

Other Things Being Equal: A Paired Testing Study of Discrimination in Mortgage Lending

Margery Austin Turner, Erin Godfrey, Stephen L. Ross, and Robin R. Smith
3/7/2003

When first-time homebuyers begin shopping for a house, they need to learn about mortgages for which they can qualify and about house prices they can afford. This information can be obtained from a variety of sources, including mortgage lending institutions, real estate agents, and mortgage brokers. But if potential homebuyers cannot obtain full and fair access to information about mortgage financing, they may give up on their pursuit of homeownership, their housing search may be restricted, or they may be unable to negotiate the most favorable loan terms. Thus, pre-application inquiries about mortgage financing options represent a critical phase in the homebuying process.

This article summarizes a recently completed study, funded by the U.S. Department of Housing and Urban Development (HUD) and conducted by the Urban Institute, of discrimination against black and Hispanic homebuyers when they visit mortgage lending institutions in two major metropolitan markets to make pre-application inquiries. It represents the first application of paired testing to rigorously measure discrimination in the mortgage lending process. The paired tests disclosed significant levels of adverse treatment on the basis of race and ethnicity, with African Americans and Hispanics receiving less information and assistance than comparable whites, even at this very early stage in the application process.

Discrimination in the Mortgage Lending Process

More than three decades after the passage of the Fair Housing Act (Title VIII of the 1968 Civil Rights Act), African American and Hispanic homebuyers still do not enjoy equal access to home ownership. Widespread evidence indicates that minority homebuyers are less likely than whites to obtain mortgage loans and, if they are successful, receive less favorable loan amounts and terms. However, considerable disagreement persists about the extent to which discrimination is the cause of these unequal outcomes, or whether they result primarily from unequal qualifications and creditworthiness.¹ A recent review of existing social science evidence concluded that minority homebuyers in the United States do face discrimination from mortgage lending

¹ See, for example, Munnell et al (1986), Ross and Yinger (2002), Goering and Wienk (1996), Carr and Megbolugbe (1993); Ladd (1998); Yezer (1995).

institutions.² But serious gaps remain in what we know and more rigorous information is needed about the forms it takes and the stages at which it occurs in order to refine and target enforcement strategies, to enable lending institutions to monitor their own performance, and to design remedies to reduce discrimination in home mortgage lending.

Paired testing provides a uniquely powerful tool for investigating both the incidence and the forms of adverse treatment based on race or ethnicity. In a paired test, two individuals—one white and one minority—pose as homebuyers and inquire about the availability and terms for home mortgage loans.³ Because the two members of a tester team present themselves as equally qualified borrowers in every respect except their race or ethnicity, systematic differences in the treatment they receive provide direct evidence of disparate treatment.⁴

The paired testing methodology has been used widely to detect and measure adverse treatment by rental and sales agents, but only a few, relatively small-scale investigative studies—primarily by the National Fair Housing Alliance (NFHA)—have been applied to mortgage lending. During the early 1990s, NFHA conducted tests in seven cities (Atlanta, Chicago, Dallas, Denver, Detroit, Oakland, and Richmond). Testers posed as first-time homebuyers and refinancers inquiring about financing terms and conditions at the pre-application stage. A reanalysis of these testing data by the Urban Institute concluded that differential treatment occurred at significant levels in at least some cities. Minorities were less likely to receive information about loan products, received less time and information from loan officers, and were quoted higher interest rates in most of the cities where tests were conducted.

Paired Testing Methodology

Building upon the experience and findings of previous testing efforts, the Department of Housing and Urban Development (HUD) contracted with the Urban Institute to further refine and apply methods for conducting paired testing at the pre-

² Turner and Skidmore (1999).

³ For more information on paired testing and its role in both measurement and enforcement, see Fix and Turner (1999).

⁴ It is important to note that discrimination in mortgage lending may involve either disparate treatment (unequal treatment of equals) or disparate impact (policies that, though equally applied, systematically disadvantage minorities). The paired testing methodology captures disparate treatment, but not disparate impact discrimination. However, it is possible that disparate treatment discrimination at the pre-application stage might cause minority borrowers to apply for different loan amounts or products, thereby contributing to a pattern of disparate treatment.

application stage of the mortgage lending process.⁵ Because of the complexity of the mortgage application process and the challenges it presents for paired testing, we divided this effort into two basic stages—a *pre-test* stage and a *pilot* stage.

The pre-test stage was used to experiment with a fairly wide variety of paired testing scenarios, and to assess the feasibility of testing several different sources of mortgage financing information. A total of 78 tests were conducted in two markets—Orange County, California and New Orleans, Louisiana – targeting six different information sources, reflecting the range of possible sources of mortgage information used by homebuyers, including conventional mortgage lenders, sub-prime mortgage lenders, mortgage brokers, real estate agents, new home sales offices, and mobile home dealers. The pre-tests were also used to experiment with different testing scenarios. Some testers indicated that they already had a particular house in mind while others were trying to find out how much they could potentially afford. Some testers posed as well-qualified borrowers, while others were assigned marginal qualifications. Some testers visited lenders as couples, posing as husband and wife inquiring together, while others conducted one-person visits, but indicated that they were married.

Based on the pre-test experience, the project's pilot phase conducted approximately 250 paired tests of a representative sample of mortgage lending institutions in Los Angeles, California and Chicago, Illinois.⁶ These tests followed a single, standardized set of protocols in order to yield statistically rigorous measures of adverse treatment against African Americans and Hispanics in the two metropolitan housing markets. Specifically, testers posing as first-time homebuyers visited mortgage lending institutions to make a general, uninformed request for information about how much house they could afford and what loan products might be available to them.

All of the testers were assigned financial profiles that qualified them for products targeted to borrowers with A- credit in their respective housing markets. Specifically, their assigned income and asset levels were sufficient to purchase a median-priced house in their metropolitan area, assuming a 30-year conventional fixed-rate loan at 8 percent interest with a 5 percent downpayment. The financial profiles made the testers moderately asset-constrained, meaning that their available downpayment resources limited the loan amount for which they could qualify. In addition, all of the testers were assigned one or two minor credit blemishes, usually a late payment of some kind. The

⁵ In principle, paired testing might be used to measure discrimination at the application and underwriting stage of the mortgage lending process as well. However, most experts believe that federal laws prohibiting the submission of false credit applications make it impossible to extend the paired testing methodology into the application stage. For further discussion of this issue, see Turner and Skidmore (1999).

⁶ Lessons from the pre-test phase were also incorporated into a package of tools for enforcement testing. See Freiberg and Herbig(2002).

two members of each tester pair were given virtually identical financial and household characteristics, with the minority partner always slightly better qualified than the white.

In both Chicago and Los Angeles, testers visited a representative sample mortgage lending institutions in the metropolitan area that reported under the Home Mortgage Disclosure Act (HMDA),⁷ accepted at least 90 mortgage loan applications per year, and had offices in the region that a first-time homebuyer could realistically find and visit. In order to draw a market representative sample, lenders were selected with replacement, with the probability of selection based on loan volume. Lending institutions with very large application volumes not only had a high probability of selection, but were likely to appear in the sample more than once. This sampling strategy allows us to draw conclusions about the incidence of differential treatment by large lending institutions in Chicago and Los Angeles that are directly accessible to first-time homebuyers.⁸

In the Los Angeles metropolitan area, 35 lenders were selected for black/white testing and 34 were selected for Hispanic/Anglo testing see (Exhibit 1). In each categories, the lenders selected accounted for half of the application activity captured in HMDA data. In Chicago, the lenders selected for black/white testing covered 47 percent of the application activity, while those tested for treatment of Hispanics and Anglos covered 51 percent.

Exhibit 1: Final Sample of Lending Institutions

Lenders Tested	# of Institutions		Application Volume		% of HMDA Activity	
	<u>LA</u>	<u>Chicago</u>	<u>LA</u>	<u>Chicago</u>	<u>LA</u>	<u>Chicago</u>
Black/White	35	49	81,031	78,655	50	47*
Hispanic/Anglo	34	51	80,447	85,214	50	51*

* During the course of testing, one institution in our sample merged with another large lender. HMDA application volume for this lender is difficult to determine and is not represented in these totals.

While the sample was drawn based on lending institution, each test was conducted by visiting an individual branch office, which was randomly selected from all of the institution's local offices. To select a branch office for a test, a list of the target lender's local branches was compiled. Urban institute staff made calls to local offices to verify addresses and determine which branches potential borrowers could visit to receive information on mortgage loans.⁹ Once the list of local branches was reduced to those

⁷ HMDA requires all independent mortgage companies and mortgage lenders owned by depository institutions that make at least 100 home purchase and/or refinancing loans in a given year to report on the demographic and locational characteristics of all applications and loans.

⁸ Levels and patterns of discrimination may be different for smaller lending institutions or when inquiries are made by telephone or internet rather than in person.

⁹ Information on the volume of activity conducted by each branch was not available.

providing mortgage information, the branch to be tested was selected randomly from the list.

The testing protocols that testers were followed when conducting a test can be summarized in five basic steps.

- *Step #1 - Obtain an Appointment.* All testers called to arrange “in person” visits with lenders. Testers were provided detailed instructions on how to complete such calls and how to avoid protracted conversations with loan officers over the phone.
- *Step #2 - Make the Initial Request.* When testers arrived for their appointments, their first step was to very clearly state (up to three times, if necessary) that the purpose of their visit was to obtain help in figuring out a price range of housing that they might be able to afford and an estimated loan amount for which they might qualify.
- *Step #3 - Exchange Personal/Financial Information.* Testers were trained to be forthcoming and provide income, debts, assets, credit information and other personal and financial characteristics when requested by a lender. Testers were instructed to be precise when providing their financial information and refer to their “cheat sheets” if necessary. Under no circumstances, however, did testers provide a social security number or date of birth or authorize a credit check.
- *Step #4 - Record Information on Financing Options Recommended.* Testers were required to take notes and record information provided by the lender such as suggested home price range, an estimated loan amount, and details about any financing options recommended.
- *Step #5 - End the Visit.* Testers were instructed to thank the lender for any assistance and allow the lender to suggest any follow-up contact.

Following every test visit, testers were completed a *Test Report Form* and a *Test Narrative Form*. The *Test Report Form* recorded their responses to specific questions about the test experience and the information that was provided by the lender. The *Test Narrative Form* provided a detailed, chronological, account of the test experience in the tester’s own words. Additional narrative forms were completed by testers following any phone contact by a lender. Testers were instructed to complete all forms as soon as possible following contact with a lender. Testers completed the forms based on their recollection of what occurred during the test, and on their review of notes taken and materials obtained during the test.¹⁰

¹⁰ Because of the complexity of lender testing, the detailed narratives played a particularly important role in quality control. Specifically, test coordinators, quality control supervisors, and Urban

Data from paired testing can be used to construct three different types of measures: 1) measures of the *gross incidence* of differential treatment; 2) measures of the *net incidence* of differential treatment; and 3) measures of the *severity* of differential treatment. Each of these measures is briefly explained in turn.

A *gross incidence* measure is defined as the share of all tests in which the minority receives less favorable treatment than his or her white Anglo partner. Gross incidence measures provide very simple and understandable indicators of how often minorities are treated less favorably than equally qualified white Anglos. However, there are also cases in which minority testers receive better treatment than their white Anglo partners. Thus, we constructed measures of the gross incidence of minority-favored treatment as well as measures of the gross incidence of white-favored treatment.

Net incidence measures focus on the difference between these two gross incidence measures by subtracting the gross incidence of minority-favored treatment from the gross incidence of white-favored treatment for a given indicator. A large net incidence measure would suggest that—even though minorities are sometimes favored over white Anglos—unfavorable treatment of minorities is substantially more prevalent. A small net incidence measure, on the other hand, suggests that while lenders do not always provide comparable treatment to similar customers, they are just as likely to treat minority customers favorably as white Anglo customers.

Net incidence measures are often interpreted as estimates of systematic discrimination against minorities. If one assumes that white Anglos rarely experience systematic adverse treatment, then all cases of minority-favored treatment can be interpreted as random differences in treatment, unrelated to race or ethnicity. If this assumption is correct, then by subtracting cases of minority-favored treatment from the cases of white-favored treatment, the net incidence measure removes the element of random error and reflects the true incidence of discrimination against minorities. However, if the assumption is incorrect (and systematic discrimination against white Anglos does sometimes occur), then the net measure may actually understate the incidence of discrimination against minorities.¹¹

In the analysis presented here, gross incidence measures are reported for both white-favored and minority-favored treatment. When these two gross measures are

Institute staff all reviewed the narrative reports to ensure that testers adhered to HTP protocols and accurately completed the more structured, closed-ended *Test Report Form*.

¹¹ For a more extensive discussion of the interpretation of gross and net measures, see Turner et al (2002).

significantly different from one another (the net measure is significantly greater than zero), then we conclude that a systematic pattern of differential treatment based on race or ethnicity has occurred. If, on the other hand, the incidence of minority-favored treatment is essentially the same as the incidence of white-favored treatment, we cannot conclude that these differences are systematically based on race or ethnicity. Because our sample sizes are relatively small, and the data may not be normally distributed, conventional tests of statistical significance may fail to detect differences in treatment that are actually significant. Therefore, we used the Sign Test to determine whether the incidence of white-favored treatment was significantly different than the incidence of minority-favored treatment.¹²

It is important to note that even when no statistical pattern of race-based differential treatment is observed, individual cases of discrimination may have occurred. Specifically, for variables where the gross measures of white-favored and minority-favored treatment is essentially equal, there may in fact be instances of race-based discrimination, even though the overall pattern does not systematically favor one group. Finally, even when treatment in one area appears to favor either the white Anglo or the minority tester, this does not necessarily mean that the entire test favors that tester. A qualitative review of the entire test file might be needed to assess the overall outcome across multiple measures.

Gross and net incidence measures are complemented by measures of the *severity* of differential treatment, which reflect the size or magnitude of differences in treatment between minority testers and their white Anglo partners.¹³ Severity measures can only be constructed for forms of treatment that yield continuous differences, such as dollars of loan amount or number of loan products. These measures do not apply to simpler “yes/no” forms of treatment, such as whether a customer receives any information or whether a customer is told she is qualified for any loan products. Again because of our small sample sizes and the potential that the data are not normally distributed, we employed the more sensitive Wilcoxon Rank Sum Test to measure the statistical significance of these severity measures.¹⁴

Unequal Information and Assistance for Minority Homebuyers

Even at the pre-application stage, inquiries about mortgage products and terms are complicated interactions and differences in treatment can take many forms. Based

¹² See Heckman and Siegelman(1993) and Ramsey and Schafer (1997).

¹³ Note that severity measures are inherently net measures, since they reflect the average of differences between minority and white outcomes. Thus, cases where minorities were favored effectively cancel out cases where whites were favored.

¹⁴ See Ramsey and Schafer (1997)..

upon the testing protocols and procedures described above, we explored six major questions about the information and assistance that lending institutions provided:¹⁵

- 1) Did testers receive the information they requested about loan amounts and house prices they could afford?
- 2) How much were testers told they could afford to borrow and/or buy?
- 3) How many specific products were testers told about?
- 4) How much “coaching” did testers receive to help them qualify for a loan?
- 5) Did testers receive follow-up calls from lenders?
- 6) Were testers encouraged to consider FHA loans as an option?

In each of these areas, the experience of the white Anglo and minority members of a tester pair were compared to determine whether both were treated equally, the white Anglo tester was favored, or the minority tester was favored. We then calculated the *incidence* of differential treatment across tests, to reflect the frequency with which equally qualified partners were treated differently from one another when they visited lending institutions to inquire about mortgage products. In addition, for some forms of treatment (such as maximum loan amount), we calculated the *severity* of differential treatment, by comparing the average amount quoted to white Anglo testers to the average amount quoted to minority testers.

Did testers receive the information they requested? When testers visited mortgage lending institutions, they posed as first-time homebuyers trying to obtain basic information about how much they might be able to borrow and how much house they can realistically afford. Thus, the first group of treatment measures focuses on whether white Anglo and minority testers were equally successful in getting the basic information they requested. Exhibit 2 reports the incidence of differential treatment for four indicators in this category:

- Was a loan amount provided?
- Was a house price provided?
- Were specific financing options discussed?

¹⁵ Testers also recorded the terms and conditions of specific loan product they were offered. Ideally, one would want to compare terms and conditions for comparable loan products offered to both white and minority customers. However, because HTP protocols called for testers to approach lending institutions with a very general request for information, the product-specific information they received was very diverse. Similar products were listed in different order and given different names, and testers did not always receive a complete set of terms and conditions for every product discussed. Therefore, it is not possible to match products and compare terms and conditions. Annex C presents an exploratory analysis that simply compares average terms and conditions across the specific loan products discussed with testers.

- Were financial details exchanged (in other words, did the loan officer provide information about loan amount and/or house price after finding out about the tester's financial circumstances)?

In addition, we constructed a summary score for this group of treatment indicators, reflecting both the amount and quality of information testers received. This summary score encompasses all four of the individual indicators that make up this category. It ranks and scores the indicators to create an overall measure of the level of information received. Scores were assigned to each tester as follows:

Summary Score	Treatment
5	Financial details were exchanged and the tester received a loan amount or house price and specific financing options were discussed
4	Financial details were exchanged and the tester received a loan amount or house price but no specific options were discussed
3	Financial details were not exchanged but the tester received a loan amount or house price and specific financing options were discussed
2	Financial details were not exchanged and the tester received a loan amount or house price but no specific options were discussed
1	The tester did not receive a loan amount or house price

Then, for each test, scores were compared for the two testers to determine whether one was favored overall. The tester with the higher summary score was considered favored in this area. This ranking scheme reflects the view that homebuyers are best served when loan officers provide information about affordable house prices and loan amounts that are based on knowledge of the customer's financial circumstances.

Exhibit 2: Who Received the Information Requested?

Black/white tests

Treatment Indicators	Los Angeles			Chicago		
	% White Favored	% Black Favored	Significance Test	% White Favored	% Black Favored	Significance Test
Loan amount provided	6.8	4.1	0.73	6.3	1.3	0.22
House price provided	6.8	6.8	1.00	5.1	3.8	1.00
Specific options discussed	6.8	5.5	1.00	6.3	0.0	0.06 *
Financial details exchanged	13.7	11.0	0.81	17.7	8.9	0.19
Overall info provided	9.5	8.1	1.00	8.9	5.1	0.55

Hispanic/Anglo Tests

Treatment Indicators	Los Angeles			Chicago		
	% Anglo Favored	% Hispanic Favored	Significance Test	% Anglo Favored	% Hispanic Favored	Significance Test
Loan amount provided	11.4	1.3	0.02**	5.1	3.8	1.00
House price provided	10.1	1.3	0.04**	6.4	6.4	1.00
Specific options discussed	11.5	1.3	0.02**	5.7	8.6	0.75
Financial details exchanged	17.7	10.1	0.29	10.1	13.9	0.65
Overall info provided	12.7	1.3	0.01**	7.5	8.8	1.00

* difference between % Anglo favored and % Hispanic favored is statistically significant at a 90% confidence level or higher

** difference between % Anglo favored and % Hispanic favored is statistically significant at a 95% confidence level or higher

Testers received the same treatment most of the time on black/white tests in both Los Angeles and Chicago. The incidence of both white-favored and minority-favored treatment was almost always below 10 percent, with slightly higher incidences only in the financial details exchanged category. For most indicators, there were no statistically significant differences between the incidence of white-favored treatment and the incidence of minority-favored treatment. The only statistically significant difference was found on specific financing options. In 6.3 percent of black/white tests in Chicago, the lender discussed specific financing options with the white tester but not the black tester in a pair. Black testers were never favored on this treatment variable.

For Hispanic/Anglo tests, testers also received the same treatment most of the time, but Anglo testers were favored overall in Los Angeles. Results for the Hispanic/Anglo tests were similar to black/white tests in that testers received the same treatment most of the time. However, in Los Angeles, when differences did occur, they were significantly more likely to favor Anglos than Hispanics. Specifically, the incidence of Anglo-favored treatment for loan amount provided, house price

Although a white male tester declined to authorize a credit check, the loan officer pre-qualified him for a maximum loan amount of \$200,000. Four days later, the same loan officer met with a Hispanic male tester and refused to provide any information or service. The loan officer told the Hispanic tester, “we usually don’t meet with anyone without doing the credit check, it would be a waste of time for you and for me” and added “you can go to other lenders, they might be able to help you without first pulling out your credit as every mortgage corporation has a different policy.” When the tester pressed one more time to obtain an estimate, the loan officer stated, “I’m sorry, but I can’t answer your questions without first pulling your credit.”

provided, and specific financing options discussed was about 10 to 12 percent, while the incidence of Hispanic-favored treatment on these indicators was only about 1 percent. For the overall information summary score, Anglos were favored 12.7 of the time in Los Angeles, while Hispanics were favored only 1.3 percent of the time. In Chicago, on the

other hand, the Hispanic/Anglo results were comparable to the black/white results discussed earlier, with no significant pattern of differential treatment favoring either Hispanics or Anglos.

How much were testers told they could afford? Testers asked loan officers for help in figuring out a price range for their housing search and loan amounts for which they would be qualified. Because the minority and white Anglo members of each tester team were so closely matched with respect to their financial characteristics, they should have received comparable estimates from the lending institutions they visited. If minority homebuyers do not receive estimates of affordable loan amounts and house prices that are comparable to those provided to white Anglo homebuyers, their housing search may be discouraged or misdirected. They may conclude that they cannot afford the kind of house they want, or they may limit their search to lower-priced neighborhoods than they could in fact afford. The next group of treatment indicators focuses on the maximum loan amounts and house prices that testers were told they could afford. This analysis is limited to tests in which both testers were able to exchange accurate financial information and obtain either a loan amount or a house price from the loan officer.¹⁶

The two members of a tester pair were often quoted different house prices or loan amounts, but in general these differences did not systematically favor either white Anglos or minorities.¹⁷ Exhibit 3 reports the share of cases in which one tester was quoted a higher loan amount or house price than his or her partner. For black/white tests in both Los Angeles and Chicago, and for Hispanic/Anglo tests in Los Angeles no statistically significant patterns of either white-favored or minority-favored treatment occurred.

¹⁶ Results are the same when all tests in which both testers received a loan amount or house price are included in the analysis, regardless of whether or not financial information was exchanged.

¹⁷ Although minority and white tester financial characteristics were closely matched, testing protocols did call for the minority tester to be slightly more qualified than his or her white partner. The incidence indicators used for loan amount and house price analysis incorporate a 5 percent threshold to account for the slight variations in qualifications. More specifically, one tester in a pair was considered to be favored over the other if the house price or loan amount he was quoted exceeded his partner's by at least 5 percent.

Exhibit 3: Who Was Quoted A Higher Loan Amount or House Price?

Black/white tests

Treatment Indicators	Los Angeles			Chicago		
	% White Favored	% Black Favored	Significance Test	% White Favored	% Black Favored	Significance Test
Maximum loan amount	32.6	26.1	0.70	26.9	17.3	0.40
Maximum house price	36.2	27.7	0.58	25.5	21.6	0.84
Average Amounts	White	Black	Significance Test	White	Black	Significance Test
Maximum loan amount	\$262,438	\$266,484	0.92	\$182,730	\$183,122	0.28
Maximum house price	\$285,783	\$291,186	0.89	\$193,824	\$193,312	0.56

Hispanic/Anglo Tests

Treatment Indicators	Los Angeles			Chicago		
	% Anglo Favored	% Hispanic Favored	Significance Test	% Anglo Favored	% Hispanic Favored	Significance Test
Maximum loan amount	42.9	32.7	0.51	51.9	19.2	0.01**
Maximum house price	44.0	30.0	0.32	51.0	13.7	0.00**
Average Amounts	Anglo	Hispanic	Significance Test	Anglo	Hispanic	Significance Test
Maximum loan amount	\$271,570	\$266,172	0.34	\$190,193	\$180,301	0.00**
Maximum house price	\$283,846	\$284,737	0.64	\$199,932	\$188,055	0.00**

* difference between % Anglo favored and % Hispanic favored is statistically significant at a 90% confidence level or higher

** difference between % Anglo favored and % Hispanic favored is statistically significant at a 95% confidence level or higher

However, for Hispanic/Anglo tests in Chicago, Hispanics testers were systematically disfavored on loan amounts and house prices. Anglo testers were quoted higher loan amounts or house prices than their Hispanic partners about 50 percent of the time, while Hispanics were favored less than 20 percent of the time. In addition to this significant incidence finding, the severity, or magnitude of loan amount and house price differences was also statistically significant. The average loan amount quoted to Anglo testers in Chicago was \$190,200. This amount was almost \$10,000 higher than the average of

A loan officer pre-qualified a white male tester for a home price of \$185,000 and a maximum loan amount of \$175,750. The loan officer also provided the white tester with a "Pre-Qualification Certificate." Eight days later, the same loan officer met with an African American male tester and pre-qualified him for a home price of \$165-175,000 and a maximum loan amount of \$160,000. The loan officer did not provide the African American tester with a "Pre-qualification Certificate."

\$180,300 quoted to their Hispanic partners. The average house price suggested to Anglo testers—about \$200,000—was \$12,000 higher than the average suggested to comparable Hispanic homebuyers—\$188,000.

How many loan products were testers told about? In addition to loan amounts and house prices, testers inquired about specific loan products that might meet their needs, including products recommended by the loan officer. Therefore, we compared the experience of minority and white Anglo testers in terms of the number of loan products described. Homebuyers presumably benefit when they are able to obtain information about a variety of financing options. If minorities receive less complete information about available loan products than comparable white Anglo homeseekers, they may have to spend more time and effort to find out about available products, or they may actually be unable to take advantage of beneficial products.

In many cases, white Anglo testers learned about more products than their minority partners. (See Exhibit 4). In the Chicago black/white tests, lenders discussed more products with white testers than their black partners 48.2 percent of the time, while blacks were favored in only 21.4 percent of tests. Overall, whites learned about an average of 2.95 products per visit while blacks learned about only 2.43 products. The same pattern occurred for Hispanic/Anglo tests in both Los Angeles and Chicago. In Los Angeles, although the *incidence* of differential treatment was not statistically significant, the magnitude measure shows that Anglos were told about an average of 2.92 products when Hispanic testers were told about only 2.33 (this difference is statistically significant). In Chicago, both the incidence and severity measures were statistically significant for the Hispanic/Anglo tests. Anglos were told about more products than their Hispanic partners 55.6 percent of the time, while Hispanics were only favored in 27.8 percent of tests. Moreover, Anglos learned about an average of 2.83 products per visit, compared to only 2.44 products per visit for Hispanics.

A loan officer told a white male tester about five possible loan options, including both conventional and FHA products. Eight days later, the same loan officer only discussed one loan option (an FHA product) with the Hispanic tester.

Exhibit 4: Who Learned About More Products?

Black/white tests

Treatment Indicators	Los Angeles			Chicago		
	% White Favored	% Black Favored	Significance Test	% White Favored	% Black Favored	Significance Test
Number of products	38.8	32.7	0.74	48.2	21.4	0.02**
Average Amounts	White	Black	Significance Test	White	Black	Significance Test
Number of products	1.2.53	2.41	0.44	2.95	2.43	0.01**

Hispanic/Anglo Tests

Treatment Indicators	Los Angeles			Chicago		
	% Anglo Favored	% Hispanic Favored	Significance Test	% Anglo Favored	% Hispanic Favored	Significance Test
Number of products	51.9	34.6	0.23	55.6	27.8	0.04**
Average Amounts	Anglo	Hispanic	Significance Test	Anglo	Hispanic	Significance Test
Number of products	2.92	2.33	0.03**	2.83	2.44	0.05*

* difference between % Anglo favored and % Hispanic favored is statistically significant at a 90% confidence level or higher

** difference between % Anglo favored and % Hispanic favored is statistically significant at a 95% confidence level or higher

How much “coaching” did testers receive? One of the important services loan officers provide at the pre-application stage is “coaching”—advice and assistance to potential homebuyers about how they can improve their qualifications as borrowers. If minority customers receive less coaching than comparable white Anglos, they may lack the information they need to correct problems that might prevent them from obtaining mortgage financing. Thus, coaching represents a subtle but potentially important form of differential treatment in the mortgage lending process at the pre-application stage.

In the Homeownership Testing Project, testers reported whether loan officers provided advice or guidance about any of the following:

- paying down debts;
- debt consolidation;
- downpayment assistance (e.g., gift from family, special program, etc.);
- seller assistance (e.g., paying points, downpayment, closing costs, etc.);

- pre-qualification or pre-approval letter; or
- homebuying seminar.

All of these information items would help potential borrowers improve their prospects for obtaining a mortgage loan. We compared the number of these items that were discussed with each of the two members of a tester pair; Exhibit 5 presents the results of this analysis.

Whites received more coaching than their minority partners in a statistically significant share of the black/white tests in both Los Angeles and Chicago. In both markets, whites were favored over half the time, while their minority partners were favored less than one fourth of the time. In Los Angeles, white testers received advice about 1.42 topics per visit compared to only 0.8 topics per visit for their black partners. In Chicago, white testers received advice about 1.54 topics per visit compared to only 1.06 topics per visit for blacks.

A loan officer informed a white male tester about the possibility of having the seller pay the closing costs and referred the tester to a real estate agent. In making the referral, the agent stated "Let me give you someone's card. He's a good guy. You'll get along with him, I know. He's your age" and added that "he'll know what you are looking for." The same loan officer did not inform the African American male tester about the possibility of a seller paying closing costs and did not refer the tester to a real estate agent.

The same pattern of unequal coaching occurred for Hispanic/Anglo tests in Chicago, where Anglos were favored in 41.3 percent of the tests while their Hispanic partners were favored in only 23.8 percent of tests. On average, Anglos received information about 0.99 topics per visit compared to only 0.64 topics per visit for Hispanics. There was no statistically significant pattern of unequal coaching for Hispanic/Anglo tests in Los Angeles.

Exhibit 5: Who Received More Coaching?

Black/White Tests

Treatment Indicators	Los Angeles			Chicago		
	% White Favored	% Black Favored	Significance Test	% White Favored	% Black Favored	Significance Test
Positive Coaching	43.2	18.9	0.01**	50.6	29.1	0.03**
Average Amounts	White	Black	Significance Test	White	Black	Significance Test
Positive Coaching	1.84	1.3	0.01**	1.74	1.53	0.33

Hispanic/Anglo Tests

Treatment Indicators	Los Angeles			Chicago		
	% Anglo Favored	% Hispanic Favored	Significance Test	% Anglo Favored	% Hispanic Favored	Significance Test
Positive Coaching	39.2	36.7	0.80	40.5	15.2	0.01*
Average Amounts	Anglo	Hispanic	Significance Test	Anglo	Hispanic	Significance Test
Positive Coaching	1.67	1.60	1.00	1.25	1.14	1.00

* difference between % Anglo favored and % Hispanic favored is statistically significant at a 90% confidence level or higher

** difference between % Anglo favored and % Hispanic favored is statistically significant at a 95% confidence level or higher

Did testers receive follow-up contact? After the pre-application visit ended, some testers received follow-up telephone calls or mail from loan officers. These contacts may serve an important marketing and educational function, providing additional information to potential homebuyers, encouraging them to continue the homebuying process, and letting them know that they are valued customers. If minorities receive less follow-up of this kind from lending institutions they visit, they may become discouraged about their housing search or may take longer to find suitable mortgage financing. Follow-up contacts to testers were systematically recorded, and we compared the number of contacts received by the white Anglo and minority members of each tester pair (see Exhibit 6).¹⁸

Exhibit 6: Who Received More Follow-Up?

Black/White Tests

Treatment Indicators	Los Angeles			Chicago		
	% White Favored	% Black Favored	Significance Test	% White Favored	% Black Favored	Significance Test
Follow-Up Contact	5.4	5.4	1.00	12.7	1.3	0.01**
Average Amounts	White	Black	Significance Test	White	Black	Significance Test
Follow-Up Contact	0.11	0.05	0.67	0.22	0.03	0.01**

¹⁸ Mail and telephone follow-up were given equal weight, with each piece of mail or telephone call counted as one follow-up contact.

Hispanic/Anglo Tests

Treatment Indicators	Los Angeles			Chicago		
	% Anglo Favored	% Hispanic Favored	Significance Test	% Anglo Favored	% Hispanic Favored	Significance Test
Follow-Up Contact	6.3	13.9	0.21	7.6	13.9	0.33
Average Amounts	Anglo	Hispanic	Significance Test	Anglo	Hispanic	Significance Test
Follow-Up Contact	0.19	0.38	0.06*	0.16	0.28	0.21

* difference between % Anglo favored and % Hispanic favored is statistically significant at a 90% confidence level or higher

** difference between % Anglo favored and % Hispanic favored is statistically significant at a 95% confidence level or higher

Only the black/white tests in Chicago show a statistically significant pattern of unequal follow-up. In all three other tracks of tests (black/white in Los Angeles and Hispanic/Anglo in both Los Angeles and Chicago), differential treatment was commonplace, but it did not systematically favor either white Anglos or minorities. The Chicago results are quite dramatic, however. White testers received more follow-up than their black partners in 12.7 percent of the tests, compared to only 1.3 percent of tests in which the black testers were favored. On average, whites in Chicago received 0.22 follow-up calls per visit to a lending institution, while blacks received only 0.03 follow-up calls per visit. In Los Angeles, although the *incidence* of differential treatment was not statistically significant, the magnitude measure suggests that Hispanics received more follow-up than their Anglo partners. Specifically, Hispanics received an average of 0.38 follow-up contacts per visit compared to an average of only 0.19 for Anglos (this difference is statistically significant).

After meeting with a white and a black tester, the loan officer asked both for their addresses. The white tester received follow-up contact in the form of a business card and a brochure about an additional financing option to consider. The black tester received no follow-up mail or telephone calls.

Were testers encouraged to consider FHA? One type of product that lenders may suggest to prospective homebuyers is an FHA loan. Because FHA guidelines are relatively flexible, they serve some borrowers who do not meet conventional underwriting standards. However, FHA loans can cost more than conventional loans over the long term and may also permit lenders to charge higher fees. Moreover, in Los Angeles, the testers' financial characteristics qualified them for a conventional loan above the FHA maximum, so that recommending FHA might unnecessarily constrain their search options.

We constructed two variables that reflect the information provided by lenders about FHA. The first indicates whether testers were *encouraged* to consider FHA, and includes cases in which an FHA loan was among the specific products described to the tester as well as cases in which the loan officer recommended FHA. The second variable indicates whether FHA was explicitly *discouraged*, and includes cases in which the loan officer told testers that FHA would not be a good option for them to consider. Exhibit 7 reports the share of tests in which FHA was either encouraged or discouraged.

Black testers in Los Angeles were substantially more likely to be encouraged to consider FHA than their white partners. Specifically, FHA was encouraged for black testers but not their white partners in 30.8 percent of tests, while FHA was encouraged to whites but not blacks in only 7.7 percent of tests. None of the results in the remaining groups of tests (black/white in Chicago and Hispanic/Anglo in Los Angeles and Chicago) provide evidence of systematic differences in treatment on the basis of race or ethnicity. In general, white Anglos and minorities received the same encouragement or discouragement, and when differences occurred they did not favor either white Anglos or minorities at statistically significant levels.

Exhibit 7: Who Was Encouraged to Consider FHA?

Black/White Tests

Treatment Indicators	Los Angeles			Chicago		
	% White Higher	% Black Higher	Significance Test	% White Higher	% Black Higher	Significance Test
FHA encouraged	7.7	30.8	0.10*	9.8	14.6	0.75
FHA discouraged	7.7	3.8	0.63	12.2	4.9	0.45

Hispanic/Anglo Tests

Treatment Indicators	Los Angeles			Chicago		
	% Anglo Higher	% Hispanic Higher	Significance Test	% Anglo Higher	% Hispanic Higher	Significance Test
FHA encouraged	15.0	30.0	0.51	13.0	13.0	1.00
FHA discouraged	0.0	0.0	—	2.2	2.2	1.00

* difference between % Anglo higher and % Hispanic higher is statistically significant at a 90% confidence level or higher

** difference between % Anglo higher and % Hispanic higher is statistically significant at a 95% confidence level or higher

Sensitivity Analysis. For all of the treatment indicators reported above, additional analysis stratified the tests by three potentially important factors: whether both

testers in the pair met with the same or different loan officers, whether the lender was a prime or sub-prime institution, and whether testers met with a minority agent. Because of the small sample sizes, these stratified analyses were based on pooled data from both Los Angeles and Chicago.

Although not mandated by the testing protocols, in approximately half of all tests both testers in a pair happened to meet with the same loan officer. In the other half of tests, testers met with different loan officers from the same branch of the same institution. The question of whether testers saw the same loan officer is particularly important because it is often used by lending institutions to explain away observed differences in treatment. When tests were stratified by whether testers saw the same loan officer or different loan officers, the patterns of differential treatment remained essentially the same. There was no evidence that statistically significant differences in treatment were more prevalent when the two testers saw different loan officers.

Tests were also stratified based on whether the lending institution had been identified by HUD as a sub-prime lender. Although our sampling protocols did not mandate the inclusion of sub-prime institutions, they were tested to the extent that they were represented among HMDA-reporters in our two sites. Tests at sub-prime institutions made up 10.5 percent of all tests conducted in Los Angeles, and only 5 percent of tests completed in Chicago. Only 7.7 percent of all tests in the sample as a whole were conducted at sub-prime lending institutions. Since statistical analyses on such a small number of tests are not reliable, a true stratification analysis based on sub-prime/prime lenders could not be conducted. However, when sub-prime tests were excluded from the analysis, patterns of differential treatment were essentially the same for all treatment indicators.¹⁹ Therefore, there was no evidence from this testing effort that sub-prime lenders were either more or less likely to treat minorities unfavorably at the pre-application stage.

Finally, we explored differences in test results based upon the race or ethnicity of the loan officer. In about 30 percent of all tests, testers met with a minority loan officer. One hypothesis is that minority loan officers may treat minority customers more favorably than white Anglos. In fact, we found that white-favored treatment occurred somewhat more frequently when the loan officer was a white Anglo than when the loan officer was minority. But there was no increase in minority-favored treatment when the loan officer was minority.

¹⁹ To identify subprime lenders, we used HUD's 1999 list of sub-prime lenders. More information about this list, and a copy of the list itself (table A.1) can be found at: <http://www.huduser.org/datasets/manu.html>.

Implications for Future Research and Enforcement Testing

Based upon the results of this pilot testing effort, we conclude that paired testing at the pre-application stage of the mortgage lending process is both feasible and effective for both enforcement and research purposes. The results demonstrate convincingly that unequal treatment of minority homebuyers at the pre-application stage of the mortgage lending process remains a significant problem. Additional testing clearly should be conducted. But because of the complexity of this kind of testing and the differences between metro area results, we do not recommend a large-scale national study of adverse treatment at the pre-application stage of the mortgage lending process. Instead, systematic studies should be conducted on a site-by-site basis, with ample time and resources for effective training, test coordination, and quality control.

The testing protocols implemented for this study focused on only one source on pre-application information, one set of borrower profiles, and one type of pre-application inquiry. They do not tell us how often adverse treatment occurs for other types of borrowers, making different types of pre-application inquiries to other information providers. Therefore, future testing efforts can and should focus on other information sources, including mortgage brokers, new home sales agents, mobile home dealers, and referrals from real estate agents. They should include testers posing as less well-qualified borrowers as well as homebuyers with a particular house price in mind. And they should employ different protocols, including scenarios in which testers approach lending institutions with a specific house price in mind, and request detailed information about available loan products.

Some of the tests conducted for this pilot evidence provided anecdotal evidence that lenders may encourage or discourage customers from considering particular neighborhoods. This may include comments that steer minority customers away from predominantly white areas or that discourage whites from investing in minority neighborhoods. However, our testing protocols were not designed to systematically elicit or record comments of this kind or to analyze the characteristics of communities discussed by loan officers. Future testing efforts should consider strategies for systematically capturing and analyzing this type of differential treatment.

A loan officer informed a white female tester that Monterey Park is "Asian-infested" and that homes quickly sell to relatives of existing Asian residents. The loan officer told the white tester that she could consider homes in Pasadena, Glendale, Silverlake and the San Gabriel Valley and that she might want to consider Alhambra because it has easy access to the freeway and she could "live quietly in a safe neighborhood." The African American tester received no comments or suggestions about specific neighborhoods.

Finally, our experience demonstrates that research and enforcement testing can be linked without compromising the integrity of either. It is important to distinguish testing for research from testing conducted primarily for law enforcement purposes.

Testing for research generally seeks to produce generalizable results regarding the prevalence of adverse treatment for a market area or for the nation as a whole. To achieve these generalizable results, tests are randomized using an accepted sampling frame, and quite large numbers of tests are conducted in order to support statistically significant comparisons. To generate reliable and objective comparisons of minority and white experiences across a large number of tests, researchers usually use highly structured reporting forms, with closed-ended, “check the box”-type items.

By contrast, the purpose of an enforcement test is to establish legal violations and to correct them either through settlement or litigation. Testing for enforcement is often complaint driven, and typically is targeted to a single firm or a selected set of firms. Enforcement testing often requires multiple tests of a single firm, but generally does not involve the large number of tests typical of research testing. As a consequence, enforcement testing report forms tend to be much more open-ended, requiring test partners to provide greater narrative detail, rather than check boxes. These forms are generally analyzed pair-by-pair by a knowledgeable analyst who compares the treatment of test partners across all aspects of the encounter, including subjective as well as objective information.

Although research and enforcement testing differ in significant ways, the distinctions between the two should not be overdrawn. Both are based on the same core methodology and protocols, differing primarily in the way test results are recorded and analyzed. Thus, innovative testing methodologies developed for one application can be adapted for the other. Moreover, randomized testing of large numbers of market transactions need not be limited to research. They can and should be applied in targeting for enforcement. Furthermore, research and enforcement testing can be conducted in tandem, yielding *both* market-wide estimates of the incidence of adverse treatment *and* case-specific evidence of individual violations (evidence that might be followed-up with enforcement oriented testing).

Conclusion

Rigorously designed and implemented paired testing demonstrates that in both Los Angeles and Chicago, African American and Hispanic homebuyers face a significant risk of receiving less favorable treatment than comparable whites when they visit mortgage lending institutions to inquire about financing options. In the majority of cases, minorities and whites received equal treatment, or when differences occurred, they were equally likely to favor the minority as the white.²⁰ But in both metropolitan areas, paired

²⁰ It is important to recognize that even when we do not observe a statistically significant pattern of systematically unequal treatment, discrimination may have occurred in individual cases. For some treatment variables, differences in treatment occurred quite frequently but favored minorities just as often as whites.

testing revealed statistically significant patterns of unequal treatment that systematically favored whites.²¹

Unequal treatment took different forms in the two metropolitan areas and for the two minority groups. In Los Angeles –

- Blacks were offered less coaching than comparable white homebuyers, and were more likely to be encouraged to consider an FHA loan.
- Hispanics were denied basic information about loan amount and house price, told about fewer products, and received less follow-up compared to Anglo homebuyers.

In Chicago –

- Blacks were denied basic information about loan amount and house price, told about fewer products, offered less coaching, and received less follow-up than comparable white homebuyers.
- Hispanics were quoted lower loan amounts or house prices, told about fewer products, and offered less coaching than comparable Anglo homebuyers.

These patterns of unequal treatment occurred regardless of whether the two members of a tester pair met with the same loan officer or with different loan officers.

The frequency of unfavorable treatment varied considerably from one category to another. For some categories, equal treatment occurred in the vast majority of cases, but when differences occurred, the white tester was dramatically more likely to be favored than his or her minority partner. For other treatment categories, differences in treatment

Two female testers, one white and one black, visited the same Los Angeles area lender two days apart and met with the same loan officer. The testers told the loan officer that they were first-time homebuyers and needed assistance in figuring out a home price range and a loan amount for which they might qualify. The loan officer requested and obtained detailed information on household income, debts, and assets from both testers and asked about their respective credit situations. He then estimated that the white tester would qualify for a \$332,500 loan to purchase a \$350,000 home and, but estimated that the black tester would qualify for a \$237,500 loan to purchase a \$245,000 home. The loan officer told the white tester that a seller would likely pay some of the closing costs, but no mention was made about seller assistance to the black tester. The loan officer also told the white tester that it was a good idea to have a home inspection conducted prior to purchase, while the loan officer did not mention anything about the value of a home inspection to the black tester. The loan officer provided a complete loan application package to the white tester, but not to the black tester.

These differences may result from random variations in the behavior of loan officers, but they may also include cases of discrimination on the basis of race or ethnicity.

²¹ Results are reported as statistically significant when the difference between the rate of white/Anglo-favored treatment and the rate of minority-favored treatment is significant at a 90 percent confidence level or higher.

occurred much more often, but again were substantially more likely to favor the white than the minority tester. The series of charts below illustrate the levels of unfavorable treatment for each treatment category where statistically significant differences were observed.

Despite the variations in the forms that differential treatment takes, these findings demonstrate that African American and Hispanic homebuyers in both Los Angeles and Chicago face a significant risk of unequal treatment when they visit main-stream mortgage lending institutions to make pre-application inquiries. Discriminatory treatment at this early stage in the mortgage lending process has the potential to discourage some minorities from continuing their housing search, to limit their search to lower cost homes than they could actually afford, or to put them at a disadvantage in the loan application and underwriting process by failing to inform them about the most favorable loan products or about ways to improve their creditworthiness.

References

- Carr, James H. and Isaac F. Megbolugbe. 1993. "The Federal Reserve Bank of Boston Study on Mortgage Lending Revisited." *Journal of Housing Research* 4(2):277-314.
- Fix, Michael and Margery Austin Turner.,eds. 1999. *A National Report Card on Discrimination in America: The Role of Testing*. Washington, D.C.: The Urban Institute.
- Freiberg, Fred and Carla Herbig. 2002. *Guide to Enforcement Tools for Fair Lending Testing*. Washington, D.C.: The Urban Institute
- Goering, John M. and Ron Wienk, eds. 1996. *Mortgage Lending, Racial Discrimination, and Federal Policy*. Washington, D.C.: The Urban Institute Press.
- Heckman, James J. and Peter Siegelman. 1993. "The Urban Institute Audit Studies: Their Methods and Findings" in *Clear and Convincing Evidence: Testing for Discrimination in America* (Fix, Michael, and Raymond J. Struyk eds.). Washington, DC: The Urban Institute Press.
- Ladd, Helen. 1998. "Evidence on Discrimination in Credit Markets." *Journal of Economic Perspectives* 12(Spring):41-62.
- Munnell, Alicia, et al. 1986. "Mortgage Lending in Boston: Interpreting HAD Data." *American Economic Review* 86(March): 25-53.
- Ramsey, Fred L. and Daniel W. Schafer. 1997. *The Statistical Sleuth: A Course in Methods of Data Analysis*. Belmont, CA: Wadsworth Publishing Company.
- Ross, Stephen and John Yinger. 2002. *The Color of Credit: Mortgage Discrimination, Research Methodology, and Fair-Lending Enforcement*. Cambridge, MA: The MIT Press.
- Turner, Margery A., Stephen Ross, George Galster, and John Yinger. 2002. *Discrimination in Metropolitan Housing Markets: Results from Phase I of HDS2000*. Washington, D.C: U.S. Department of Housing and Urban Development:..
- Turner, Margery A. and Felicity Skidmore, eds. 1999. *Mortgage Lending Discrimination: A Review of Existing Evidence*. Washington, D.C.: The Urban Institute .

Yezer, Anthony M. ed. 1995. *Fair Lending Analysis: A Compendium of Essays on the Use Of Statistics*. Washington, D.C: American Bankers Association..