APPENDIX I.
Description of Data Sources for Marketplace Analyses

To perform the marketplace analyses presented in Appendices E through H, Keen Independent used data from the following secondary data sources:

- Integrated Public Use Microdata Series (IPUMS) from the 1990 Decennial Census;

- Integrated Public Use Microdata Series (IPUMS) from the 2000 Decennial Census;

- Integrated Public Use Microdata Series (IPUMS) data from the 2008-2012 (five-year) American Community Survey (ACS);

- Federal Reserve Board's 1998 Survey of Small Business Finances (SSBF);

- Federal Reserve Board's 2003 Survey of Small Business Finances (SSBF);

- 2007 Survey of Business Owners (SBO) conducted by the U.S. Census Bureau;

- 2006 Home Mortgage Disclosure Act (HMDA) data provided by the Federal Financial Institutions Examination Council (FFIEC);

- 2009 Home Mortgage Disclosure Act (HMDA) data provided by the Federal Financial Institutions Examination Council (FFIEC); and

- 2012 Home Mortgage Disclosure Act (HMDA) data provided by the Federal Financial Institutions Examination Council (FFIEC).

The following sections provide further detail on each data source, including how the study team used it in its quantitative marketplace analyses.

A. Integrated Public Use Microdata Series (IPUMS) Data

The Minnesota Population Center is home to the Integrated Public Use Microdata Series (IPUMS), the largest repository of national and international Census microdata for social and economic research. Researchers may access the IPUMS program and retrieve customized, accurate datasets.[1] The IPUMS-USA data consist of more than 50 samples of the American population. These samples are drawn from censuses (1850 to 2000) and from the ACS (2000-2012).

IPUMS data offer several features ideal for the analyses reported in this study, including historical cross-sectional data, stratified national and state-level samples, and large sample sizes that enable analysis with a high level of statistical confidence, even for subsets of the population

---

[1] Steven Ruggles, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. *Integrated Public Use Microdata Series: Version 5.0* [Machine-readable database]. Minneapolis: University of Minnesota, 2011.

(e.g., racial/ethnic and occupational groups). Because the design of these surveys has changed over time, they have a wide range of record layouts and coding schemes. The IPUMS data files are specifically formulated to standardize the U.S. Census Bureau Public Use Microdata Sample (PUMS) data from year to year. Variables that cannot be compared across years are removed from the dataset. In multiyear files, IPUMS inflates dollar values to the most recent year in the sample. IPUMS also provides some additional geographic and family interrelationship variables. Most importantly, IPUMS provides strata and cluster variables for survey samples prior to 2005, as well as replicate weights for survey samples since 2005, to account for the complexity of the sample design in the measurement of standard errors.

The study team obtained selected Decennial Census and ACS IPUMS data from the University of Minnesota Population Center.

Focusing on the construction and engineering industries, the Keen Independent study team used IPUMS data to analyze workers and households in Dane County by examining:

- Demographic characteristics;

- Measures of financial resources;

- Educational attainment; and

- Self-employment (business ownership).

For the analyses contained in this report, the study team used the 1990 and 2000 Census 5 percent samples and 2008-2012 ACS samples.

2000 Census data. The 2000 U.S. Census sample contains 14,081,466 individual observations, weighted to represent 281,421,906 people. The Wisconsin sub-sample contains 272,879 observations weighted to represent 5,357,182 people and the sample for Dane County includes 15,683 observations weighted to represent 429,839 people.

Categorizing individual race/ethnicity. To define race/ethnicity for the 2000 Census dataset, the study team used the IPUMS race/ethnicity variables — RACED and HISPAN — to categorize individuals into one of seven groups:

- Non-Hispanic white;

- Hispanic American;

- African American;

- Asian-Pacific American;

- Subcontinent Asian American;

- Native American; and

- Other minority (unspecified).

An individual was considered "non-Hispanic white" if they did not report Hispanic ethnicity and indicated being white only — not in combination with any other race group. All self-identified Hispanics (based on the HISPAN variable) were considered Hispanic American, regardless of any other race or ethnicity identification. For the five other racial groups, an individual's race/ethnicity was categorized by the first (or only) race group identified in each possible race-type combination. The study team used a rank ordering methodology similar to that used in the 2000 Census data dictionary. An individual who identified multiple races was placed in the reported race category with the highest ranking in the study team's ordering. African American is first, followed by Native American, Asian-Pacific American, and then Subcontinent Asian American. For example, if an individual identified himself or herself as "Korean," that person was placed in the Asian-Pacific American category. If the individual identified himself or herself as "Korean" in combination with "Black," the individual was considered African American.

■ The Asian-Pacific American category included the following race/ethnicity groups: Bhutanese, Burmese, Cambodian, Chamorro, Chinese, Filipino, Guamanian, Hmong, Indonesian, Japanese, Korean, Laotian, Malaysian, Mongolian, Nepalese, Okinawan, Samoan, Tahitian, Taiwanese, Thai, Tongan, and Vietnamese. This category also included other Polynesian, Melanesian, and Micronesian races, as well as individuals identified as Pacific Islanders.

■ The Subcontinent Asian American category included these race groups: Asian Indian (Hindu), Bangladeshi, Pakistani, and Sri Lankan. Individuals who identified themselves as "Asian," but were not clearly categorized as Subcontinent Asian were placed in the Asian-Pacific American group.

■ American Indian, Alaska Native, Native Hawaiian and Latin American Indian groups were considered Native American.

■ If an individual was identified with any of the above groups and an "other race" group, the individual was categorized into the known category. Individuals identified as "other race" or "white and other race" were categorized as "other minority."

For some analyses — those in which sample sizes were small — the study team combined minority groups.

Business ownership. Keen Independent used the Census "labor force status" variable (LABFORCE) and the detailed "class of worker" variable (CLASSWKD) to determine self-employment.[2] Individuals were classified into the following categories.

■ Self-employed for a non-incorporated business;

■ Self-employed for an incorporated business;

---

[2] The labor force consists of the civilian labor force (employed and unemployed) as well as active duty members of the U.S. Armed Forces. Civilians 16 years and older who are not classified in the labor force include students, homemakers, retired workers, seasonal workers interviewed in an off season who were not seeking work, persons doing incidental unpaid family work of less than 15 hours and the institutionalized population (see http://www.census.gov/acs/www/Downloads/data_documentation/SubjectDefinitions/2010_ACSSubjectDefinitions.pdf for more information).

- Wage or salary employee for a private firm;

- Wage or salary employee for a non-profit organization;

- Employee of the Federal government;

- Employee of a State government;

- Employee of a local government; or

- Unpaid family worker.

The study team counted individuals who reported being self-employed — either for an incorporated or a non-incorporated business — as business owners.[3]

Study industries. The marketplace analyses focus on two study industries: construction and engineering-related services. Keen Independent used the IND variable to identify individuals as working in one or the other industry. The variable reports the industry in which a person performed an occupation and includes several hundred industry and subindustry categories. Figure I-1 identifies the IND codes used to define each study area for the 2000 Census and 2008-2012 ACS analyses.

Figure I-1.
2000 Census and 2008-2012 ACS industry codes used for construction

Study industry	2000 Census/ 2008-2012 ACS IND codes	Description
Construction	77/770	Construction industry

Source: Keen Independent Research from the IPUMS program: http://usa.ipums.org/usa/.

---

[3] For the analysis of business ownership, the study team excluded active duty members of the U.S. Armed Forces and all other wage/salary workers.

Industry occupations. The study team also examined workers by occupation within the construction industry using the PUMS variable OCC. Figure I-2 summarizes the 2000 Census and 2008-2012 ACS OCC codes used in the study team's analyses.

Figure I-2.
2000 Census and 2008-2012 ACS occupation codes used to examine workers in construction

Census 2000/ 2008-2012 ACS occupational title and code	Job description
Construction managers 22/220	Plan, direct, coordinate, or budget, usually through subordinate supervisory personnel, activities concerned with the construction and maintenance of structures, facilities, and systems. Participate in the conceptual development of a construction project and oversee its organization, scheduling, and implementation. Include specialized construction fields, such as carpentry or plumbing. Include general superintendents, project managers, and constructors who manage, coordinate, and supervise the construction process.
First-line supervisors/managers of construction trades and extraction workers 620/6200	Directly supervise and coordinate the activities of construction or extraction workers.
Brickmasons, blockmasons and stonemasons 622/6220	Lay and bind building materials, such as brick, structural tile, concrete block, cinder block, glass block, and terra-cotta block, construct or repair walls, partitions, arches, sewers, and other structures. Build stone structures, such as piers, walls, and abutments and lay walks, curbstones, or special types of masonry for vats, tanks, and floors.
Carpenters 623/6230	Construct, erect, install, or repair structures and fixtures made of wood, such as concrete forms, building frameworks including partitions, joists, studding, rafters, wood stairways, window and door frames, and hardwood floors.
Carpet, floor, and tile installers and finishers 624/6240	Apply shock-absorbing, sound-deadening, or decorative coverings to floors. Lay carpet on floors and install padding and trim flooring materials. Scrape and sand wooden floors to smooth surfaces, apply coats of finish. Apply hard tile, marble, wood tile, walls, floors, ceilings, and roof decks.

Figure I-2 (continued).
2000 Census and 2008-2012 ACS occupation codes used to examine workers in construction

Census 2000 and 2008-2012 ACS occupational title and code	Job description
Cement masons, concrete finishers and terrazzo workers 625/6250	Smooth and finish surfaces of poured concrete, such as floors, walks, sidewalks, or curbs using a variety of hand and power tools. Align forms for sidewalks, curbs or gutters; patch voids; use saws to cut expansion joints. Terrazzo workers apply a mixture of cement, sand, pigment or marble chips to floors, stairways, and cabinet fixtures.
Construction laborers 626/6260	Perform tasks involving physical labor at building, highway, and heavy construction projects, tunnel and shaft excavations, and demolition sites. May operate hand and power tools of all types: air hammers, earth tampers, cement mixers, small mechanical hoists, surveying and measuring equipment, and a variety of other equipment and instruments. May clean and prepare sites, dig trenches, set braces to support the sides of excavations, erect scaffolding, clean up rubble and debris, and remove asbestos, lead, and other hazardous waste materials. May assist other craft workers. Exclude construction laborers who primarily assist a particular craft worker, and classify them under "Helpers, Construction Trades."
Paving, surfacing and tamping equipment operators 630/6300	Operate equipment used for applying concrete, asphalt, or other materials to road beds, parking lots, or airport runways and taxiways, or equipment used for tamping gravel, dirt, or other materials. Include concrete and asphalt paving machine operators, form tampers, tamping machine operators, and stone spreader operators.
Miscellaneous construction equipment operators, including pile-driver operators 632/6320	Operate one or several types of power construction equipment, such as motor graders, bulldozers, scrapers, compressors, pumps, derricks, shovels, tractors, or front-end loaders to excavate, move, and grade earth, erect structures, or pour concrete or other hard surface pavement. Operate pile drivers mounted on skids, barges, crawler treads, or locomotive cranes to drive pilings for retaining walls, bulkheads, and foundations of structures, such as buildings, bridges, and piers.
Drywall installers, ceiling tile installers and tapers 633/6330	Apply plasterboard or other wallboard to ceilings or interior walls of buildings, mount acoustical tiles or blocks, strips, or sheets of shock-absorbing materials to ceilings and walls of buildings to reduce or reflect sound.
Electricians 635/6350,6355	Install, maintain, and repair electrical wiring, equipment, and fixtures. Ensure that work is in accordance with relevant codes. May install or service street lights, intercom systems, or electrical control systems. Exclude "Security and Fire Alarm Systems Installers." The 2000 category includes electrician apprentices.

Figure I-2 (continued).
2000 Census and 2008-2012 ACS occupation codes used to examine workers in construction

Census 2000 and 2008-2012 ACS occupational title and code	Job description
Glaziers 636/6360	Install glass in windows, skylights, store fronts, display cases, building fronts, interior walls, ceilings, and tabletops.
Painters, construction and maintenance 642/6420	Paint walls, equipment, buildings, bridges, and other structural surfaces using brushes, rollers, and spray guns. Remove old paint to prepare surfaces prior to painting and mix colors or oils to obtain desired color or consistency.
Pipelayers, plumbers, pipefitters and steamfitters 644/6440	Lay pipe for storm or sanitation sewers, drains, and water mains. Perform any combination of the following tasks: grade trenches or culverts, position pipe, or seal joints. Excludes "Welders, Cutters, Solderers, and Brazers." Assemble, install, alter, and repair pipelines or pipe systems that carry water, steam, air, or other liquids or gases. May install heating and cooling equipment and mechanical control systems. Includes sprinklerfitters.
Plasterers and stucco masons 646/6460	Apply interior or exterior plaster, cement, stucco, or similar materials and set ornamental plaster.
Roofers 651/6510,6515	Cover roofs of structures with shingles, slate, asphalt, aluminum, and wood. Spray roofs, sidings, and walls with material to bind, seal, insulate, or soundproof sections of structures.
Iron and steel workers, including reinforcing iron and rebar workers 653/6530	*Iron and steel workers* raise, place, and unite iron or steel girders, columns, and other structural members to form completed structures or structural frameworks. May erect metal storage tanks and assemble prefabricated metal buildings. *Reinforcing iron and rebar workers* position and secure steel bars or mesh in concrete forms in order to reinforce concrete. Use a variety of fasteners, rod-bending machines, blowtorches, and hand tools. Include rod busters.
Helpers, construction trades 660/6600	All construction trades helpers not listed separately.

Figure I-2 (continued).
2000 Census and 2008-2012 ACS occupation codes used to examine workers in construction

Census 2000 and 2008-2012 ACS occupational title and code	Job description
Driver/sales workers and truck drivers 913/9130	*Driver/sales workers* drive trucks or other vehicles over established routes or within an established territory and sell goods, such as food products, including restaurant take-out items, or pick up and deliver items, such as laundry. May also take orders and collect payments. Include newspaper delivery drivers. *Truck drivers (heavy)* drive a tractor-trailer combination or a truck with a capacity of at least 26,000 GVW, to transport and deliver goods, livestock, or materials in liquid, loose, or packaged form. May be required to unload truck. May require use of automated routing equipment. Requires commercial drivers' license. *Truck drivers (light)* drive a truck or van with a capacity of under 26,000 GVW, primarily to deliver or pick up merchandise or to deliver packages within a specified area. May require use of automatic routing or location software. May load and unload truck. Exclude "Couriers and Messengers."
Crane and tower operators 951/9510	Operate mechanical boom and cable or tower and cable equipment to lift and move materials, machines, or products in many directions. Exclude "Excavating and Loading Machine and Dragline Operators."
Dredge, excavating and loading machine operators 952/9520	*Dredge operators* operate dredge to remove sand, gravel, or other materials from lakes, rivers, or streams; and to excavate and maintain navigable channels in waterways. *Excavating and loading machine and dragline operators* Operate or tend machinery equipped with scoops, shovels, or buckets, to excavate and load loose materials. *Loading machine operators, underground mining*, Operate underground loading machine to load coal, ore, or rock into shuttle or mine car or onto conveyors. Loading equipment may include power shovels, hoisting engines equipped with cable-drawn scraper or scoop, or machines equipped with gathering arms and conveyor.

Source: 2000 Census occupational titles and codes at http://usa.ipums.org/usa/volii/00occup.shtml. 2008-2012. ACS occupational titles and codes at https://usa.ipums.org/usa/volii/c2ssoccup.shtml.

Education variables. Keen Independent used the variable indicating respondents' highest level of educational attainment (EDUCD) to classify individuals into six categories:

■   Less than high school;

■   High school diploma or equivalent;

■   Some college but no degree;

■   Associate's degree;

■   Bachelor's degree; and

■   Advanced degree.

Definition of workers. The universe for the class of worker, industry, and occupation variables includes workers 16 years of age or older who are "gainfully employed" and those who are unemployed but seeking work. "Gainfully employed" means that the worker reported an occupation as defined by the Census code OCC.

1990 Census data. The study team compared 2000 Census data with data for the 1990 Census to analyze changes in worker demographics, educational attainment and business ownership over time. The 1990 Census five percent sample includes 12,501,046 observations weighted to represent 248,107,628 people. At the state level, the sample includes 245,498 observations in Wisconsin weighted to represent 4,881,441 people. For Dane County, the sample includes 15,692 observations weighted to represent 364,684 individuals.

2008-2012 American Community Survey (ACS) data. The study team also examined 2008-2012 ACS data from IPUMS. The U.S. Census Bureau conducts the ACS which uses monthly samples to produce annually updated data for the same small areas as the 2000 Census long-form.[4] Since 2005, the ACS has expanded to a roughly 1 percent sample of the population, based on a random sample of housing units in every county in the United States (including District of Columbia and Puerto Rico). The 2008-2012 ACS estimates represent the average characteristics over the five-year period of time.

There were 3,063,887 observations included in the U.S. sample, weighted to represent 309,376,285 people.[5] The Wisconsin sub-sample contains 293,331 observations weighted to represent 5,687,219 people; the Dane County sample includes 18,289 observations weighted to represent 490,233 people.

Changes in race/ethnicity categories between censuses. The 2000 Census 5 percent sample and the 2008-2012 ACS IPUMS data use essentially the same categories for the detailed race variable (RACED). However, in some cases, the numerical code assignment is different and the study team accounted for those differences. Categories for the Hispanic variable (HISPAN) remained consistent between the two datasets. Comparison to the 1990 data was not as straightforward. In 2000, the Census Bureau introduced categories representing a combination of race types to allow individuals to select multiple races when responding to the questionnaire. For example, an individual who is primarily white with Native American ancestry could choose the "white and American Indian/Alaska Native" race group in 2000. However, if the same individual received the 1990 Census questionnaire, he/she would need to choose a single race group – either "white" or "American Indian/Alaska Native." Such a choice would ultimately depend on unknowable factors including how strongly the individual identifies with her Native American heritage. The variability introduced by allowing multiple race selection complicates direct comparisons between race data from the 2000 Census and previous censuses. Even so, 98 percent of survey respondents in 2000 indicated a single race.[6]

---

[4] U.S. Census Bureau. *Design and Methodology: American Community Survey.* Washington D.C.: U.S. Government Printing 2009. Available at http://www.census.gov/acs/www/SBasics/desgn_meth.htm.

[5] Analyses of the 2008-2012 American Community Survey PUMS data at the national level were based on the one percent sample; at the state and county level, analyses were based on the five percent sample.

[6] Grieco, Elizabeth M. and Rachel C. Cassidy. "Overview of Race and Hispanic Origin." Census 2000 Brief, March 2001.

B. Survey of Small Business Finances (SSBF)

The study team used the SSBF to analyze the availability and characteristics of small business loans. The Federal Reserve Board conducted the SSBG every five years through 2003, but stopped after that year.

The SSBF collects financial data from non-governmental for-profit firms with fewer than 500 employees. The survey uses a nationally representative sample, structured to allow for analysis of specific geographic regions, industry sectors, and racial and gender groups. The SSBF is unique as it provides detailed data on both firm and owner financial characteristics. For the purposes of this report, Keen Independent used the surveys from 1998 and 2003, which are available at the Federal Reserve Board website.[7]

Categorizing owner race/ethnicity and gender. In the SSBF, businesses were able to give responses on owner characteristics for up to three different owners. The data also included a fourth variable that is a weighted average of other answers provided for each question. In order to define race/ethnicity and gender variables, the study team used the final weighted average for variables on owner characteristics. Definition of race and ethnic groups in the SSBF are slightly different than the classifications used in the Decennial Censuses and 2008-2012 ACS.

The SSBF classified race and ethnicity of businesses according to the following five groups:

- Non-Hispanic white;

- Hispanic American;

- African American;

- Asian American;

- Native American; and

- Other (unspecified).

A business was considered Hispanic American-owned if more than 50 percent of the business was owned by Hispanic Americans, regardless of race. All businesses that reported 50 percent or less Hispanic American ownership were included in the racial group that owned more than half of the company. No firms reported the race/ethnicity of their owners as "other."

Similar to race, firms were classified as female-owned if more than 50 percent of the firm was owned by women. Firms owned half by women and half by men were classified as male-owned.

---

[7] The Federal Reserve Board. *Survey of Small Business Finances, 2003*. Available online at http://www.federalreserve.gov/pubs/.

Defining selected industry sectors. In the 2003 SSBF, each business was classified according to Standard Industrial Classification (SIC) code and placed into one of seven industry categories:

- Construction;

- Mining;

- Transportation, communications, and utilities;

- Finance, insurance, and real estate;

- Trade;

- Engineering; or

- Services (excluding engineering).

In 1998, the SSBF allowed an additional industry category of Agriculture, forestry and fishing.

Region variables. The SSBF divides the United States into nine Census Divisions. Along with, Illinois, Indiana, Michigan and Ohio, Wisconsin is in the East North Central Census Division.

Loan denial variables. In the survey, firm owners were asked if they have applied for a loan in the last three years and whether loan applications were always approved, always denied, or sometimes approved and sometimes denied. For the purposes of this study, only firms that were always denied were considered when analyzing loan denial.

Data reporting. Due to missing responses to survey questions in SSBF datasets, data were imputed to fill in missing values. The missing values in the 2003 dataset were imputed using a different method than in previous SSBF studies. In the 1998 survey data, the number of observations in the dataset matches the number of firms surveyed. However, the 2003 data includes five implicates, each with imputed values that have been filled in using a randomized regression model.[8] Thus, there are 21,200 observations in the 2003 data, five for each of the 4,240 firms surveyed. Across the five implicates, all non-missing values are identical, whereas imputed values may differ.

As discussed in a recent paper about the 2003 imputations by the Finance and Economics Discussion Series, missing survey values can lead to biased estimates and inaccurate variances and confidence intervals.[9] Those problems can be corrected through the use of multiple implicates. For summary statistics using 2003 SSBF data, Keen Independent utilized all five implicates and included observations with missing values in the analyses. For the probit regression models presented in

---

[8] For a more detailed explanation of imputation methods, see the "Technical Codebook" for the *2003 Survey of Small Business Finances.*

[9] Lieu N. Hazelwood, Traci L. Mach and John D. Wolken. *Alternative Methods of Unit Nonresponse Weight Adjustments: An Application from the 2003 Survey of Small Businesses.* Finance and Economics Discussion Series Divisions of Research and Statistics and Monetary Affairs, Federal Reserve Board. Washington, D.C., 2007. http://www.federalreserve.gov/pubs/feds/2007/200710/200710pap.pdf.

Appendix G, the study team used the first implicate and did not include observations with imputed values for the dependent variables.

C. Survey of Business Owners (SBO)

Keen Independent used data from the 2007 SBO to analyze mean annual firm receipts. The SBO is conducted every five years by the U.S. Census Bureau. Data for the most recent publication of the SBO were collected in 2007. (In early 2015, the Census Bureau was still conducting research for the 2012 SBO.)

Response to the survey is mandatory, which ensures comprehensive economic and demographic information for business and business owners in the U.S. All tax-filing businesses and nonprofits were eligible to be surveyed, including firms with and without paid employees. In 2007, almost 8 million firms were surveyed. The study team examined SBO data relating to the number of firms, number of firms with paid employees, and total receipts. That information is available by geographic location, industry, gender and race/ethnicity.

The SBO uses the 2002 North American Industry Classification System (NAICS) to classify industries. The study team analyzed data for firms in all industries and for firms in selected industries that corresponded closely to construction and engineering-related services.

To categorize the business ownership of firms reported in the SBO, the Census Bureau uses standard definitions for women-owned and minority-owned businesses. A business is defined as female-owned if more than half of the ownership and control is by women. Firms with joint male-/female-ownership were tabulated as an independent gender category. A business is defined as minority-owned if more than half of the ownership and control is by African Americans, Asian Americans, Hispanic Americans, Native Americans, or by another minority group. Respondents had the option of selecting one or more racial groups when reporting business ownership.

The study team reported business receipts for the following race/ethnicity and gender groups:

- African Americans;

- Asian Americans;

- Hispanic Americans;

- Native Americans;

- Non-Hispanic whites;

- Men; and

- Women.

D. Home Mortgage Disclosure Act (HMDA) Data

Keen Independent analyzed mortgage lending in Wisconsin and nationwide using HMDA data that the Federal Financial Institutions Examination Council (FFIEC) provides. HMDA data provide information on mortgage loan applications that financial institutions, savings banks, credit unions and some mortgage companies receive. Those data include information about the location, dollar amount and types of loans made, as well as race/ethnicity, income and credit characteristics of loan applicants. Data are available for home purchase, home improvement, and refinance loans.

Financial institutions were required to report 2012 HMDA data if they had assets of more than $41 million ($35 million for 2007), had a branch office in a metropolitan area and originated at least one home purchase or refinance loan in the reporting calendar year. Mortgage companies were required to report HMDA if they are for-profit institutions, had home purchase loan originations exceeding 10 percent of all loan obligations in the past year, were located in an MSA (or originated five or more home purchase loans in an MSA), and either had more than $10 million in assets or made at least 100 home purchase or refinance loans in the calendar year.

The study team used those data to examine loan denial rates and subprime lending rates for different racial and ethnic groups in 2006, 2009 and 2012. Note that the HMDA data represent the entirety of home mortgage loan applications reported by participating financial institutions in each year examined. Those data are not a sample. Appendix G provides a detailed explanation of the methodology that the study team used for measuring loan denial and subprime lending rates.