN6].

[119] *See* Fishbein & Essene, *supra* note 112, at 10.

[120] *See id.* at 14.

efforts, or confining the data collected simply for use by government actors only, reflects an unduly pessimistic view of the capacity of public and private actors to properly control for the adverse consequences of improper *p*-hacking.

3. *Limited Information Disclosure*

In response to the problem of *p*-hacking identified in Part V, a regulatory agency could continue to collect data on private actors in its supervisory capacities but limit the subsequent disclosure of this data to the general public. For instance, with respect to information collected on home mortgage lending, financial regulators under this particular disclosure model would still continue to collect data from lending institutions in exactly the same way as they presently do. The sole difference is that the full dataset would be released only to those lenders that submitted loan data to the government and would not be more broadly disclosed to the public.[121] The primary advantage of this model of limited information disclosure is that lenders can conduct internal corporate compliance, but other private actors cannot use the improper data mining techniques highlighted in Part V to uncover spurious correlations in the publicly disclosed data.[122]

Finally, to establish which of the following three models of information disclosure is optimal in a given regulatory context, the type of ratios introduced in Part V are instructive. In the case of redlining, the ratios reported in Tables 5-7 are relatively high, suggesting that full information disclosure is optimal. Many regulatory agencies have limited resources, and, in some administrations, limited motivation to enforce certain civil rights laws.[123] Private enforcement is a substantial benefit in this regard, both as a supplement to, and as a spur for, public enforcement. In addition, HMDA data is used by many parties other than fair lending litigants; for instance, investors and academics rely upon HMDA as an important data source in better understanding

[121] The data could be shared with other regulatory agencies, as well as with researchers deemed likely to understand the problem of multiple comparisons and investors who are seeking greater market knowledge and not potential litigation targets.

[122] As for regulatory agencies themselves, in using the data as a means to screen for the presence of redlining, regulators can normally be relied upon to make the necessary corrections to avoid the negative consequences of improper *p*-hacking. Regulatory agencies have experts who are expected to understand the problem of multiple comparisons and possess the technical knowledge and expertise to make the proper adjustments. Of course, this might not always hold true in each case, and a valid argument can be made that regulators should be treated just like private litigants. For instance, it might make sense for certain regulatory agencies, such as the CFPB, to construct an information barrier (commonly referred to as a "Chinese wall") between the agency's enforcement and market monitoring functions. This barrier would be designed to prevent exchanges of information (or other communications) that could lead to litigation in which enforcement tests its claim (or hypothesis) of discrimination by an individual lender on the very same dataset that was used by market monitoring to generate the hypothesis in the first place.

[123] *See generally* Zachary S. Price, *Politics of Nonenforcement*, 65 CASE W. RES. L. REV. 1119 (2015).

the operation of residential mortgage markets.[124] In this case, the benefits of added enforcement and greater market knowledge likely offset the costs of improper data mining. If the reported ratios are relatively low (and close to one), by contrast, then full information disclosure is likely *not* optimal, because the costs of improper *p*-hacking by means of multiple comparisons become sufficiently high such that these costs, in all likelihood, exceed the benefits of added enforcement and greater market knowledge.

B. The Role of Litigation Under Different Models of Information Disclosure

The role of litigation differs under varying models of information disclosure by regulatory agencies. Most interestingly, under a model of limited, or no, information disclosure, litigation serves as a formal means to correctly implement statistical hypothesis testing and the scientific method more generally. Rather than first *p*-hack data made publicly available by a regulatory agency to arrive at a potential litigation hypothesis, under a model of restricted information disclosure, a plaintiff starts with a single testable litigation hypothesis: a given lender has engaged in redlining in a given MSA during a given time period in violation of the FHA and ECOA. This hypothesis can be generated through customer complaints or other qualitative evidence of discrimination and is determined prior to examination of any quantitative lending data. Importantly, that the litigation hypothesis is generated by means of qualitative evidence implies that this hypothesis will never be tested on the same data used to generate this hypothesis. To test its litigation hypothesis given limited publicly disclosed data by regulatory agencies, a plaintiff must use discovery to obtain the quantitative information necessary to assess the statistical significance of its original hypothesis. During the discovery process, the defendant-lender must voluntarily turn over to the plaintiff the data necessary to test the hypothesis that there exists a statistically significant difference between the defendant's lending in majority-minority areas and such lending by its peer institutions. This data must include not only the loan application data of the lender but also the loan application data for the defendant-lender's peer group. In all likelihood, this information should be readily available to the plaintiff (in electronic format), as this type of data is necessary for any lender to implement a meaningful corporate compliance program related to fair lending.[125]

To reiterate, the main advantage of a model of limited, or no, information disclosure by regulatory agencies is that, by construction, the data used during the course of litigation to test the claim that a lender has failed to

---

[124] *See, e.g.*, Brent Smith, *Turmoil in the Residential Mortgage Market: A Review and Compilation of Research and Policy*, 19 J. HOUSING RES. 65 (2010).

[125] *See, e.g.*, Kenneth A. Bamberger, *Technologies of Compliance: Risk and Regulation in a Digital Age*, 88 TEX. L. REV. 669, 685–702 (2009).

provide sufficient credit to minority areas *cannot* also be first used to identify or distinguish the lender as one against whom such a claim should be pled. Unlike under a model of full information disclosure, a potential litigant cannot *p*-hack publicly available data under a model of limited, or no, information disclosure to generate a set of hypotheses that will be quite predictably confirmed if later tested on the same data during the course of a lawsuit. Rather, a plaintiff must formulate a hypothesis prior to examination of the quantitative data and obtain the information necessary to test this hypothesis only through discovery. Under limited, or no, information disclosure, a plaintiff cannot simply download data made freely available to the public by regulatory agencies and run the applicable statistical analysis. Instead, the plaintiff must state a claim against a lender and subsequently acquire the necessary quantitative information to prove this claim by incurring the costs of discovery, which are often large.[126] In this way, restricted information disclosure by regulatory agencies (as has historically been the case), coupled with costly discovery, ensures that hypotheses (concerning violations of the law) are tested in a proper application of the scientific method, and statistical hypothesis testing in particular, with the formulation of a theory of liability *preceding* empirical analysis of the data ultimately used to test that theory.[127]

Of course, the main disadvantage of restricted information disclosure by regulatory agencies is that it is now more difficult for a plaintiff to state a discrimination claim. Under a full information disclosure model, a plaintiff can state a claim of redlining against a lender by using publicly disclosed HMDA data to demonstrate that there exists a statistically significant difference in lending rates in majority-minority areas.[128] Under a restricted information disclosure model, by contrast, a plaintiff cannot include in the complaint such direct statistical evidence and must, instead, rely upon relatively less compelling circumstantial or qualitative evidence.[129]

Recent lawsuits filed by municipalities against national banks in response to the large number of residential foreclosures that accompanied the Great Recession to recover for harms suffered as a result of alleged discrimi-

[126] *See generally* Martin H. Redish & Colleen McNamara, *Back to the Future: Discovery Cost Allocation and Modern Procedural Theory*, 79 Geo. Wash. L. Rev. 773 (2011).

[127] *See, e.g.*, Bryman, *supra* note 3.

[128] *See supra* Part IV.A.

[129] *See* Roy L. Brooks, Conley *and* Twombly: *A Critical Race Theory Perspective*, 52 How. L.J. 31, 58 (2008) (contending that the applicable pleading standard "disadvantages the prosecution of civil rights cases by imposing a difficult, if not impossible, burden on the plaintiff to make specific factual allegations about evidence (or 'proof') known only to defendants"); *see also* Douglas A. Blaze, *Presumed Frivolous: Application of Stringent Pleading Requirements in Civil Rights Litigation*, 31 Wm. & Mary L. Rev. 935, 957 (1990) (noting that plaintiff would not normally have the requisite factual predicate to show a "custom and practice" of discrimination pre-discovery, making it "nearly impossible" for a civil rights claim to escape 12(b)(6) dismissal); A. Benjamin Spencer, *Understanding Pleading Doctrine*, 108 Mich. L. Rev. 1, 26 (2009) ("[A] standard that dismisses valid claims at the very front end of the system based on an inability to offer facts that claimants are, at this early stage, unlikely or unable to know blocks access to the courts in a way that is fundamentally improper.").

natory mortgage lending practices show the difficulty of surviving a motion to dismiss in the face of restricted information disclosure and a relatively heightened pleading standard. In *Mayor and City Council of Baltimore v. Wells Fargo Bank*,[130] for example, the U.S. District Court for the District of Maryland concluded that the total number of properties located in Black neighborhoods on which Wells Fargo made loans that were foreclosed upon—163 in total—was insufficient to demonstrate the plausibility of a causal connection between the bank's alleged discriminatory activities and the generalized type of damages claimed.[131] Likewise, in *City of Birmingham v. Citigroup, Inc.*,[132] the district court granted the bank's motion to dismiss, contending that the city's pleadings relied upon merely "a series of speculative inferences,"[133] while the court itself speculated that minority borrowers likely defaulted on their home mortgage loans as a result of individualized economic problems.[134] Recognizing the difficulty faced by potential victims of redlining, courts should, to the full extent possible, apply a pleading standard that does not require the plaintiff to make precise statistical allegations based upon lending data that will generally not be available to the plaintiff prior to discovery under a model of limited, or no, information disclosure.

Notably, the Supreme Court has made statements that appear to lend support to such an approach, stating that a plaintiff in a discrimination case must allege, at the pleading stage, only a short and plain statement of the claim so as to provide the defendant with fair notice of the claim and the corresponding grounds upon which the claim rests.[135] The Court has held that a plaintiff is not required to plead all of the facts that will ultimately be used by the plaintiff to prove its prima facie case of discrimination.[136] Nor is a plaintiff required to show how the defendant's asserted explanations are merely pre-textual.[137] Accordingly, in the context of redlining, the appropriate application of the pleading standard might, in theory, allow a plaintiff to

[130] 677 F. Supp. 2d 847 (D. Md. 2010).

[131] *See id.* at 850.

[132] No. CV-09-BE-467-S, 2009 WL 8652915 (N.D. Ala. Aug. 19, 2009).

[133] *Id.* at *3.

[134] *See id.* at *4 ("[I]t is quite speculative that the depreciation in value of the neighboring homes in the City was caused by the foreclosures of minority borrowers' properties rather than as a result of 'a myriad of other factors,' which . . . could include 'rising unemployment in the region, changes in the housing market, or other economic conditions.'") (quoting Tingley v. Beazer Homes Corp., No. 3:07–CV–176, 2008 WL 1902108, at *5 (W.D.N.C. Apr. 25, 2008)).

[135] *See* Swierkiewicz v. Soreman, 534 U.S. 506, 512 (2002); *see also* United States v. Union Auto Sales, Inc., 490 Fed. App'x 847, 848 (9th Cir. 2012) (reversing dismissal of discrimination claim under ECOA and stating that, at the pleading stage, plaintiff "is not required to demonstrate discrimination, but merely to allege facts sufficient to make a discrimination claim plausible").

[136] *See id.* at 510 ("The prima facie case . . . is an evidentiary standard, not a pleading requirement."); *see also* Ring v. First Interstate Mortg., Inc., 984 F.2d 924, 926 (8th Cir. 1993) ("Under the Federal Rules of Civil Procedure, an evidentiary standard is not a proper measure of whether a complaint fails to state a claim.").

[137] *See, e.g.*, United States v. Badgett, 976 F.2d 1176, 1178 (8th Cir. 1992).

plead a prima facie case purely on the basis of certain non-statistical factual allegations, such as the lender's strategic decision to primarily market its residential mortgage lending products to residents of majority-white neighborhoods to the exclusion of majority-minority neighborhoods. In the complaint, the plaintiff might further state, as a factual matter, that the mortgage lending data to be later obtained in discovery from the defendant will *prove* its prima facie case of redlining, demonstrating, at a significance level of at least five percent, that the defendant-lender has generated disproportionately fewer loan applications from majority-minority areas relative to its peer lenders in the same market.

In theory, pleading standards strive to balance different considerations that argue for, or against, creating a barrier to discovery—and that balancing act is especially delicate in the case of a disparate impact claim.[138] On the one hand, the basic allegation in a disparate impact case is a statistical one. However, those statistics could be evidence of a random correlation, rather than a truly meaningful one, even if the likelihood of that is small. Also, given the large reputational harm to a defendant from a discriminatory lending claim, there is a strong incentive to settle if the complaint survives the motion-to-dismiss stage.[139] These factors leave businesses vulnerable to ongoing harassment from litigation based upon spurious statistical correlations. On the other hand, however, there must not be unlawful discrimination in the marketplace. The theoretically optimal pleading standard in a disparate impact claim must strike the right balance between these competing factors. This optimal balancing, however, may vary depending upon the existing model of information disclosure. If the existing model is full information disclosure, for example, then a special pleading standard may be required.[140] For instance, if a private litigant finds a statistical correlation, then the complaint could be filed under seal, initially, to minimize the reputational harm to the defendant, and a regulatory agency could be required to investigate further, enabling "regulatory discovery."[141] Assuming that the agency finds further evidence of discrimination, the suit could then be unsealed and permitted to proceed. Otherwise, the lawsuit would be dismissed. If, by contrast, the existing model is limited, or no, information disclosure, then the correct pleading standard, when there is no other relevant external data available, might allow plaintiffs "limited discovery" on any missing data in the defendant's possession necessary to establish the alleged claim of disparate impact. The important point here is simply that the opti-

[138] *See generally* Martin H. Reddish, *Pleading, Discovery, and the Federal Rules: Exploring the Foundation of Modern Procedure*, 64 Fla. L. Rev. 845 (2012).

[139] *See, e.g.*, Marc Galanter, *The Vanishing Trial: An Examination of Trials and Related Matters*, 1 J. Empirical Legal Stud. 459, 468 (2004) (noting that "trials were 19.7 percent of all civil rights dispositions in 1970 and 3.8 percent in 2002.").

[140] The author thanks Professor Adam Levitin of Georgetown University Law Center for this insight.

[141] This "regulatory discovery" could proceed similarly to a HUD investigation conducted in response to a fair lending complaint.

mal pleading standard should correspond, or be sensitive, to the level of information disclosure by regulatory agencies.

VII. Conclusion

This Article has shown that data mining techniques can be used to locate statistical outliers that are incorrectly characterized as evidence of unlawful conduct and can lead to violations of the key statistical principle that a hypothesis should not be tested on the same data that was used to construct the hypothesis. The example considered at length is the redlining of majority-minority neighborhoods by lending institutions. Using home mortgage loan data made publicly available by financial regulators, the Article found that approximately three percent of all lender-MSA pairs (or approximately seven to nine percent of all lending institutions) flagged by a standard data mining exercise is attributable to a failure to correct for the multiple hypothesis testing problem. In addition, for a large number of lenders, the percentage of home mortgage loan applications received from majority-minority census tracts is significantly different from peer institutions—a finding that was robust to model specification. A better understanding as to why this finding is true, and the extent to which these differences in lending rates are truly a product of discrimination, or some other feature of the mortgage market, is left open as an important topic for future scholarly research.

Three models of information disclosure by regulatory agencies were considered: (1) full information disclosure, (2) no information disclosure, and (3) limited information disclosure. The primary advantage of a limited information disclosure model is that lenders can conduct internal corporate compliance, but other private actors cannot use improper data mining techniques to uncover spurious correlations in publicly disclosed administrative datasets. Litigation plays an important role under this disclosure model, serving as a means to correctly implement statistical hypothesis testing and the scientific method more generally. Under a model of limited information disclosure, a potential plaintiff cannot *p*-hack publicly available data to generate a set of hypotheses (concerning potential violations of the law) that will be confirmed if later tested on the same data during the course of a lawsuit. Rather, the plaintiff must formulate a hypothesis prior to examination of the data and obtain the data necessary to test this hypothesis only through discovery. In this way, limited information disclosure by regulatory agencies, coupled with costly discovery, ensures that hypotheses (regarding violations of the law) are tested in a proper application of the scientific method, with the formulation of a theory of liability preceding empirical analysis of the data ultimately used to test that theory.

Tables

Table 1: Descriptive Statistics for Potential Targets (2012-2016)

	Mean	**Stan. Dev.**	**Minimum**	**Maximum**	**Range**
Lender-Level					
Loan Applications	471	1,338	21	64,396	64,375
% Apps in Maj-Min Tracts	18.15%	21.19%	0.00%	100.00%	100.00%
Loan Originations	298	848	0	38,965	38,965
% Origs in Maj-Min Tracts	16.87%	21.05%	0.00%	100.00%	100.00%
MSA-Level					
Number of Tracts	124	263	4	3,392	3,388
% Maj-Min Tracts	27.81%	22.20%	0.78%	98.82%	98.04%
Loan Applications	68,560	81,926	44	442,141	442,097
% Apps in Maj-Min Tracts	18.96%	19.64%	0.00%	99.98%	99.98%
Loan Originations	43,078	50,971	18	280,445	280,427
% Origs in Maj-Min Tracts	17.21%	19.24%	0.00%	100.00%	100.00%
Number of MSAs			476		
Number of Lender-MSAs			17,568		
Number of Lender-MSA-Years			87,840		

Table 2: Potential Targets Using HMDA Data (2012-2016)

	1-Yr*	**1-Yr**	**2-Yr**	**3-Yr**	**4-Yr**	**5-Yr**
Targets						
Total Targets	8,290	7,021	4,778	3,426	2,423	1,571
% MSA-Bank	47.47%	40.21%	27.36%	19.62%	13.88%	9.00%
Unique Target	2,119	1,897	1,553	1,290	1,056	806
% Bank	61.60%	55.15%	45.15%	37.50%	30.70%	23.43%
Lending						
Mean Apps	600	655	748	929	823	896
% Min-Apps	15.79%	16.44%	15.87%	14.94%	14.29%	14.00%
Peers	49	50	52	50	54	55
Shortfall	20	24	35	48	58	72
MSAs						
# Tracts	147	154	178	211	224	268
% Min Tracts	27.31%	28.12%	30.02%	30.80%	30.50%	33.14%

TABLE 3: POTENTIAL TARGETS USING RANDOMLY-GENERATED DATA (2012-2016)

	1-Yr*	1-Yr	2-Yr	3-Yr	4-Yr	5-Yr
Targets						
Total Targets	3,118	1,464	48	0	0	0
% MSA-Bank	17.86%	8.38%	0.27%	0.00%	0.00%	0.00%
Unique Target	1,100	622	41	0	0	0
% Bank	31.98%	18.08%	1.19%	0.00%	0.00%	0.00%
Lending						
Mean Apps	575	649	809	-	-	-
% Min-Apps	21.91%	23.14%	25.75%	-	-	-
Peers	47	46	50	-	-	-
Shortfall	2	3	7	-	-	-
MSAs						
# Tracts	157	180	442	-	-	-
% Min Tracts	27.76%	28.91%	37.59%	-	-	-

TABLE 4: POTENTIAL TARGETS USING RANDOMLY-GENERATED DATA WITH POSITIVE INTRACLASS CORRELATION (ICC) (2012-2016)

	1-Yr*	1-Yr	2-Yr	3-Yr	4-Yr	5-Yr
75% ICC						
Total Targets	3,055	1,811	794	420	194	82
% MSA-Bank	17.50%	10.37%	4.55%	2.41%	1.11%	0.47%
Unique Target	1,126	789	425	260	140	64
% Bank	32.73%	22.94%	12.35%	7.56%	4.07%	1.86%
85% ICC						
Total Targets	2,181	1,209	607	351	216	97
% MSA-Bank	12.49%	6.92%	3.48%	2.01%	1.24%	0.56%
Unique Target	904	595	362	232	163	85
% Bank	26.28%	17.30%	10.52%	6.74%	4.74%	2.47%
95% ICC						
Total Targets	1,308	653	437	325	240	159
% MSA-Bank	7.49%	3.74%	2.50%	1.86%	1.37%	0.91%
Unique Target	618	363	270	214	172	124
% Bank	17.97%	10.55%	7.85%	6.22%	5.00%	3.60%
100% ICC						
Total Targets	688	312	312	312	312	312
% MSA-Bank	3.94%	1.79%	1.79%	1.79%	1.79%	1.79%
Unique Target	389	212	212	212	212	212
% Bank	11.31%	6.16%	6.16%	6.16%	6.16%	6.16%

TABLE 5: POTENTIAL TARGETS USING AGGREGATED HMDA DATA (2012-2016)

	100/0		**100/2**		**0/20**	
	1-Sided	*2-Sided*	*1-Sided*	*2-Sided*	*1-Sided*	*2-Sided*
HMDA Data						
Total Targets	9,130	8,218	8,896	8,028	7,055	6,477
% MSA-Bank	36.72%	33.05%	37.00%	33.39%	40.18%	36.89%
Unique Target	2,255	2,135	2,232	2,117	2,015	1,926
% Bank	58.15%	55.05%	57.91%	54.93%	58.52%	55.94%
Randomly-Generated Data						
Total Targets	1,368	656	1,257	667	988	528
% MSA-Bank	5.50%	2.64%	5.23%	2.77%	5.63%	3.01%
Unique Target	583	330	525	324	424	246
% Bank	15.03%	8.51%	13.62%	8.41%	12.31%	7.14%
Ratios						
MSA-Bank	6.67	12.53	7.08	12.04	7.14	12.27
Bank	3.87	6.47	4.25	6.53	4.75	7.83

TABLE 6: ROBUSTNESS CHECKS USING AGGREGATED HMDA DATA (2012-2016)

	Originations	Peer Select	No HUD/NCUA	Maj-Black	Conventional
HMDA Data					
Total Targets	5,340	6,438	3,124	4,060	5,579
Total MSA-Bank	17,559	17,468	8,865	11,839	15,633
% MSA-Bank	30.41%	36.86%	35.24%	34.29%	35.69%
Unique Target	1,719	1,922	1,066	1,473	1,796
Total Bank	3,443	3,439	1,895	2,866	3,438
% Bank	49.93%	55.89%	56.25%	51.40%	52.24%
Randomly-Generated Data					
Total Targets	455	501	207	293	450
Total MSA-Bank	17,559	17,468	8,865	11,839	15,633
% MSA-Bank	2.59%	2.87%	2.34%	2.47%	2.88%
Unique Target	243	244	97	159	226
Total Bank	3,443	3,439	1,895	2,866	3,438
% Bank	7.06%	7.10%	5.12%	5.55%	6.57%
Ratios					
MSA-Bank	11.74	12.85	15.09	13.86	12.40
Bank	7.07	7.88	10.99	9.26	7.95

	Middle/High-Income Tracts
HMDA Data	
Total Targets	4,259
Total MSA-Bank	12,873
% MSA-Bank	33.08%
Unique Target	1,489
Total Bank	2,697
% Bank	55.21%
Randomly-Generated Data	
Total Targets	387
Total MSA-Bank	12,873
% MSA-Bank	3.01%
Unique Target	209
Total Bank	2,697
% Bank	7.75%
Ratios	
MSA-Bank	11.01
Bank	7.12