Spatial Variation in GSE Mortgage Purchase Activity

Bradford Case
Board of Governors of the Federal Reserve System

Kevin Gillen
Susan M. Wachter
University of Pennsylvania

Abstract

This study analyzes the government-sponsored enterprises' (GSEs') mortgage purchase patterns over the period from 1993 through 1996 and focuses on their share of the secondary mortgage market in specific market segments identified by borrower income, borrower race, and other indicators of policy interest. Using a database on GSE loan purchases from HUD's Public Use Database (PUDB) combined with data on non-GSE loan purchases reported under the Home Mortgage Disclosure Act in 44 of the largest metropolitan areas in the country, we provide a picture of GSE mortgage purchase patterns in a variety of urban areas.

We report a series of cross-tabulations estimating the market share of each GSE by borrower and neighborhood characteristics, coupled with a logistic regression analysis on the influence of specific borrower and neighborhood characteristics on the probability that a given loan will be purchased by one of the GSEs. These analyses suggest that during the period covered by the study, Fannie Mae and Freddie Mac provided a lower proportion of funding for mortgage lending to lower income and minority borrowers than to higher income or White borrowers. The GSEs also had lower market shares in lower income neighborhoods than in higher income neighborhoods, in central-city areas compared to suburban areas, and in neighborhoods that are geographically targeted according to HUD's mandates for GSE loan purchase activity compared to nontargeted neighborhoods. The logistic regressions further suggest that the GSEs were more likely to purchase loans in racially mixed tracts than in predominantly White tracts.

Finally, we focus on spatial differences in GSE mortgage purchase patterns using clustering methods and find that GSE purchases differ in all included California metropolitan areas (along with Boston, Newark, New York, and Washington) compared with the rest of the metropolitan statistical areas (MSAs) studied. Loans made to borrowers with relatively high loan balances were less likely to be purchased by the GSEs in the California-plus metropolitan areas than in the remaining metropolitan areas. This may reflect the relatively high housing prices in the California-plus

metropolitan areas. Because GSEs are prohibited from purchasing jumbo loans that exceed a conforming loan limit, they can be expected, other things being equal, to have a smaller market share in areas with higher housing prices because a larger share of loans can be expected to exceed the loan limit. Moreover, loans originated to minority borrowers are more likely to be purchased by the GSEs in the California-plus MSAs, a difference that may be attributable in part to the different mix of minority borrowers in the California MSAs, which have higher population proportions of Asian Americans compared to African Americans.

During the post-World War II era, the housing finance system in the United States has succeeded in making mortgage credit available, and homeownership affordable, for the majority of American households. Notwithstanding this success, it is evident that mortgage credit is systematically less available to particular groups of households, including minorities and those living in predominantly minority, low-income, and central-city neighborhoods. In an effort to eliminate these disparities in mortgage lending activity, Congress passed the Federal Housing Enterprises Financial Safety and Soundness Act of 1992 (FHEFSSA), which called on HUD to establish performance targets to help ensure that Fannie Mae and Freddie Mac, as government-sponsored enterprises (GSEs), would adequately promote the public purposes specified in the charter acts for both GSEs:

- To provide stability in the secondary market for residential mortgages.
- To respond appropriately to the private capital market.
- To provide ongoing assistance to the secondary market for residential mortgages (including activities relating to mortgages on housing for low- and moderate-income families involving a reasonable economic return that may be less than the return earned on other activities) by increasing the liquidity of mortgage investments and improving the distribution of investment capital available for residential mortgage financing.
- To promote access to mortgage credit throughout the Nation (including central cities, rural areas, and underserved areas) by increasing the liquidity of mortgage investments and improving the distribution of investment capital available for residential mortgage financing.

Federal Government sponsorship of Fannie Mae and Freddie Mac directly serves an important policy function in the Nation's housing finance system. Federal sponsorship, by enabling mortgage lenders to offer housing finance at lower mortgage interest rates, makes homeownership affordable to a wider range of households. The cost advantage that makes such reduced interest rates possible, however, also benefits both GSEs in their capacities as profit-making enterprises. In return for the implied guarantee of Federal sponsorship, the government requires that the GSEs operate in a manner that serves the public policy interests specified above.

In particular, HUD established housing goals under the 1992 FHEFSSA legislation designed to encourage GSE purchases of loans to segments of the population that have limited access to mortgage credit. In 1993 HUD set the following goals for both GSEs in the transition period 1993–95:

- **Low- and Moderate-Income Goal.** Thirty percent of units financed by mortgage purchases (28 percent for Freddie Mac in 1993) should be either owner-occupied units for which the borrower's income is less than or equal to area median income or

rental units with rents (adjusted for unit size) not exceeding 30 percent of area median income.

- **Geographically Targeted Goal.** Thirty percent of units financed by mortgage purchases (26 percent for Freddie Mac and 28 percent for Fannie Mae in 1993) should be located in central cities as defined by the Office of Management and Budget.
- **Special Affordable Goal.** Mortgage purchases of $16.4 billion for Fannie Mae and $11.9 billion for Freddie Mac in 1993 and 1994 combined, and $4.6 billion for Fannie Mae and $3.4 billion for Freddie Mac in 1995, should finance owner-occupied units for which the borrower's income was less than or equal to 60 percent of area median income, less than or equal to 80 percent of area median income and located in low-income census tracts or nonmetropolitan counties, or rental units affordable at these income levels.

In 1995 HUD revised these goals for the period 1996–99, increasing the required proportion of total mortgage purchases meeting each of the goals and changing the definition of geographically targeted areas:

- **Low- and Moderate-Income Goal.** Forty percent of mortgage purchases (42 percent in 1997–99) should be of mortgages originated to households with incomes less than or equal to the area median income.
- **Geographically Targeted Goal.** Twenty-one percent of purchases (24 percent in 1997–99) should be of mortgages on dwelling units in census tracts with minority concentration of at least 30 percent with tract median income less than or equal to 120 percent of area median income, or census tracts with, in metropolitan areas, median family income less than or equal to 90 percent of area median income or, in nonmetropolitan areas, median family income less than or equal to 95 percent of the greater of State or national nonmetropolitan median income.
- **Special Affordable Goal.** Twelve percent of purchases (14 percent in 1997–99) should be of mortgages originated to households with income less than or equal to 60 percent of area median income, or less than or equal to 80 percent of area median income and located in low-income areas.

Several studies show that since the goals were established, the GSEs have increased the proportion of their total mortgage purchase volume that represents mortgages from the market segments identified as *underserved.* For example, Harold Bunce and Randall Scheessele (1996) found that "both GSEs have significantly improved their performance over the past four years" (1992–95), and a subsequent followup (Bunce and Scheessele, 1998) showed that "both GSEs have improved their performance over the longer period from 1992 to 1996." Paul Manchester, Sue Neal, and Bunce (1998) found that "both GSEs' performance exceeded their low-mod goals in every year during the transition period." These studies indicate that both GSEs have made significant progress toward meeting or exceeding the mandates for mortgage purchases from underserved market segments as a proportion of their total mortgage purchases.

The mandates focus on the proportion of each GSE's total mortgage purchases that represents mortgages from underserved market segments. The mandates address the probability that a loan was made in an underserved market, given that it was purchased by one of the GSEs. This can be expressed in probability notation as:

$$\text{Prob}[\text{underserved} \mid \text{GSE}],$$

where "Prob[underserved]" indicates the probability that any loan was originated in a segment of the population with limited access to mortgage credit, and "| GSE" indicates that the probability is conditional on the loan having been purchased by a GSE—that is, that we are considering only those loans in the GSE-purchased portion of the mortgage market.

It is also of interest, however, to focus on a somewhat different question—the GSE purchases of mortgages as a share of the total origination of loans in underserved markets. This alternative can be viewed as the probability that a mortgage originated to an underserved borrower is sold to one of the GSEs. That is, given that a mortgage has been approved and originated to a low-income or otherwise underserved borrower, this approach evaluates how likely it is that the mortgage will be purchased by the GSEs. In probability notation, this is:

$$\text{Prob}[\text{GSE} \mid \text{underserved}],$$

where "Prob[GSE]" indicates the probability that any loan was purchased by a GSE, and "| underserved" indicates that the probability is conditional on the loans having originated in the underserved portion of the mortgage market—that is, that we are considering only those loans originated among households with limited access to mortgage credit.

Glenn B. Canner, Wayne Passmore, and Dolores S. Smith (1994) researched this question and found that the proportion of GSE mortgage purchases from several market segments—including low-income and minority borrowers and those in low-income, central-city, and predominantly minority areas—lagged behind the proportion of non-GSE mortgages from these same market segments. Bunce and Scheessele (1996) also found that "the shares of the GSEs' business going to lower income borrowers and underserved neighborhoods typically fall short of the corresponding shares of other market participants."

The alternative way of evaluating GSE mortgage purchase activity described here is relevant because the implicit Federal guarantee of GSE debt is believed to reduce finance costs for mortgage borrowers whose loans are sold to the GSEs, and this measure focuses on the proportion of underserved borrowers who benefit from the indirect Federal subsidy. This alternative measure addresses more explicitly the particular concern raised over the aggregate funding flows to the underserved, since increasing these flows and the volume of lending directly supports the goal of increased homeownership among the underserved.

In order to evaluate the proportion of loans in underserved market segments that are purchased by the GSEs, it is essential to have information on the non-GSE portion of the secondary mortgage market: that is, on the number of loans that are originated in underserved market segments but that are not purchased by the GSEs as well as on the number of loans that are originated in underserved market segments and that are purchased by the GSEs. The only database that includes information on mortgages from both the GSE and the non-GSE portions of the mortgage market is the database collected under the Home Mortgage Disclosure Act (HMDA). Thus, HMDA data were used by Canner, Passmore, and Smith (1994); Bunce and Scheessele (1996, 1998); Manchester, Neal, and Bunce (1998); and Manchester (1998).

However, the HMDA database is subject to several significant problems that may limit its utility for studies of GSE mortgage purchase activity. James Berkovec and Peter Zorn (1996), for example, compared HMDA loans identified as having been sold to Freddie Mac with actual Freddie Mac purchases, and found that "HMDA data covered about 66

percent of mortgage loans sold to Freddie Mac in 1992, in the range of other estimates of HMDA coverage."[1] Moreover, many of the loans recorded in the HMDA database are incomplete and some appear to be inaccurate.

HUD's Public Use Database (PUDB) of loans purchased by both GSEs provides essentially universal coverage of the GSE portion of the secondary mortgage market. Although most of the earlier studies of GSE mortgage purchase patterns were based on HMDA data, more recent studies (Bunce and Scheessele, 1996, 1998; William Segal and Edward Szymanoski, 1997; Manchester, 1998; and Manchester, Neal, and Bunce, 1998) employed the loan-level data on GSE mortgage purchases contained in the PUDB. Because the PUDB does not cover the non-GSE portion of the market, however, we combine the PUDB with HMDA data on the non-GSE portion of the mortgage market to develop a more complete picture of the role of the GSEs relative to the entire mortgage market in the provision of credit to homebuyers.

This report presents the results of our research project conducted using the recently released data on GSE mortgage purchase activity available in the PUDB, in combination with data collected under HMDA, to investigate spatial variation in mortgage purchase activity by Fannie Mae and Freddie Mac. The combined PUDB-HMDA data set was employed in a series of cross-tabulations and multiple regressions for 44 of the Nation's largest metropolitan statistical areas (MSAs)[2] to pinpoint those market segments in which the GSEs tend to provide funding for a relatively high (or relatively low) share of mortgage loans originated.

There are two caveats to our analysis. It is possible that HMDA data represent a nonrandom sample of all mortgage originations. Indeed, Berkovec and Zorn (1996) found that the sample of loans recorded in the HMDA database was selective; in particular, loans from low-income areas were more likely to be included in the sample of loans recorded in the HMDA database than were loans from high-income areas, so that lending in higher income areas would be underrepresented relative to lending in lower income areas.[3] If this is true, then combining the PUDB and HMDA data sets may itself result in biased estimates of the GSE market shares. Specifically, if it is true that loans from lower income areas are more likely to be included in the HMDA database, then combining the two data sets may result in the appearance of a relatively low share of GSE purchases of loans for the low end of the market—that is, those who are underserved—and this appearance would be simply an artifact of HMDA underreporting. We test for this by replicating our cross-tabulation results using HMDA data alone. We show cross-tabulations for borrower income categories across all MSAs and for all categories of policy interest across a subset of MSAs. Our results do not suggest that the GSEs' market share for the underserved is biased downward by use of the combined data sets.

In addition to the data reporting and other problems presented by the HMDA data, the use of this database to investigate the suggested alternative measure of GSE mortgage purchase activity—the proportion of loans to underserved borrowers purchased by the GSEs—may be subject to a second important caveat: reporting of mortgage originations that are not purchased by the GSEs may have increased over the past several years.[4] If true, this means that the non-GSE portion of the secondary mortgage market may appear to be increasing more than it actually is, and therefore, that GSE purchase activity, as a proportion of the total secondary mortgage market, may appear to be increasing less rapidly than it actually is. Thus, the GSE share of geographically targeted mortgage originations (like the GSE share of nontargeted originations) may in fact be increasing over time more rapidly than the available data suggest.

The remainder of this article is organized into three sections. "Merging the PUDB With HMDA Mortgage Data" describes in detail the methodology used in combining the PUDB and HMDA databases to create the combined data set employed in the analysis. The second section presents the results of statistical cross-tabulations conducted using the combined PUDB-HMDA data set in 44 MSAs. The final section offers the results of a series of logistic regression analyses designed to control for several sources of variation simultaneously. An appendix presents detailed empirical results for Philadelphia as an example of the results summarized in this article. (All empirical results summarized here can, of course, be reproduced in the same way that the Philadelphia results were produced.)

Merging the PUDB With HMDA Mortgage Data

This research project was designed to take full advantage of HUD's PUDB by combining its data with data collected under HMDA to provide a fuller picture of mortgage lending activity for nationwide analysis. As noted, the PUDB data provide universal coverage of loans purchased by either Fannie Mae or Freddie Mac during the period 1993–96, but do not include any loans that were originated but not sold to either GSE during that time period. In contrast, the HMDA data include loans originated during the same period regardless of whether or not they were purchased by one of the GSEs, but does not provide universal coverage of the mortgage market.

We believe that combining the PUDB data with HMDA data is the only satisfactory way to analyze mortgage purchase activity in particular market segments by each GSE, not only relative to the other GSE, but also relative to the non-GSE portion of the mortgage market. However, while both PUDB and HMDA databases include loan-level data, loans are not given any identifier that can be used to match loans uniquely across the two files. This makes the task of combining PUDB and HMDA data sets exceedingly difficult. Fortunately, the two databases contain common variables, and these variables can be used to develop a *match* across the databases by aggregating the data upward to synthetic pools based on unique permutations of these variables.

We use eight common variables of policy interest to combine data on individual loans into GSE and non-GSE pools:

- State.
- County.
- MSA.
- Census tract.
- Borrower race (White/minority).
- Borrower income (categorized).
- GSE purchaser (Fannie Mae/Freddie Mac).
- Loan origination/acquisition year.[5]

To match individual loan observations across the two databases, we formed pools of loans in each database defined by a unique permutation of these eight common variables. For example, one pool might consist of all loans from each database that originate in Pennsylvania, in Chester County, in the Philadelphia MSA, in census tract 3001.011, to a minority borrower with income in the top decile of the MSA income

distribution, with the loan originated/acquired in 1994, and sold to Fannie Mae. The pool containing all of the loans from the PUDB having this particular permutation of values on the common variables was then matched with the corresponding pool containing all of the loans from the HMDA database having the same permutation of values on the common variables.

It is important to point out that the majority of the synthetic pools identified using this process consisted of only one loan. That is, using the example above, there may have been only one loan to a minority borrower of the highest income decile in census tract 3001.011, Chester County, Pennsylvania, Philadelphia MSA, that was originated and sold to Fannie Mae in 1994. This means that, although there is no way for us to know for certain whether this procedure yields an accurate matching of loans, we have confidence that a large share of the matched pools did in fact accomplish a correct matching of individual loans across the two databases.

In performing this pooling and matching process, we did not use all of the variables common to the two databases, because a few of them had too many missing or implausible values to provide a reliable basis for matching. For example, we tested the possibility of pooling and matching using data on borrower gender as well as on coborrower race and gender in addition to the eight common variables listed above, but the number of pools that failed to match across databases suggested that those variables were insufficiently reliable. Similarly, we decided not to pool and match on the basis of loan balance because the difference in variable definitions across the databases (amount at origination in the HMDA database and unpaid balance at acquisition in the PUDB), along with rounding and other data recording problems, made it impossible to rely on these data for matching.

The borrower race variable provided information in five specific racial/ethnic categories (American Indian or Alaskan Native, Asian or Pacific Islander, Black, Hispanic, and White) as well as four other categories (Other, Information Not Provided, Not Applicable, and Not Available). Although we believe that the identification of borrowers as members of minority (non-White) racial/ethnic groups may have been reasonably consistent, the HMDA and PUDB data suggested that the detailed identification of minority borrowers as members of one of the four separate minority racial/ethnic groups may have been much less consistent. For this reason, we elected to pool and match loan observations only according to whether the loan was originated to a White or minority (non-White) borrower, or to a borrower identified in one of the other categories for which reliable data on borrower race was missing.

In some cases the value for one or more of the common geographic variables (State, county, or MSA) was missing or implausible, and we were able to supply or correct it before performing the pooling and pool-matching process. For example, the values given for county and State might be used to fill in a missing value or correct an implausible value for MSA.

It is useful to recognize that this pooling and matching process involves a tradeoff between two goals. The larger the set of common variables (or values for categorical variables) that are used, the more likely it is that each synthetic pool will consist of a single loan. While achieving a nearly loan-level match may seem appealing, however, it is important to note that it would also introduce two types of problems. Many of the loans may be mismatched across databases because of the problems of inconsistent variable definitions and simple data entry errors. Other loans are likely not to match across databases at all. Mismatching and nonmatching could be minimized by pooling and merging

only the most reliable variables, such as geography (State, MSA, county, and census tract) and GSE purchaser; on the other hand, this would mean that each loan pool would be quite heterogeneous with respect to borrower and loan data. Thus, the choice of which variables to use in the loan aggregation and matching process involves a tradeoff between the goals of increasing the homogeneity of loans in each pool (and increasing the number of loan-level matches) and increasing the accuracy of loan matches made. Before deciding on the specific set of eight common variables and the specific categorizations described above, we tested several possible aggregation schemes to strike what we believe is the appropriate balance between these two goals.

Exhibit 1 presents a summary of the loan pooling and data merging process for each of the 44 large MSAs included in the analysis. Column 1 in this exhibit shows, for each MSA, the total number of loans listed in the GSE PUDB. Column 2 shows the number of loan pools formed by aggregating loans into the synthetic pools according to the eight common variables. For example, in Philadelphia, the PUDB lists 263,550 loans acquired by Fannie Mae or Freddie Mac during the study period 1993–96, and the aggregation process resulted in 99,327 pools.[6] Column 3 in exhibit 1 shows the total number of Philadelphia MSA loan originations listed in the HMDA database. The next two columns show the number of pools formed by aggregating according to the eight common variables, with column 4 showing the number of pools of loans that were identified as not having been sold to one of the GSEs. Column 5 states the number of pools of loans not identified as having been sold to a GSE. Loan pools in the HMDA data that are identified as GSE acquisitions are then matched with loan pools from the PUDB, while non-GSE loan pools are appended to reflect the non-GSE portion of the mortgage market. In Philadelphia, for example, the HMDA data set lists 524,909 loan originations. These loans were aggregated into 80,231 GSE loan pools and 96,646 non-GSE loan pools.[7]

After the loan aggregation was complete, the PUDB loan pools were matched with the GSE loan pools formed from the HMDA database according to the same set of eight common variables. Column 6 shows the number of pools in each MSA that matched across the GSE and HMDA data sets (for example, 62,257 loan pools matched for Philadelphia) indicating the extent to which information would have been lost if we had relied only on HMDA data. Finally, the pools of non-GSE loans from the HMDA database were added, as were the nonmatched pools of PUDB loans to create the final data set used in our subsequent empirical analysis. The final number of pools shown in column 7, then, is the sum of column 5, non-GSE pools in the HMDA database, and column 2, GSE pools in the PUDB. Thus, column 7 is the total number of observations in the final MSA data set used for all our subsequent analysis.

In summary, we used the following decision rules to combine the data: all GSE data in the PUDB are included regardless of whether or not a corresponding match for each observation is found in the HMDA database, and loans included in the HMDA database but not acquired by a GSE are simply appended onto the PUDB. Consequently, GSE loans for which no match is found in the HMDA data simply have missing values for those fields found only in the HMDA database. The only observations that are dropped altogether are those HMDA loans coded as GSE acquisitions for which no match is found in the GSE PUDB.[8]

There are important differences between the PUDB and HMDA data sets in their coverage of specific segments of the market, in addition to the question of their coverage of the overall market. These differences affect the loan-matching process as well as the interpretation of results. For example, both GSEs are prohibited from purchasing mortgages with balances that exceed a *conforming loan limit* established by statute; thus all

Exhibit 1

Summary of Loan Pooling and Database Merging Process

MSA	(1) Loans PUDB	(2) Pools PUDB	(3) Loans HMDA	(4) GSE Pools HMDA	(5) Non-GSE Pools HMDA	(6) Matched Pools	(7) Pools [(2)+(5)]	(8) GSE Coverage Rate [(6)/(2)] (%)	(9) HMDA Coverage Rate [(6)/(4)] (%)
ANA Anaheim	189,626	70,202	377,333	53,259	55,032	44,416	125,234	63.3	83.4
ATL Atlanta	257,507	62,798	579,461	48,699	54,158	39,917	116,956	63.6	82.0
BAL Baltimore	130,154	45,461	321,339	36,077	48,999	28,834	94,460	63.4	79.9
BIR Birmingham	44,695	14,068	101,366	10,841	15,305	8,433	29,373	59.9	77.8
BOS Boston	215,620	69,359	391,642	57,551	53,579	46,030	122,938	66.4	80.0
BUF Buffalo	52,835	19,050	98,506	15,490	18,244	12,827	37,294	67.3	82.8
CHI Chicago	547,726	171,214	1,073,052	131,042	151,695	108,155	322,909	63.2	82.5
CIN Cincinnati	112,307	32,957	225,619	25,558	30,978	21,006	63,935	63.7	82.2
CLE Cleveland	104,932	44,565	279,559	30,986	50,297	25,440	94,862	57.1	82.1
COL Columbus	108,411	30,736	212,579	25,023	28,836	20,745	59,572	68.3	82.9
DAL Dallas	138,089	45,744	340,943	36,685	42,671	28,662	88,415	62.7	78.1
DEN Denver	178,087	51,796	411,241	43,516	43,378	35,549	95,174	68.6	81.7
DET Detroit	350,852	102,937	659,260	88,796	85,123	75,105	188,060	73.0	84.6
FTW Ft. Worth	58,544	21,491	165,955	17,260	22,674	13,202	44,165	61.4	76.5
HAR Hartford	49,399	21,177	105,237	15,246	19,629	12,277	40,806	58.0	80.5
HOU Houston	169,294	52,022	316,282	42,367	45,463	34,096	97,485	65.5	80.5
IND Indianapolis	105,184	29,317	219,931	23,934	27,467	19,837	56,784	67.7	82.9
KAN Kansas City	98,036	31,247	229,118	25,615	31,359	20,258	62,606	64.8	79.1
LAN Los Angeles	444,386	184,412	879,997	132,243	146,498	108,848	330,910	59.0	82.3

Exhibit 1 (continued)

Summary of Loan Pooling and Database Merging Process

MSA	(1) Loans PUDB	(2) Pools PUDB	(3) Loans HMDA	(4) GSE Pools HMDA	(5) Non-GSE Pools HMDA	(6) Matched Pools	(7) Pools [(2)+(5)]	(8) GSE Coverage Rate [(6)/(2)] (%)	(9) HMDA Coverage Rate [(6)/(4)] (%)
MEM Memphis	36,216	11,126	128,478	7,983	18,998	5,895	30,124	53.0	73.8
MIA Miami	104,226	29,557	224,112	21,851	26,190	18,048	55,747	61.1	82.6
MIL Milwaukee	106,043	29,270	189,066	24,497	28,811	19,443	58,081	66.4	79.4
MIN Minneapolis	220,141	63,848	491,229	53,037	54,858	44,531	118,706	69.7	84.0
NWO New Orleans	44,700	19,267	119,765	15,746	24,583	11,507	43,850	59.7	73.1
NYC New York City	329,203	154,911	692,115	113,329	153,353	83,259	308,264	53.7	73.5
NEW Newark	84,923	36,998	182,294	29,044	37,002	21,420	74,000	57.9	73.8
NOR Norfolk	48,813	19,244	204,112	15,770	30,288	11,998	49,532	62.3	76.1
OKL Oklahoma City	38,148	15,904	109,531	13,600	22,398	9,649	38,302	60.7	70.9
PHI Philadelphia	263,550	99,327	524,909	80,231	96,646	62,257	195,973	62.7	77.6
PHO Phoenix	177,497	51,115	452,207	41,910	47,679	34,673	98,794	67.8	82.7
PIT Pittsburgh	60,432	28,481	188,634	20,449	44,926	15,280	73,407	53.6	74.7
POR Portland	150,710	43,670	257,174	34,121	30,989	29,110	74,659	66.7	85.3
PRO Providence	51,292	18,664	100,170	14,826	17,903	11,611	36,567	62.2	78.3
ROC Rochester	50,015	18,470	96,723	15,383	17,481	12,593	35,951	68.2	81.9
STL St. Louis	149,090	41,458	355,642	34,392	38,922	28,014	80,380	67.6	81.5
SAL Salt Lake City	104,207	28,481	244,771	25,037	23,361	20,491	51,842	71.9	81.8
SAN San Antonio	40,152	14,218	124,276	11,623	18,835	8,563	33,053	60.2	73.7
SBE San Bernardino	158,159	44,236	378,540	33,636	38,113	27,506	82,349	62.2	81.8

Exhibit 1 (continued)

Summary of Loan Pooling and Database Merging Process

MSA	(1) Loans PUDB	(2) Pools PUDB	(3) Loans HMDA	(4) GSE Pools HMDA	(5) Non-GSE Pools HMDA	(6) Matched Pools	(7) Pools [(2)+(5)]	(8) GSE Coverage Rate [(6)/(2)] (%)	(9) HMDA Coverage Rate [(6)/(4)] (%)
SDI San Diego	176,559	56,773	351,485	42,497	45,523	36,029	102,296	63.5	84.8
SFR San Francisco	96,425	39,721	224,726	28,119	32,423	23,626	72,144	59.5	84.0
SJO San Jose	132,951	46,755	257,121	35,024	33,406	30,422	80,161	65.1	86.9
SEA Seattle	206,576	60,796	374,430	45,133	43,593	38,660	104,389	63.6	85.7
TAM Tampa	119,342	40,130	280,165	29,801	40,180	24,601	80,310	61.3	82.6
WAS Washington	307,817	108,565	748,770	86,441	99,924	68,547	208,489	63.1	79.3

nonconforming (*jumbo*) loans must be held in the lender's portfolio or sold to an investor other than Fannie Mae or Freddie Mac. These nonconforming loans are included in the HMDA data but not in the PUDB data, because both GSEs are prohibited from purchasing them. Nonconforming loans are included in this analysis in order to form as complete as possible a picture of the overall mortgage market and the roles of Fannie Mae and Freddie Mac in that overall market. Because nonconforming loans are generally originated to higher income borrowers, this means that the GSEs may appear to purchase a small share of all loans to higher income borrowers.

As another example, in order to maintain their financial safety and soundness, both GSEs have generally refrained from purchasing *subprime* loans—that is, loans originated to borrowers with credit problems at a higher interest rate that includes a risk premium reflecting the additional default risk associated with these borrowers. Subprime loans are included in the HMDA data but do not generally appear in the PUDB data because the GSEs generally have not purchased subprime loans.[9] As with conforming loans, subprime loans are included in this analysis to form a picture of the activities of the GSEs in the overall mortgage lending market that is as complete as possible. Because subprime loans are generally originated to lower income borrowers, this means that the GSEs may appear to purchase a smaller share of all loans originated to lower income borrowers. Moreover, as there is some evidence that the volume of subprime loans increased during the study period of 1993–96, it may appear that the GSEs' share of loans originated to lower income borrowers is declining simply because the volume of subprime loans to these borrowers is increasing.

Finally, it is important to recognize the difference between loan origination year (recorded in the HMDA database) and loan acquisition year (recorded in the PUDB), and the difficulty that this inconsistency in variable definitions posed for matching loans across the two databases. Fortunately, most loans are sold to Fannie Mae or Freddie Mac shortly after they are originated, if at all. We believe that by matching records from the HMDA database by year of origination (for loans recorded as having been sold to a GSE) to records from the PUDB by year of acquisition, it is likely that most loans originated and sold to a GSE were matched correctly. There are two groups of loans, however, for which this matching procedure would be unlikely to yield satisfactory results. The first group consists of loans guaranteed by the Federal Housing Administration (FHA). Because GSE guidelines require that FHA loans be seasoned at least 1 year before origination, we matched FHA loans from the HMDA database that were originated in one year with loans from the PUDB that were acquired in the following year and shared the same values on the eight common variables. The second group consists of loans originated near the end of any year, since it typically takes some 1 to 3 months to accomplish the sale of a loan to Fannie Mae or Freddie Mac; we explored matching loans from the PUDB that were acquired in January of one year with loans from the HMDA database that were originated the previous year. The number of successful loan matches was higher, however, when we simply matched origination and acquisition years without adjusting in this way. Although we have no way of knowing the extent to which these matching guidelines were successful in pairing records correctly across the two databases, we believe that they reflect as closely as possible the actual pattern of origination and sale of mortgages and significantly improve the accuracy of the resulting data set.

It is important to keep in mind that without HMDA data on non-GSE mortgages we would be unable to determine market share of GSE purchases; thus, we cannot rely on PUDB data alone for our analysis. On the other hand, the PUDB provides more complete coverage of GSE mortgage purchases than does HMDA, since HMDA coverage is not designed to be universal. In addition, the HMDA data set may not correctly indicate whether the loan was purchased by a GSE. The strategy of combining data sets may itself

give rise to biases. To test for this, we computed the GSE market shares in specific market segments using HMDA data alone and using the combined PUDB-HMDA data set, and report those sets of results.[10, 11]

Cross-Tabulations of Loan Purchase Activity in Different Market Segments

This section presents the results of a series of cross-tabulations performed using the PUDB-HMDA data set in each of the 44 large MSAs included in the analysis. The cross-tabulations facilitate comparisons among the GSEs and the non-GSE portion of the market in several different market segments and enable us to detect general patterns in the extent to which each GSE is an active purchaser of loans made to borrowers in each market segment. In addition, since the PUDB and HMDA databases cover substantially all loans originated over the entire period 1993–96, we also conducted cross-tabulations separately for each of the 4 years in order to investigate whether the performance of either GSE in any given market segment seems to have changed systematically over the time period covered by the analysis.

The appendix presents an illustrative example (for Philadelphia) of the statistical cross-tabulations summarized in this report. The cross-tabulations detail GSE market share in market segments defined by the following variables:

- **Borrower Income Category.** Fifteen categories defined, as described in Section 2, by the ratio of borrower income to MSA median income.
- **Borrower Race.** White versus minority.
- **Ratio of Tract Median Income to MSA Median Income.** Five categories: 0 to 60 percent, 60 to 80 percent, 80 to 100 percent, 100 to 120 percent, and greater than 120 percent.
- **Tract Percentage Minority.** Five categories: 0 to 10 percent, 10 to 15 percent, 15 to 30 percent, 30 to 50 percent, and greater than 50 percent.
- **Central-City Versus Suburban Tracts.**[12]
- **Geo-Targeted Census Tracts.** Compared with nontargeted tracts.

The appendix tables show the total amount (as measured by aggregate loan balance[13]) of loans originated by Fannie Mae, Freddie Mac, and, in the case of exhibit A–1, which reports data for all purchasers, non-GSE ("Other") mortgage market participants in each market segment in Philadelphia. The appendix exhibits also show the share of each market participant in the total volume of loans for that market segment in Philadelphia. For example, the first section of the "All Purchasers, 1993–96" report in appendix exhibit A–1 shows that the total dollar balance of loans made during the period was $60.8 billion (bottom, column 8). Fannie Mae purchased loans valued at $13.7 billion or 22 percent of the total, while Freddie Mac purchased $11.5 million or 19 percent of the total; the remaining $35.6 billion or 59 percent were held by non-GSE market participants (bottom of columns 2, 4, and 6).

Borrower Income Category: Illustrative Results for Philadelphia MSA

Of the total amount of loans in Philadelphia, appendix exhibit A–1 shows that $1 billion (or 2 percent, as shown in the last column) consisted of loans to borrowers whose incomes were less than 50 percent (0–0.5) of the MSA median. Of these loans made

to the lowest income borrowers, Fannie Mae purchased $127.7 million (or 12 percent), meaning that Fannie Mae's market share among the lowest income borrowers in Philadelphia was lower than its market share among all borrowers (12 percent versus 22 percent). Freddie Mac purchased $53.5 million or 5 percent of loans made to the lowest income borrowers, meaning that Freddie Mac's market share among the lowest income borrowers in Philadelphia was also lower than its market share among all borrowers (5 percent versus 19 percent). In contrast, loans valued at $871.0 million, or 83 percent of loans made to the lowest income borrowers, were originated but not purchased by the GSEs. This means that the non-GSE market share among the lowest income borrowers in Philadelphia was higher than the non-GSE market share among all borrowers (83 percent versus 59 percent).

A useful contrast with the GSEs' market share among the lowest income borrowers is provided by an upper middle-income market segment, borrowers whose incomes were between 150 percent and 200 percent (1.5–2.0) of MSA median income. Exhibit A–1 shows that loans made to these borrowers in Philadelphia totaled $11.1 billion during the study period, making it the largest market segment by borrower income at 18 percent of the total. Fannie Mae purchased loans made to these higher income borrowers totaling $3.1 billion or 28 percent, above its overall market share of 22 percent. Freddie Mac purchased loans in this market segment totaling $2.9 billion or 27 percent, also above its overall market share of 19 percent.

In comparing the GSE and non-GSE market shares of loans made to borrowers in different categories, it is important to keep in mind two very important restrictions on mortgage purchase activity that apply to both GSEs. First, the GSEs are prohibited from acquiring jumbo loans. These loans, therefore, must be held by non-GSE market participants. Second, the need to maintain a sound financial condition for the enterprises may reduce their ability to purchase loans from riskier segments of the mortgage market.[14] In the context of evaluating GSE mortgage purchase activity by borrower income category, the first restriction tends to prevent GSEs from acquiring some loans originated to the highest income borrowers, since many of these loans are likely to exceed the conforming loan limit. In contrast, the second restriction tends to prevent GSEs from acquiring some loans originated to the lowest income borrowers, since many of these loans are likely to present greater risk of default.

Borrower Income Category: Summary Results for 44 MSAs

The results presented in the cross-tabulations computed for all 44 large MSAs can be summarized simply in a table indicating whether each GSE's market share in a given market segment was less than, greater than, or approximately the same as its overall market share. Moreover, this tabular form enables us to discern whether there are any clear patterns in the extent to which each of the GSEs finances mortgage lending in particular market segments across multiple MSAs.

Exhibit 2 presents such a tabular summary of results, showing Fannie Mae's relative share of the total market by borrower income category across all 44 MSAs. The left column shows, for each borrower income category, the MSAs in which Fannie Mae's market share for this market segment was less than its overall market share in that MSA. In other words, it shows the MSAs in which Fannie Mae finances proportionately less mortgage lending in that market segment than do Freddie Mac and non-GSE market participants. The middle column shows the MSAs in which Fannie Mae's market share in that market segment is approximately equal to its overall market share,[15] and the right column shows the MSAs in which Fannie Mae finances proportionately more mortgage lending to borrowers in that market segment compared to other market participants.

Exhibit 2

Fannie Mae Relative Share of Total Market by Borrower Income Category

Borrower Income Category	Segment Market Share < Overall Market Share	Segment Market Share = Overall Market Share	Segment Market Share > Overall Market Share
0–0.5	ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, SEA, TAM, WAS		IND, SFR, SJO
0.5–0.6	ATL, BAL, BIR, BUF, CHI, CIN, CLE, COL, DEN, DET, FTW, HAR, IND, KAN, MEM, MIL, MIN, NWO, NYC, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, TAM	DAL, SBE	ANA, BOS, HOU, LAN, MIA, NEW, SDI, SFR, SJO, SEA, WAS
0.6–0.7	BAL, BIR, BUF, CIN, CLE, COL, DEN, FTW, IND, KAN, MEM, MIL, MIN, NWO, NOR, OKL, PIT, PRO, ROC, STL, SAN, TAM	ATL, DET, PHI, PHO, SAL, SBE	ANA, BOS, CHI, DAL, HAR, HOU, LAN, MIA, NYC, NEW, POR, SDI, SFR, SJO, SEA, WAS
0.7–0.8	BIR, BUF, CIN, CLE, COL, FTW, HAR, KAN, MEM, MIL, NWO, NOR, OKL, PIT, PRO, ROC, STL, SAL, SAN, TAM	BAL, DEN, IND	ANA, ATL, BOS, CHI, DAL, DET, HOU, LAN, MIA, MIN, NYC, NEW, PHI, PHO, POR, SBE, SDI, SFR, SJO, SEA, WAS
0.8–0.9	BUF, COL, IND, MEM, NWO, NOR, OKL, PIT, ROC, SAN, TAM	BIR, CIN, HAR, MIL, STL, SBE	ANA, ATL, BAL, BOS, CHI, CLE, DAL, DEN, DET, FTW, HOU, KAN, LAN, MIA, MIN, NYC, NEW, PHI, PHO, POR, PRO, SAL, SDI, SFR, SJO, SEA, WAS
0.9–1.0	BIR, IND, MEM, NWO, NOR, OKL, PRO, SAN	BAL, CIN, CLE, FTW, KAN, PIT, ROC, SBE, TAM	ANA, ATL, BOS, BUF, CHI, COL, DAL, DEN, DET, HAR, HOU, LAN, MIA, MIL, MIN, NYC, NEW, PHI, PHO, POR, STL, SAL, SDI, SFR, SJO, SEA, WAS
1.0–1.1	BUF, MEM, NOR, OKL, SAN	NWO, PIT	ANA, ATL, BAL, BIR, BOS, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MIA, MIL, MIN, NYC, NEW, PHI, PHO, POR, PRO, ROC, STL, SAL, SBE, SDI, SFR, SJO, SEA, TAM, WAS
1.1–1.2	MEM, SAN		ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MIA, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SBE, SDI, SFR, SJO, SEA, TAM, WAS

Exhibit 2 (continued)

Fannie Mae Relative Share of Total Market by Borrower Income Category

Borrower Income Category	Segment Market Share < Overall Market Share	Segment Market Share = Overall Market Share	Segment Market Share > Overall Market Share
1.2–1.3	NOR, OKL, SAN	MEM	ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MIA, MIL, MIN, NWO, NYC, NEW, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SBE, SDI, SFR, SJO, SEA, TAM, WAS
1.3–1.4		SAN	ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SBE, SDI, SFR, SJO, SEA, TAM, WAS
1.4–1.5			ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, SFR, SJO, SEA, TAM, WAS
1.5–2.0		SJO	ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, SFR, SEA, TAM, WAS
2.0–2.5	ANA, BOS, NEW, SFR, SJO, WAS	HAR, SEA	ATL, BAL, BIR, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, TAM
2.5–3.0	ANA, ATL, BAL, BOS, CHI, CLE, DET, HAR, LAN, MIL, NEW, PHI, POR, SDI, SFR, SJO, SEA, WAS		BIR, BUF, CIN, COL, DAL, DEN, FTW, HOU, IND, KAN, MEM, MIA, MIN, NWO, NYC, NOR, OKL, PHO, PIT, PRO, ROC, STL, SAL, SAN, SBE, TAM

Exhibit 2 (continued)

Fannie Mae Relative Share of Total Market by Borrower Income Category

Borrower Income Category	Segment Market Share < Overall Market Share	Segment Market Share = Overall Market Share	Segment Market Share > Overall Market Share
3.0	ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SBE, SDI, SFR, SJO, SEA, TAM, WAS	NOR	SAN
Missing	ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, SFR, SJO, SEA, TAM, WAS		CLE

For example, the first section of exhibit A–1 shows that Fannie Mae's market share in the lowest borrower income category (0 to 50 percent of MSA median) in Philadelphia was 12 percent, which is less than Fannie Mae's overall market share of 22 percent in Philadelphia. Exhibit 2, therefore, shows Philadelphia in the left column in that lower borrower income category. For borrowers with incomes between 60 percent and 70 percent of MSA median, Fannie Mae's market share (22 percent) was approximately equal to its overall market share, so exhibit 2 shows Philadelphia in the middle column for that borrower income category. Finally, for borrowers in the higher income category (70 to 80 percent of MSA median), Fannie Mae's market share (23 percent) was higher than its overall market share, so exhibit 2 shows Philadelphia in the right column for that borrower income category.

The tabular presentation enables us to discern a clear pattern in terms of Fannie Mae's performance in providing mortgage finance to borrowers in different income categories. For example, among the lowest income borrowers (those with incomes less than 50 percent of MSA median), Fannie Mae's market share is lower than its overall market share in 41 of the 44 MSAs. The only MSAs in which Fannie Mae has a relatively high market share among these borrowers are Indianapolis, San Francisco, and San Jose. The pattern holds, although less strikingly, in the next-higher borrower income category: among borrowers with incomes between 50 percent and 60 percent of MSA median, Fannie Mae's market share was low in 31 MSAs, the same as its overall market share in 2 MSAs, and high in the remaining 11 MSAs.

As exhibit 2 shows, Fannie Mae's market share is relatively high in the majority of MSAs for borrowers with incomes at least 80 percent of the MSA median. In fact, exhibit 2 shows that for borrowers whose incomes are between 140 and 150 percent of MSA median, Fannie Mae's market share is relatively high in all 44 of the MSAs evaluated. Finally, as noted, among the highest-income borrowers (those whose mortgages are most likely to

exceed the conforming loan limit), Fannie Mae's market share is lower than its overall market share—Fannie Mae has a relatively low market share in 42 of the 44 MSAs for borrowers with incomes at least three times the MSA median.

Exhibit 3 presents the same type of information as does exhibit 2, but this time for Freddie Mac instead of Fannie Mae. The pattern is the same: Freddie Mac provides less mortgage finance than do other secondary market participants to borrowers in the lowest income market segments in the large majority of cities, and provides more mortgage finance to borrowers in the middle- and upper income market segments (except those at the highest incomes). Indeed, the pattern is perhaps even more striking for Freddie Mac than for Fannie Mae. For example, among the lowest income borrowers (those with incomes less than 50 percent of MSA median), Freddie Mac had a lower market share in 43 of the 44 MSAs studied, the only exception being San Francisco.

Exhibit 4 is constructed to facilitate comparison of the performance of the GSEs relative to each other by showing Fannie Mae's relative share of the GSE portion of the mortgage market according to each borrower income category. Freddie Mac's relative share is simply the opposite of Fannie Mae's and is, therefore, indicated by reversing the left and right column headings. The left column shows those MSAs in which Fannie Mae's segment market share is less than its overall share of the GSE market and, therefore, those MSAs in which Freddie Mac's segment market share is more than its overall GSE market share. The middle column shows those MSAs in which Fannie Mae's segment market share—and therefore Freddie Mac's as well—is approximately the same as its overall GSE market share. The right column shows those MSAs in which Fannie Mae's segment market share is more than its overall GSE market share and, therefore, those MSAs in which Freddie Mac's segment market share is less than its overall GSE market share.

For example, the first section of the "GSE Purchasers Only, 1993–96" report for Philadelphia presented in appendix exhibit A–2 shows that the total dollar amount of loans purchased by both GSEs during 1993–96 was $25.2 billion (bottom of the sixth column). Fannie Mae's aggregate purchase of $13.7 billion represented 54 percent of this GSE-only total, while Freddie Mac's $11.5 billion represented the remaining 46 percent of the GSE-only total. In the lowest borrower income category (those with incomes less than 50 percent of MSA median), Fannie Mae purchased loans totaling $127.7 million or 70 percent of total GSE purchases, higher than its overall share of the GSE market. Freddie Mac's purchases of $53.5 million, on the other hand, represented just 30 percent of total GSE purchases, below its overall share of the GSE market. For comparison, among upper middle-income borrowers with incomes between 150 percent and 200 percent (1.5 to 2.0) of MSA median, Fannie Mae purchased loans totaling $3.2 billion or 52 percent of the GSE-only market in Philadelphia during this period, meaning that Fannie Mae's market share in this market segment was only somewhat less than its overall share of the GSE-only market. In contrast, Freddie Mac purchased loans totaling $2.9 billion from borrowers in this upper middle-income market segment, and its 48 percent share of this market segment is higher than its overall market share for GSE loans.

Exhibit 4 enables us to see a pattern concerning the performance of each GSE in purchasing mortgage loans made to borrowers at different income levels. Among lower income borrowers, Fannie Mae's market share is relatively high—and Freddie Mac's relatively low—in almost all of the 44 MSAs studied. For example, among the lowest-income borrowers (those with incomes less than 50 percent of the MSA median), Fannie Mae has a relatively high market share in every metropolitan area except New York City. This pattern holds for all borrowers with incomes below MSA median and less significantly for borrowers with incomes up to 120 percent of MSA median. In contrast, among higher income borrowers, Fannie Mae's market share is relatively low, and Freddie Mac's

Exhibit 3

Freddie Mac Relative Share of Total Market by Borrower Income Category

Borrower Income Category	Segment Market Share < Overall Market Share	Segment Market Share = Overall Market Share	Segment Market Share > Overall Market Share
0–0.5	ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, SJO, SEA, TAM, WAS		SFR
0.5–0.6	ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, TAM, WAS		ANA, LAN, SDI, SFR, SJO, SEA
0.6–0.7	ATL, BAL, BIR, BUF, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, MEM, MIA, MIL, MIN, NWO, NYC, NOR, OKL, PHI, PHO, PIT, PRO, ROC, STL, SAL, SAN, SBE, TAM	BOS, CHI, WAS	ANA, LAN, NEW, POR, SDI, SFR, SJO, SEA
0.7–0.8	ATL, BAL, BIR, BUF, CHI, CIN, CLE, COL, DAL, DEN, FTW, HAR, HOU, IND, KAN, MEM, MIA, MIL, NWO, NYC, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, TAM	MIN	ANA, BOS, DET, LAN, NEW, SBE, SDI, SFR, SJO, SEA, WAS
0.8–0.9	BUF, COL, HAR, HOU, IND, MEM, MIA, MIL, NWO, NOR, OKL, PHI, PIT, POR, ROC, STL, SAN, SBE, TAM	BAL, BIR, CIN, CLE, DEN, FTW, PHO	ANA, ATL, BOS, CHI, DAL, DET, KAN, LAN, MIN, NYC, NEW, PRO, SAL, SDI, SFR, SJO, SEA, WAS
0.9–1.0	BAL, BIR, CIN, FTW, HOU, IND, KAN, MEM, MIA, NWO, NOR, OKL, PIT, PRO, ROC, SAN, TAM	BUF, CLE, DAL, HAR, MIL, MIN, PHO, SAL, SBE	ANA, ATL, BOS, CHI, COL, DEN, DET, LAN, NYC, NEW, PHI, POR, STL, SDI, SFR, SJO, SEA, WAS
1.0–1.1	BUF, FTW, MEM, NWO, NOR, OKL, PIT, SAN	BIR, CLE, HOU, PRO, ROC	ANA, ATL, BAL, BOS, CHI, CIN, COL, DAL, DEN, DET, HAR, IND, KAN, LAN, MIA, MIL, MIN, NYC, NEW, PHI, PHO, POR, STL, SAL, SBE, SDI, SFR, SJO, SEA, TAM, WAS
1.1–1.2	MEM, NWO, SAN	OKL	ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MIA, MIL, MIN, NYC, NEW, NOR, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SBE, SDI, SFR, SJO, SEA, TAM, WAS

Exhibit 3 (continued)

Freddie Mac Relative Share of Total Market by Borrower Income Category

Borrower Income Category	Segment Market Share < Overall Market Share	Segment Market Share = Overall Market Share	Segment Market Share > Overall Market Share
1.2–1.3	OKL, SAN	MEM, NOR	ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MIA, MIL, MIN, NWO, NYC, NEW, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SBE, SDI, SFR, SJO, SEA, TAM, WAS
1.3–1.4		MEM, NWO	ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MIA, MIL, MIN, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, SFR, SJO, SEA, TAM, WAS
1.4–1.5		MEM, SAN	ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MIA, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SBE, SDI, SFR, SJO, SEA, TAM, WAS
1.5–2.0		SJO	ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, SFR, SEA, TAM, WAS
2.0–2.5	ANA, BOS, SFR, SJO, WAS	NEW	ATL, BAL, BIR, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, SEA, TAM
2.5–3.0	ANA, BOS, CHI, DET, HAR, LAN, NEW, POR, SDI, SFR, SJO, SEA, WAS	ATL, BAL, MIL	BIR, BUF, CIN, CLE, COL, DAL, DEN, FTW, HOU, IND, KAN, MEM, MIA, MIN, NWO, NYC, NOR, OKL, PHI, PHO, PIT, PRO, ROC, STL, SAL, SAN, SBE, TAM

Exhibit 3 (continued)

Freddie Mac Relative Share of Total Market by Borrower Income Category

Borrower Income Category	Segment Market Share < Overall Market Share	Segment Market Share = Overall Market Share	Segment Market Share > Overall Market Share
3.0	ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SBE, SDI, SFR, SJO, SEA, TAM, WAS	NOR	SAN
Missing	ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, SFR, SJO, SEA, TAM, WAS		

relatively high. Among the highest income borrowers (those with incomes at least 3 times MSA median), Fannie Mae's segment market share is lower than its overall market share in 28 MSAs, approximately the same in 13 MSAs, and relatively high in only 3 of the 44 MSAs studied.

In summary, the data on market share suggest that both GSEs have a relatively low market share among lower income borrowers and a relatively high market share among higher income borrowers—except those in the highest income category, whose loans are most likely to exceed the conforming loan limit. Although this general pattern holds for both GSEs, it appears more pronounced for Freddie Mac than for Fannie Mae.

Borrower Race

The second section of each appendix exhibit shows the relative market share of Fannie Mae and Freddie Mac in the Philadelphia MSA according to the race of the borrower (non-Hispanic White versus minority, including White Hispanic). Appendix exhibit A–1 also shows the relative share for non-GSE market participants. Exhibits 5 and 6 present a tabular summary of the relative share of Fannie Mae and Freddie Mac, respectively, in the total mortgage market.[16] These exhibits show a very similar pattern of relatively high market share among White borrowers and relatively low market share among minority borrowers. For example, Fannie Mae's market share among White borrowers is more than its overall market share in 36 of the 44 MSAs studied, approximately the same in 3 others, and less than its overall market share in only 5 (Anaheim, Los Angeles, Miami, San Francisco, and San Jose). Freddie Mac's market share among White borrowers is relatively high in 35 MSAs, approximately the same in 7 others, and relatively low in only 2 (Los Angeles and San Francisco).

Exhibit 7 shows Fannie Mae's relative share of the GSE-only portion of the market by borrower race. As this exhibit shows, Fannie Mae appears to be purchasing loans made to minority borrowers to a greater extent than does Freddie Mac. Fannie Mae's market share

Exhibit 4

Fannie Mae Relative Share of GSE Market by Borrower Income Category

Borrower Income Category	Segment Market Share < Overall Market Share	Segment Market Share = Overall Market Share	Segment Market Share > Overall Market Share
0–0.5	NYC		ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, SFR, SJO, SEA, TAM, WAS
0.5–0.6		BUF, SJO	ANA, ATL, BAL, BIR, BOS, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, SFR, SEA, TAM, WAS
0.6–0.7	ANA	SFR, SJO	ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, SEA, TAM, WAS
0.7–0.8	SFR, SJO	ANA, HAR	ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, SEA, TAM, WAS
0.8–0.9	ANA, MEM, SFR, SJO	BIR, BOS, DET, HAR, NWO, PRO, SAL, SAN	ATL, BAL, BUF, CHI, CIN, CLE, COL, DAL, DEN, FTW, HOU, IND, KAN, LAN, MIA, MIL, MIN, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, ROC, STL, SBE, SDI, SEA, TAM, WAS
0.9–1.0	ANA, SFR, SJO	BOS, COL, DET, SBE	ATL, BAL, BIR, BUF, CHI, CIN, CLE, DAL, DEN, FTW, HAR, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SDI, SEA, TAM, WAS

Exhibit 4 (continued)

Fannie Mae Relative Share of GSE Market by Borrower Income Category

Borrower Income Category	Segment Market Share < Overall Market Share	Segment Market Share = Overall Market Share	Segment Market Share > Overall Market Share
1.0–1.1	BOS, IND, NOR, OKL, SAN, SBE, SJO	ANA, BAL, COL, DEN, DET, HAR, MEM, NEW, SDI, SFR	ATL, BIR, BUF, CHI, CIN, CLE, DAL, FTW, HOU, KAN, LAN, MIA, MIL, MIN, NWO, NYC, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SEA, TAM, WAS
1.1–1.2	BIR, BOS, CLE, DET, IND, OKL, SFR, SJO	BAL, COL, HAR, LAN, MEM, NEW, PHI, PRO, SBE, SDI	ANA, ATL, BUF, CHI, CIN, DAL, DEN, FTW, HOU, KAN, MIA, MIL, MIN, NWO, NYC, NOR, PHO, PIT, POR, ROC, STL, SAL, SAN, SEA, TAM, WAS
1.2–1.3	BAL, BOS, BUF, CLE, DEN, DET, HAR, IND, MEM, NOR, OKL, PHI, POR, ROC, SFR	ATL, CHI, COL, KAN, MIL, MIN, NEW, PHO, PIT, STL, SAL, SAN, SJO, WAS	ANA, BIR, CIN, DAL, FTW, HOU, LAN, MIA, NWO, NYC, PRO, SBE, SDI, SEA, TAM
1.3–1.4	ANA, BAL, BOS, BUF, CIN, CLE, COL, DAL, DEN, DET, HAR, IND, NEW, OKL, PHI, PRO, STL, SAN, SFR, SJO	ATL, CHI, FTW, HOU, KAN, LAN, MIN, NOR, POR, ROC, SAL, SDI, TAM, WAS	BIR, MEM, MIA, MIL, NWO, NYC, PHO, PIT, SBE, SEA
1.4–1.5	ANA, BIR, BOS, BUF, CIN, CLE, COL, DAL, DEN, DET, HAR, HOU, IND, LAN, MIN, NEW, OKL, PHI, PIT, POR, PRO, ROC, STL, SAL, SJO	ATL, BAL, CHI, NWO, SBE, SDI, SFR, TAM, WAS	FTW, KAN, MEM, MIA, MIL, NYC, NOR, PHO, SAN, SEA
1.5–2.0	ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, DAL, DEN, DET, FTW, HAR, HOU, IND, LAN, MEM, NEW, OKL, PHI, PIT, POR, PRO, ROC, STL, SAL, SBE, SFR, SJO, TAM	COL, MIA, MIL, MIN, NWO, NYC, NOR, PHO, SAN, SEA, WAS	KAN, SDI
2.0–2.5	ANA, BAL, BIR, BOS, BUF, CHI, CIN, CLE, DAL, DEN, DET, FTW, HAR, HOU, IND, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SFR, SJO, TAM	ATL, KAN, SBE, SDI, SEA, WAS	COL, NOR
2.5–3.0	ANA, BAL, BOS, BUF, CHI, CIN, CLE, DAL, FTW, HAR, HOU, IND, LAN, MEM, MIA, MIL, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, ROC, STL, SAL, SAN, SDI, SFR, TAM	ATL, BIR, COL, DEN, DET, MIN, PRO, SBE, SJO, SEA, WAS	KAN

Exhibit 4 (continued)

Fannie Mae Relative Share of GSE Market by Borrower Income Category

Borrower Income Category	Segment Market Share < Overall Market Share	Segment Market Share = Overall Market Share	Segment Market Share > Overall Market Share
3.0	ANA, ATL, CHI, CLE, DAL, FTW, HOU, IND, KAN, LAN, MIA, MIL, MIN, NWO, NYC, OKL, PHI, PHO, PIT, POR, ROC, SAL, SAN, SBE, SDI, SFR, SEA, TAM	BAL, BIR, BOS, CIN, DEN, DET, MEM, NEW, NOR, PRO, STL, SJO, WAS	BUF, COL, HAR
Missing	CHI, COL, FTW, MIA, MIL, MIN, NOR, PHO, SEA, WAS	ATL, KAN, SDI	ANA, BAL, BIR, BOS, BUF, CIN, CLE, DAL, DEN, DET, HAR, HOU, IND, LAN, MEM, NWO, NYC, NEW, OKL, PHI, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SFR, SJO, TAM

Exhibit 5

Fannie Mae Relative Share of Total Market by Borrower Race

Borrower Race	Segment Market Share < Overall Market Share	Segment Market Share = Overall Market Share	Segment Market Share > Overall Market Share
White	ANA, LAN, MIA, SFR, SJO	CLE, NYC, SDI	ATL, BAL, BIR, BOS, BUF, CHI, CIN, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, MEM, MIL, MIN, NWO, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SEA, TAM, WAS
Minority	ATL, BAL, BIR, BUF, CHI, CIN, COL, DEN, DET, FTW, HAR, IND, KAN, MEM, MIL, MIN, NWO, NOR, OKL, PHI, PHO, PIT, PRO, ROC, STL, SAN, SBE, TAM		ANA, BOS, CLE, DAL, HOU, LAN, MIA, NYC, NEW, POR, SAL, SDI, SFR, SJO, SEA, WAS
Missing	ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, SFR, SJO, SEA, TAM, WAS		

Exhibit 6

Freddie Mac Relative Share of Total Market by Borrower Race

Borrower Race	Segment Market Share < Overall Market Share	Segment Market Share = Overall Market Share	Segment Market Share > Overall Market Share
White	LAN, SFR	ANA, CLE, OKL, POR, PRO, SJO, WAS	ATL, BAL, BIR, BOS, BUF, CHI, CIN, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, NOR, PHI, PHO, PIT, ROC, STL, SAL, SAN, SBE, SDI, SEA, TAM
Minority	ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, MEM, MIL, NWO, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, TAM	MIA, MIN, NEW, WAS	ANA, LAN, NYC, SDI, SFR, SJO, SEA
Missing	ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, SFR, SJO, SEA, TAM, WAS		

Exhibit 7

Fannie Mae Relative Share of GSE Market by Borrower Race

Borrower Race	Segment Market Share < Overall Market Share	Segment Market Share = Overall Market Share	Segment Market Share > Overall Market Share
White	ANA, BOS, CLE, DET, HAR, HOU, LAN, MIA, NYC, PHI, STL, SBE, SDI, SFR, SJO	BAL, BUF, CHI, CIN, COL, DAL, DEN, FTW, IND, KAN, NWO, NEW, NOR, OKL, PHO, PIT, POR, SAL, SAN, TAM, WAS	ATL, BIR, MEM, MIL, MIN, PRO, ROC, SEA
Minority	OKL, PIT	CIN, IND, LAN, MEM, MIN, PRO, ROC, SJO, WAS	ANA, ATL, BAL, BIR, BOS, BUF, CHI, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, KAN, MIA, MIL, NWO, NYC, NEW, NOR, PHI, PHO, POR, STL, SAL, SAN, SBE, SDI, SFR, SEA, TAM
Missing	ATL, BAL, BIR, CHI, CIN, CLE, COL, DAL, DEN, FTW, HOU, IND, KAN, MEM, MIA, MIL, MIN, NWO, NYC, NOR, OKL, PHO, PIT, POR, ROC, SAL, SAN, SEA, TAM, WAS	NEW, PRO	ANA, BOS, BUF, DET, HAR, LAN, PHI, STL, SBE, SDI, SFR, SJO

for minority borrowers is relatively high (and Freddie Mac's relatively low) in 33 of the 44 MSAs studied, approximately the same as its overall market share in another 9 MSAs, and relatively low (and Freddie Mac's relatively high) in only 2 MSAs. However, while Fannie Mae appears to be providing housing finance of minority borrowers more than Freddie Mac does, it is important to recall from exhibits 5 and 6 that both GSEs appear to be financing loans to minority borrowers to a lesser extent relative to the non-GSE portion of the mortgage market.

In summary, both GSEs appear to have relatively high market shares among White borrowers and relatively low market shares among minority borrowers, although this pattern is more pronounced for Freddie Mac than for Fannie Mae.

Ratio of Tract Median Income to MSA Median Income

The third section of each appendix exhibit shows the market share of each mortgage market participant according to the median income of all households located in the same census tract as the borrower's property, expressed as a percentage of the MSA median income. Exhibits 8 and 9 summarize the relative market shares of Fannie Mae and Freddie Mac, respectively, for each of the five market segments defined by tract median income. Again, the two exhibits show very similar results: both GSEs fund a lower percentage of borrowers in the lowest income census tracts—those with median incomes less than 60 percent of MSA median, between 60 and 80 percent, or, less strikingly, between 80 and 100 percent—and a higher percentage of borrowers in higher income census tracts.

Exhibit 10 presents Fannie Mae's relative share of the GSE-only portion of the market according to the same tract income categories. As this exhibit shows, Fannie Mae's share of the GSE market among borrowers in lower income tracts is relatively high (and Freddie Mac's relatively low) in almost all metropolitan areas studied. Conversely, Fannie Mae's relative market share among borrowers in the highest income tracts is low (and Freddie Mac's high) in most of the metropolitan areas.

In summary, both GSEs appear to have relatively low market shares among borrowers in lower income census tracts and relatively high market shares among borrowers in higher income census tracts, although this pattern is more pronounced for Freddie Mac than for Fannie Mae.

Tract Percentage Minority

The fourth section of each appendix exhibit presents market activity by each mortgage market participant according to the tract minority concentration, which is the percentage of population in the borrower's census tract that is minority, including White Hispanics. Exhibits 11 and 12 summarize the data from all MSAs in tabular form. The two exhibits show similar patterns, but in this case, the pattern for Fannie Mae (exhibit 11) is less striking than that for Freddie Mac (exhibit 12). For example, among borrowers living in tracts with no more than 10-percent minority population, Fannie Mae's share of the total mortgage market is relatively high in 17 of the MSAs studied, approximately equal to its overall market share in another 14 MSAs, and relatively low in the remaining 13 MSAs. For a comparison, Freddie Mac's market share in the same market segment is relatively high in 21 MSAs, approximately equal to its overall market share in another 16 MSAs, and relatively low in the remaining 7 MSAs. In predominantly minority neighborhoods, on the other hand, Fannie Mae had a relatively low market share in 35 MSAs and a relatively high market share in only 8 MSAs. Similarly, Freddie Mac had a relatively low market share in 38 MSAs and a relatively high market share in only 5 MSAs. These

Exhibit 8

Fannie Mae Relative Share of Total Market by Ratio of Tract to MSA Median Income

Tract/MSA Income Ratio	Segment Market Share < Overall Market Share	Segment Market Share = Overall Market Share	Segment Market Share > Overall Market Share
0–60	ATL, BAL, BIR, BUF, CHI, CIN, CLE, COL, DEN, DET, FTW, HAR, HOU, IND, KAN, MEM, MIA, MIL, MIN, NEW, NOR, OKL, PHI, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, TAM	BOS	ANA, DAL, LAN, NWO, NYC, PHO, SDI, SFR, SJO, SEA, WAS
60–80	BAL, BIR, BUF, CHI, CIN, CLE, COL, DEN, DET, IND, KAN, MEM, MIL, MIN, NWO, NEW, NOR, OKL, PHI, PHO, PIT, PRO, ROC, STL, SAL, SAN, SBE, TAM	ATL, FTW	ANA, BOS, DAL, HAR, HOU, LAN, MIA, NYC, POR, SDI, SFR, SJO, SEA, WAS
80–100	ATL, BAL, BIR, BUF, CIN, CLE, COL, DEN, DET, FTW, HOU, IND, KAN, MEM, MIN, NOR, OKL, PIT, ROC, STL, SAL, SAN, TAM	NWO, PHO	ANA, BOS, CHI, DAL, HAR, LAN, MIA, MIL, NYC, NEW, PHI, POR, PRO, SBE, SDI, SFR, SJO, SEA, WAS
100–120	BIR, DAL, FTW, MEM, NWO, NOR	BAL, HOU, IND, POR, PRO, SBE, TAM	ANA, ATL, BOS, BUF, CHI, CIN, CLE, COL, DEN, DET, HAR, KAN, LAN, MIA, MIL, MIN, NYC, NEW, OKL, PHI, PHO, PIT, ROC, STL, SAL, SAN, SDI, SFR, SJO, SEA, WAS
>120	ANA, BOS, CHI, DET, HAR, LAN, NYC, NEW, POR, PRO SDI, SFR, SJO, SEA, WAS	ATL, CLE, DAL, HOU, MIA, MIL, MIN, PHI, PHO	BAL, BIR, BUF, CIN, COL, DEN, FTW, IND, KAN, MEM, NWO, NOR, OKL, PIT, ROC, STL, SAL, SAN, SBE, TAM
Missing	ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HOU, IND, KAN, LAN, MIA, MIL, MIN, NWO, NYC, NEW NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, SFR, SJO, SEA, TAM, WAS	MEM	HAR

patterns suggest that both GSEs are providing less mortgage finance to borrowers in predominantly minority (or even moderately minority) neighborhoods than are non-GSE participants in the secondary mortgage market.

Exhibit 13 summarizes Fannie Mae's relative share of the GSE-only market according to the same minority-concentration categories. As this exhibit shows, Fannie Mae appears to finance loans to borrowers in high-minority neighborhoods more readily than does Freddie Mac. For borrowers in predominantly minority tracts, Fannie Mae's market share was relatively high (and Freddie Mac's relatively low) in 40 of the 44 MSAs studied,

Exhibit 9

Freddie Mac Relative Share of Total Market by Ratio of Tract to MSA Median Income

Tract/MSA Income Ratio	Segment Market Share < Overall Market Share	Segment Market Share = Overall Market Share	Segment Market Share > Overall Market Share
0–60	ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, SEA, TAM	WAS	LAN, SFR, SJO
60–80	ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, MEM, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, TAM	WAS	ANA, LAN, MIA, SDI, SFR, SJO, SEA
80–100	ATL, BAL, BIR, BUF, CHI, CIN, CLE, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, MEM, MIA, MIL, MIN, NWO, NOR, OKL, PHI, PIT, PRO, ROC, STL, SAL, SAN, SBE, TAM	NYC, PHO, POR	ANA, BOS, COL, LAN, NEW, SDI, SFR, SJO, SEA, WAS
100–120	BIR, DAL, FTW, HOU, MEM, MIA, NWO, NOR, OKL, SAN	BAL, COL, IND, KAN, PHO, PIT, PRO, ROC, SBE	ANA, ATL, BOS, BUF, CHI, CIN, CLE, DEN, DET, HAR, LAN, MIL, MIN, NYC, NEW, PHI, POR, STL, SAL, SDI, SFR, SJO, SEA, TAM, WAS
>120	ANA, BOS, CHI, HAR, LAN, NYC, NEW, POR, SDI, SFR, SJO, SEA, WAS	ATL, BIR, CLE, COL, DET, HOU, MIA, PRO	BAL, BUF, CIN, DAL, DEN, FTW, IND, KAN, MEM, MIL, MIN, NWO, NOR, OKL, PHI, PHO, PIT, ROC, STL, SAL, SAN, SBE, TAM
Missing	ATL, BAL, CIN, CLE, DAL, DEN, FTW, HAR, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NOR, OKL, PHO, PIT, POR, PRO, ROC, STL, SAN, SDI, SJO, SEA, WAS	SBE	ANA, BIR, BOS, BUF, CHI, COL, DET, NYC, NEW, PHI, SAL, SFR, TAM

Exhibit 10

Fannie Mae Relative Share of GSE Market by Ratio of Tract to MSA Median Income

Tract/MSA Income Ratio	Segment Market Share < Overall Market Share	Segment Market Share = Overall Market Share	Segment Market Share > Overall Market Share
0–60	FTW, PRO		ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, HAR, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, ROC, STL, SAL, SAN, SBE, SDI, SFR, SJO, SEA, TAM, WAS
60–80	STL, SAN	DET, MIA	ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, FTW, HAR, HOU, IND, KAN, LAN, MEM, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, SAL, SBE, SDI, SFR, SJO, SEA, TAM, WAS
80–100	BIR, COL, MEM, NOR, OKL, SAN, SJO	CLE, FTW, LAN, NWO, PIT, ROC, STL, WAS	ANA, ATL, BAL, BOS, BUF, CHI, CIN, DAL, DEN, DET, HAR, HOU, IND, KAN, MIA, MIL, MIN, NYC, NEW, PHI, PHO, POR, PRO, SAL, SBE, SDI, SFR, SEA, TAM
100–120	BIR, BOS, CIN, DEN, IND, POR, STL, SBE, SDI, SFR, SJO	ANA, ATL, BAL, BUF, HAR, LAN, MIL, MIN, NWO, NEW, PIT, PRO, ROC, SAL, SEA, TAM, WAS	CHI, CLE, COL, DAL, DET, FTW, HOU, KAN, MEM, MIA, NYC, NOR, OKL, PHI, PHO, SAN
>120	ANA, BAL, BUF, CHI, CIN, CLE, DAL, DEN, DET, FTW, HAR, HOU, IND, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, OKL, PHI, PHO, POR, SAL, SAN, SBE, SFR, SJO, SEA, WAS	ATL, BOS, KAN, NOR, PIT, PRO, ROC, SDI, TAM	BIR, COL, STL
Missing	ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, SFR, SJO, SEA, TAM, WAS		

Exhibit 11

Fannie Mae Relative Share of Total Market by Percentage Minority of Tract Population

Tract/MSA Income Ratio	Segment Market Share < Overall Market Share	Segment Market Share = Overall Market Share	Segment Market Share > Overall Market Share
0–10	ANA, DAL, HOU, LAN, MIA, POR, SBE, SDI, SFR, SJO, SEA, TAM, WAS	BAL, BOS, BUF, CHI, CIN, CLE, DET, HAR, MIN, NYC, NEW, PHI, PHO, PRO	ATL, BIR, COL, DEN, FTW, IND, KAN, MEM, MIL, NWO, NOR, OKL, PIT, ROC, STL, SAL, SAN
10–15	ANA, ATL, BIR, CIN, CLE, DET, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, PRO, STL, SAL, SBE, SDI, SFR, SJO, WAS	BAL, BOS, BUF, COL, FTW, HOU, PIT, ROC	CHI, DAL, DEN, HAR, NOR, OKL, PHI, PHO, POR, SAN, SEA, TAM
15–30	ATL, BUF, CIN, CLE, COL, DEN, DET, FTW, KAN, LAN, MIA, MIL, MIN, NWO, NYC, OKL, PIT, PRO, ROC, STL, SAL, SFR, SJO	ANA, BIR, HAR, IND, NOR, PHI, PHO	BAL, BOS, CHI, DAL, HOU, MEM, NEW, POR, SAN, SBE, SDI, SEA, TAM, WAS
30–50	ATL, BIR, BUF, CHI, CIN, CLE, COL, DEN, DET, FTW, IND, KAN, MEM, MIA, MIL, MIN, NWO, NOR, OKL, PHO, PIT, PRO, ROC, STL, SAL, SAN, TAM	BAL, PHI	ANA, BOS, DAL, HAR, HOU, LAN, NYC, NEW, POR, SBE, SDI, SFR, SJO, SEA, WAS
>50	ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, MEM, MIL, MIN, NWO, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, TAM	WAS	ANA, LAN, MIA, NYC, SDI, SFR, SJO, SEA
Missing	ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, SFR, SJO, SEA, TAM, WAS		HAR

whereas Fannie Mae's relative market share was low (and Freddie Mac's high) in only 3 MSAs. Conversely, in predominantly White tracts, Fannie Mae's relative market share was low (and Freddie Mac's high) in 30 MSAs, while Fannie Mae's relative market share was high (and Freddie Mac's low) in only a single MSA. Once again, however, the implication that Fannie Mae is financing loans in predominantly minority neighborhoods more readily than Freddie Mac must be tempered by exhibits 11 and 12, which show that both GSEs appear to lag behind non-GSE market participants in funding loans to borrowers in high-minority areas.

Exhibit 12

Freddie Mac Relative Share of Total Market by Percentage Minority of Tract Population

Tract Percent Minority	Segment Market Share < Overall Market Share	Segment Market Share = Overall Market Share	Segment Market Share > Overall Market Share
0–10	ANA, HOU, LAN, SBE, SDI, SFR, SJO	BIR, BOS, CHI, CLE, COL, DAL, HAR, IND, MIA, MIN, PHO, PIT, POR, PRO, SEA, WAS	ATL, BAL, BUF, CIN, DEN, DET, FTW, KAN, MEM, MIL, NWO, NYC, NEW, NOR, OKL, PHI, ROC, STL, SAL, SAN, TAM
10–15	ANA, ATL, BOS, BUF, CIN, CLE, COL, DET, HAR, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, OKL, PIT, PRO, ROC, STL, SAL, SFR, SJO, WAS	BAL, BIR, CHI, FTW, PHI, POR, SBE, SDI, TAM	DAL, DEN, HOU, NOR, PHO, SAN, SEA
15–30	ATL, BOS, BUF, CIN, CLE, COL, DEN, DET, FTW, HAR, IND, KAN, MIA, MIL, MIN, NWO, NYC, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SFR, SJO	BAL, BIR, CHI, DAL, LAN, NOR, TAM, WAS	ANA, HOU, MEM, NEW, SAN, SBE, SDI, SEA
30–50	ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, MEM, MIA, MIL, MIN, NWO, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SEA, TAM	NEW, SBE	ANA, LAN, NYC, SDI, SFR, SJO, WAS
>50	ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, MEM, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, TAM, WAS	SEA	ANA, LAN, MIA, SFR, SJO
Missing	ANA, ATL, CLE, DAL, FTW, HOU, IND, KAN, LAN, MEM, MIA, MIN, NWO, NYC, NOR, OKL, POR, ROC, STL, SAN, SDI, SJO, SEA, WAS	DET	BAL, BIR, BOS, BUF, CHI, CIN, COL, DEN, HAR, MIL, NEW, PHI, PHO, PIT, PRO, SAL, SBE, SFR, TAM

In summary, both GSEs appear to have relatively low market shares in mixed or predominantly minority census tracts and relatively high market shares in predominantly White census tracts, although this pattern appears to be more pronounced for Freddie Mac than for Fannie Mae.

Central-City Versus Suburban Tracts

Exhibits 14 and 15 summarize results concerning the relative market shares of Fannie Mae and Freddie Mac according to whether borrowers reside in census tracts located in

Exhibit 13

Fannie Mae Relative Share of GSE Market by Percentage Minority of Tract Population

Tract Percent Minority	Segment Market Share < Overall Market Share	Segment Market Share = Overall Market Share	Segment Market Share > Overall Market Share
0–10	ANA, BAL, BOS, CHI, CIN, CLE, DAL, FTW, HAR, HOU, IND, LAN, MIA, MIL, NWO, NYC, NEW, OKL, PHI, PHO, POR, SAL, SAN, SBE, SDI, SFR, SJO, SEA, TAM, WAS	ATL, BUF, COL, DEN, DET, KAN, MEM, MIN, NOR, PIT, PRO, ROC, STL	BIR
10–15	BIR, DEN, DET, HOU, LAN, MIA, NYC, NOR, SAN, SBE, SDI, SFR, SJO	ANA, BAL, CHI, CLE, DAL, NEW, PHO	ATL, BOS, BUF, CIN, COL, FTW, HAR, IND, KAN, MEM, MIL, MIN, NWO, OKL, PHI, PIT, POR, PRO, ROC, STL, SAL, SEA, TAM, WAS
15–30	LAN, MEM, MIA, NYC, STL, SFR, SJO	ANA, BIR, DEN, HOU, NWO, SAN, SBE, SDI, WAS	ATL, BAL, BOS, BUF, CHI, CIN, CLE, COL, DAL, DET, FTW, HAR, IND, KAN, MIL, MIN, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, SAL, SEA, TAM
30–50	LAN, MEM, MIA, NWO, PHO, WAS	ATL, NOR, SJO	ANA, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, MIL, MIN, NYC, NEW, OKL, PHI, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, SFR, SEA, TAM
>50	FTW, OKL, PRO	SAN	ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, HAR, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, NOR, PHI, PHO, PIT, POR, ROC, STL, SAL, SBE, SDI, SFR, SJO, SEA, TAM, WAS
Missing	ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, SFR, SJO, SEA, TAM, WAS		

central-city or suburban areas. As these exhibits show, both GSEs appear to finance loans to borrowers located in suburban areas to a greater extent than borrowers located in central-city areas, with the pattern more striking for Freddie Mac (exhibit 15) than for Fannie Mae. Exhibit 16 shows Fannie Mae's relative share of the market for GSE loans only and suggests that Fannie Mae is more active than Freddie Mac in purchasing loans

Exhibit 14

Fannie Mae Relative Share of Total Market by Central-City Versus Suburban Location

Tract Location	Segment Market Share < Overall Market Share	Segment Market Share = Overall Market Share	Segment Market Share > Overall Market Share
Suburbs	ANA, NYC, PHO, POR, SAN, SDI, SJO	ATL, BAL, CIN, COL, DEN, DET, FTW, IND, MIN, NEW, OKL, STL, SFR, SEA, TAM	BIR, BOS, BUF, CHI, CLE, DAL, HAR, HOU, KAN, LAN, MEM, MIA, MIL, NWO, NOR, PHI, PIT, PRO, ROC, SAL, SBE, WAS
Central City	ATL, BAL, BIR, BUF, CHI, CIN, CLE, DAL, DEN, DET, FTW, HAR, HOU, KAN, LAN, MEM, MIA, MIN, NWO, NEW, NOR, PIT, PRO, ROC, STL, SAL, TAM	IND, MIL, NYC, PHO, SDI, SEA, WAS	ANA, BOS, COL, OKL, PHI, POR, SAN, SBE, SFR, SJO
Missing	ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NOR, OKL, PHI, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, SFR, SJO, SEA, TAM, WAS	PHO	ANA, NEW

Exhibit 15

Freddie Mac Relative Share of Total Market by Central-City Versus Suburban Location

Tract Location	Segment Market Share < Overall Market Share	Segment Market Share = Overall Market Share	Segment Market Share > Overall Market Share
Suburbs	ANA, OKL, SJO	ATL, BAL, BIR, BOS, CLE, COL, DET, HAR, HOU, IND, MIA, NWO, NYC, NEW, PHI, PHO, PIT, POR, PRO, SAN, SFR, WAS	BUF, CHI, CIN, DAL, DEN, FTW, KAN, LAN, MEM, MIL, MIN, NOR, ROC, STL, SAL, SBE, SDI, SEA, TAM
Central City	ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, NOR, PHI, PIT, POR, PRO, ROC, STL, SAL, SBE, SEA, TAM, WAS	OKL, PHO, SAN, SDI, SFR	ANA, SJO
Missing	ATL, BAL, BIR, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, HAR, HOU, IND, KAN, LAN, MIA, MIL, NWO, OKL, POR, STL, SBE, SDI, SJO, TAM, WAS	BOS, MIN, PRO, SAL, SAN	ANA, FTW, MEM, NYC, NEW, NOR, PHI, PHO, PIT, ROC, SFR, SEA

Exhibit 16

Fannie Mae Relative Share of GSE Market by Central-City Versus Suburban Location

Tract Location	Segment Market Share < Overall Market Share	Segment Market Share = Overall Market Share	Segment Market Share > Overall Market Share
Suburbs	ANA, BAL, CIN, COL, DAL, DEN, FTW, HOU, IND, LAN, MEM, MIL, NWO, NYC, POR, SAN, SDI, SFR, SJO, TAM	ATL, BIR, BOS, BUF, CHI, CLE, DET, HAR, KAN, MIA, MIN, NEW, NOR, OKL, PHI, PHO, ROC, STL, SAL, SBE, SEA, WAS	PIT, PRO
Central City	PHO	CHI, MIA, NWO, OKL, SAL, SJO, WAS	ANA, ATL, BAL, BIR, BOS, BUF, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MEM, MIL, MIN, NYC, NEW, NOR, PHI, PIT, POR, PRO, ROC, STL, SAN, SBE, SDI, SFR, SEA, TAM
Missing	ANA, BAL, BIR, BOS, CHI, CLE, COL, DEN, DET, FTW, HAR, HOU, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, SFR, SJO, SEA, TAM	BUF, IND	ATL, CIN, DAL, WAS

made to borrowers in central-city areas. In summary, both GSEs appear to have relatively low market shares in central-city census tracts and relatively high market shares in suburban census tracts, although the pattern is more pronounced for Freddie Mac than for Fannie Mae.

Geographically Targeted Census Tracts

Exhibits 17 and 18 summarize market-share results according to whether borrowers reside in census tracts that are *geographically targeted* according to the criteria established in the GSE housing goals legislation.[17] As before, both exhibits show a similar pattern in which the GSEs appear to be funding a smaller percentage of loans to borrowers in geographically targeted census tracts than to borrowers in nontargeted tracts, but the pattern is more striking for Freddie Mac (exhibit 18) than for Fannie Mae. Exhibit 19 summarizes Fannie Mae's relative share of the GSE-only market and suggests very strongly that Fannie Mae is more active than Freddie Mac in serving borrowers in geographically targeted census tracts. In summary, both GSEs appear to have relatively low market shares in geographically targeted census tracts and relatively high market shares in nontargeted census tracts, although the pattern is more pronounced for Freddie Mac than for Fannie Mae.

Cross-Tabulations Using HMDA Data Only

As noted in the first section of this article, the HMDA data on which several previous studies of GSE mortgage purchase activity were based may not encompass the universe of mortgage originations. Berkovec and Zorn (1996) found that "loans in lower income

Exhibit 17

Fannie Mae Relative Share of Total Market by Geographically Targeted Census Tract

Whether Targeted	Segment Market Share < Overall Market Share	Segment Market Share = Overall Market Share	Segment Market Share > Overall Market Share
Nontargeted	ANA, LAN, MIA, NYC, POR, SDI, SFR, SJO	ATL, BOS, DAL, DET, FTW, HAR, HOU, MIN, NWO, NEW, PHO, PRO, SEA, WAS	BAL, BIR, BUF, CHI, CIN, CLE, COL, DEN, IND, KAN, MEM, MIL, NOR, OKL, PHI, PIT, ROC, STL, SAL, SAN, SBE, TAM
Targeted	ATL, BAL, BIR, BUF, CHI, CIN, CLE, COL, DEN, DET, FTW, HOU, IND, KAN, MEM, MIL, MIN, NWO, NOR, OKL, PHI, PHO, PIT, PRO, ROC, STL, SAL, SAN, SBE, TAM	DAL, HAR	ANA, BOS, LAN, MIA, NYC, NEW, POR, SDI, SFR, SJO, SEA, WAS
Missing	ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, SFR, SJO, SEA, TAM, WAS		ANA

Exhibit 18

Freddie Mac Relative Share of Total Market by Geographically Targeted Census Tract

Whether Targeted	Segment Market Share < Overall Market Share	Segment Market Share = Overall Market Share	Segment Market Share > Overall Market Share
Nontargeted	ANA, LAN, SFR, SJO, WAS	ATL, BIR, BOS, CHI, CLE, COL, DET, FTW, HAR, HOU, IND, MIA, NWO, NYC, NEW, PIT, POR, PRO, SDI, SEA	BAL, BUF, CIN, DAL, DEN, KAN, MEM, MIL, MIN, NOR, OKL, PHI, PHO, ROC, STL, SAL, SAN, SBE, TAM
Targeted	ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, TAM		ANA, LAN, SDI, SFR, SJO, SEA, WAS
Missing	ATL, BAL, CIN, CLE, DAL, DEN, FTW, HAR, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NOR, OKL, PHO, PIT, POR, PRO, ROC, STL, SAN, SDI, SJO, SEA, WAS	DET, SBE	ANA, BIR, BOS, BUF, CHI, COL, NYC, NEW, PHI, SAL, SFR, TAM

Exhibit 19

Fannie Mae Relative Share of GSE Market by Geographically Targeted Census Tract

Whether Targeted	Segment Market Share < Overall Market Share	Segment Market Share = Overall Market Share	Segment Market Share > Overall Market Share
Nontargeted	ANA, BAL, BOS, CIN, DAL, DEN, HAR, HOU, IND, LAN, MEM, MIA, MIL, NWO, NYC, OKL, PHO, POR, SAL, SAN, SFR, SJO	ATL, BUF, CHI, CLE, COL, DET, FTW, KAN, MIN, NEW, NOR, PHI, PIT, PRO, ROC, STL, SBE, SDI, SEA, TAM, WAS	BIR
Targeted	BIR, COL	NOR, STL	ANA, ATL, BAL, BOS, BUF, CHI, CIN, CLE, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, OKL, PHI, PHO, PIT, POR, PRO, ROC, SAL, SAN, SBE, SDI, SFR, SJO, SEA, TAM, WAS
Missing	ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HOU, IND, KAN, LAN, MIA, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, SFR, SJO, SEA, TAM, WAS	MEM	HAR

tracts are overrepresented in the HMDA data," so that "HMDA data may overstate the share of total mortgage market activity that occurs in low-income areas." Because there is no better source than HMDA for the non-GSE portion of the market, however, our combined database may suffer from underreporting of mortgages that are not purchased by the GSEs. If Berkovec and Zorn are correct in their evaluation of the biases in HMDA reporting, this then suggests that the non-GSE loans most likely to remain unreported in the combined PUDB-HMDA data set are those that originated in higher income neighborhoods. This implies that the market share of the GSEs computed from our combined PUDB-HMDA data set may be biased upward in higher income census tracts: that is, the actual market shares of Fannie Mae and Freddie Mac in higher income neighborhoods may not be as high as our PUDB-HMDA data set suggests. This is because GSE mortgage purchases in all types of neighborhoods should be represented completely in the PUDB data, while non-GSE originations, particularly in higher income neighborhoods, may be underreported in HMDA data.

It is possible, then, that the overall pattern of results derived from our cross-tabulations, in which the GSEs appear to finance a systematically larger share of mortgages to higher income borrowers and to White borrowers in higher income, predominantly White, suburban, and nontargeted census tracts, may arise in part because of our data set construction. In particular, we have supplemented the HMDA data by appending all loans purchased by the GSEs, including those higher income loans that Berkovec and Zorn conclude may have been underreported in the raw HMDA data set. On the other hand, we are unable to append loans that were not purchased by the GSEs; therefore, higher income loans in the non-GSE portion of the market may remain underreported.

Exhibit 20

GSE Market Shares by Borrower Income

MSA	Market Segment	Market Share: HMDA Data Only	Market Share: Combined Data	Ratio
Anaheim	Income < Median	34.3	46.7	1.36 ↑
	Income ≥ Median	24.5	30.0	1.22
Atlanta	Income < Median	23.3	37.4	1.61 ↑
	Income ≥ Median	28.8	42.2	1.47
Baltimore	Income < Median	20.7	30.4	1.47 ↑
	Income ≥ Median	28.4	38.5	1.35
Birmingham	Income < Median	19.5	31.1	1.59 ↑
	Income ≥ Median	32.0	43.7	1.37
Boston	Income < Median	34.5	46.4	1.34 ↑
	Income ≥ Median	37.0	41.1	1.11
Buffalo	Income < Median	29.7	37.4	1.26 ↑
	Income ≥ Median	40.3	50.2	1.25
Chicago	Income < Median	26.4	40.0	1.52 ↑
	Income ≥ Median	28.1	40.9	1.46
Cincinnati	Income < Median	24.6	34.8	1.41 ↑
	Income ≥ Median	31.2	42.8	1.37
Cleveland	Income < Median	19.3	29.2	1.51 ↑
	Income ≥ Median	21.7	32.3	1.49
Columbus	Income < Median	26.0	38.3	1.47 ↑
	Income ≥ Median	36.3	48.2	1.33
Dallas	Income < Median	23.5	35.1	1.49 ↑
	Income ≥ Median	32.0	42.7	1.33
Denver	Income < Median	26.4	37.1	1.41 ↑
	Income ≥ Median	36.4	47.1	1.30
Detroit	Income < Median	35.7	45.6	1.28 ↑
	Income ≥ Median	40.4	47.3	1.17
Fort Worth	Income < Median	18.9	29.3	1.55 ↑
	Income ≥ Median	31.7	42.9	1.35
Hartford	Income < Median	24.0	35.3	1.47 ↑
	Income ≥ Median	29.0	40.2	1.38
Houston	Income < Median	32.3	42.6	1.32 ↑
	Income ≥ Median	36.6	48.1	1.31
Indianapolis	Income < Median	22.8	35.8	1.57 ↑
	Income ≥ Median	33.2	45.2	1.36
Kansas City	Income < Median	22.6	34.0	1.51 ↑
	Income ≥ Median	32.5	44.9	1.38
Los Angeles	Income < Median	29.4	40.3	1.37 ↑
	Income ≥ Median	23.4	30.6	1.31
Memphis	Income < Median	14.7	15.5	1.05
	Income ≥ Median	29.7	35.2	1.19 ↓
Miami	Income < Median	22.7	35.9	1.59 =
	Income ≥ Median	24.7	39.3	1.59
Milwaukee	Income < Median	31.4	40.7	1.29
	Income ≥ Median	36.6	49.0	1.34 ↓
Minneapolis	Income < Median	25.4	38.1	1.50 ↑
	Income ≥ Median	37.2	48.9	1.32

Exhibit 20 (continued)

GSE Market Shares by Borrower Income

MSA	Market Segment	Market Share: HMDA Data Only	Market Share: Combined Data	Ratio
New Orleans	Income < Median	22.9	29.3	1.28
	Income ≥ Median	29.8	39.7	1.33 ↓
New York City	Income < Median	27.8	36.3	1.31
	Income ≥ Median	26.3	36.6	1.39 ↓
Newark	Income < Median	32.1	39.0	1.21
	Income ≥ Median	26.9	34.8	1.29 ↓
Norfolk	Income < Median	9.7	14.9	1.53 ↑
	Income ≥ Median	21.5	30.7	1.43
Oklahoma City	Income < Median	17.9	25.1	1.41 ↑
	Income ≥ Median	33.3	43.6	1.31
Philadelphia	Income < Median	31.2	38.4	1.23
	Income ≥ Median	35.5	45.1	1.27 ↓
Phoenix	Income < Median	23.5	34.0	1.45 ↑
	Income ≥ Median	32.8	42.3	1.29
Pittsburgh	Income < Median	14.4	20.9	1.46
	Income ≥ Median	22.4	32.9	1.47 ↓
Portland	Income < Median	34.3	47.4	1.38 ↑
	Income ≥ Median	34.5	47.0	1.36
Providence	Income < Median	24.2	36.7	1.52 ↑
	Income ≥ Median	34.0	46.6	1.37
Rochester	Income < Median	29.5	38.6	1.31 ↑
	Income ≥ Median	41.6	52.0	1.25
St. Louis	Income < Median	25.8	35.1	1.36 ↑
	Income ≥ Median	35.2	43.8	1.24
Salt Lake City	Income < Median	27.5	37.3	1.36 ↑
	Income ≥ Median	37.4	47.4	1.27
San Antonio	Income < Median	14.9	19.8	1.33
	Income ≥ Median	28.4	39.4	1.39 ↓
San Bernardino	Income < Median	19.3	32.3	1.67 ↑
	Income ≥ Median	28.4	40.3	1.42
San Diego	Income < Median	30.2	42.3	1.40 ↑
	Income ≥ Median	28.4	35.4	1.25
San Francisco	Income < Median	38.6	48.5	1.26 ↑
	Income ≥ Median	15.3	18.7	1.22
San Jose	Income < Median	43.4	54.3	1.25 ↑
	Income ≥ Median	22.4	27.1	1.21
Seattle	Income < Median	32.3	46.0	1.43 ↑
	Income ≥ Median	30.5	42.2	1.38
Tampa Bay	Income < Median	21.4	32.1	1.50 ↑
	Income ≥ Median	28.1	41.0	1.46
Washington	Income < Median	24.7	35.7	1.45 ↑
	Income ≥ Median	25.1	33.0	1.31

In order to evaluate whether the HMDA data are biased in the manner detected by Berkovec and Zorn (1996), we compute GSE market shares for MSAs in two market segments: borrowers with incomes below the area median and borrowers with incomes at or above the area median. The results of this analysis are presented in exhibit 20. As this exhibit shows, the combined market share of both GSEs is estimated to be higher using the combined PUDB-HMDA data set compared with the HMDA-only data set. This is to be expected because the combined PUDB-HMDA data set includes the universe of loans purchased by the GSEs whereas the HMDA database is not designed to provide universal coverage.

More important, however, the difference between the HMDA-only and combined PUDB-HMDA data in estimated GSE market shares is slightly greater among lower income borrowers than among higher income borrowers in 35 of the 44 MSAs studied (and the same in Miami), the exceptions being Memphis, Milwaukee, Newark, New Orleans, New York, Philadelphia, Pittsburgh, and San Antonio. This means that for the 1993–96 time period covered by our analysis, using the HMDA database seems to result in reporting a lower GSE share of the underserved market in most of the cities studied.

To further evaluate the differences in market share for the GSEs between the combined PUDB-HMDA data set and the raw HMDA data set, we replicate the cross-tabulations discussed above using just the HMDA data set for borrower income, and also, for all variables of policy interest, for a subset of 6 of the 44 MSAs (Cleveland, Miami, Newark, Philadelphia, St. Louis, and Seattle). We then compare the market shares computed for Fannie Mae and Freddie Mac in different market segments to determine whether there was any systematic difference between the two data sets. Of course, in all six MSAs the GSE market shares estimated using just the HMDA data are smaller than the GSE market shares estimated using the combined PUDB-HMDA data set, because the combined data set includes all GSE mortgage purchases while the HMDA data set does not have universal coverage. An important question, however, is whether the observed GSE mortgage purchase patterns can be explained by selective coverage of mortgage lending in different market segments.

In general, the estimated market shares using HMDA-only data suggest that the GSEs provide a lower proportion of funding for mortgage lending to minority borrowers than is suggested by the combined PUDB-HMDA data set. In Newark, for example, the combined PUDB-HMDA data set indicates that Fannie Mae's market share is essentially the same (19 percent) among White and minority borrowers. Using the HMDA data set alone, Fannie Mae's market share in Cleveland is estimated to be slightly higher among White borrowers (16 percent) than among minority borrowers (15 percent). Freddie Mac's market shares show a similar pattern: the combined PUDB-HMDA data set suggests that Freddie Mac's market share among White borrowers (16 percent) is slightly higher than among minority borrowers (15 percent), but the HMDA-only data suggest that the difference is greater (12 percent versus 9 percent). The other five MSAs investigated show similar differences in cross-tabulations between the combined PUDB-HMDA data set and the raw HMDA data: in general, the pattern identified using the combined PUDB-HMDA data set (higher market shares for the GSEs among White than among minority borrowers) appears more pronounced using the raw HMDA data set.

The HMDA-only data suggest that the GSEs provide a lower proportion of funding for mortgage lending in geographically targeted areas than is suggested by the combined PUDB-HMDA data set. In Newark, for example, the combined PUDB-HMDA data set indicates that Fannie Mae's market share is slightly higher in targeted areas (21 percent) than in nontargeted areas (18 percent). The HMDA-only data, however, indicate that the

pattern is reversed: that is, Fannie Mae's market share is slightly higher in nontargeted areas (15 percent) than in targeted areas (14 percent). Freddie Mac's market shares show a similar difference: the combined PUDB-HMDA data set suggests that Freddie Mac's market share in nontargeted areas (15 percent) is slightly higher than in targeted areas (14 percent), but the HMDA-only data suggest that the difference is greater (11 percent versus 9 percent). Again, the five other MSAs investigated show a similar difference between the cross-tabulation results using the combined PUDB-HMDA data set and the raw HMDA data set. In general, the GSEs appear to provide a higher proportion of mortgage funding to borrowers in nontargeted than in targeted areas, but this pattern appears more pronounced using the raw HMDA data set.

For market segments defined by borrower income category, the differences between the combined PUDB-HMDA data set and the raw HMDA data are similar. The overall pattern identified from cross-tabulations using the combined PUDB-HMDA data set is that GSE market shares are greater among higher income borrowers than among lower income borrowers, and it appears somewhat more pronounced using the raw HMDA data set. In Newark, for example, the combined PUDB-HMDA data set suggests that Fannie Mae's market share was higher among borrowers whose incomes were between 50 and 60 percent of area median (20 percent) than among those whose incomes were between 250 and 300 percent of area median (11 percent). The HMDA-only data set, however, suggests that Fannie Mae's market shares in the two market segments are more similar: 14 percent among the low-income borrowers compared to 10 percent among the high-income borrowers. Freddie Mac's market shares show a similar pattern. The combined PUDB-HMDA data set suggests that it is higher among the low-income borrowers (13 percent) than among the high-income borrowers (10 percent). The HMDA-only data set, however, suggests that Freddie Mac's market shares in the two market segments are more similar: 9 percent among the low-income borrowers compared to 8 percent among the high-income borrowers. Again, the five other MSAs investigated show similar differences between the cross-tabulation results using the combined PUDB-HMDA data set and the raw HMDA data set. In general, the GSEs appear to provide a higher proportion of mortgage funding to borrowers in nontargeted than in targeted areas, but this pattern appears more pronounced using the raw HMDA data set.[18] In short, then, the analysis of GSE market shares prompts generally the same conclusions regardless of whether the analysis is conducted using the HMDA-only data set or the combined PUDB-HMDA data set that we constructed for this analysis.

Logistic Regressions on GSE Loan Purchase Activity

While the statistical cross-tabulations presented in the previous section reveal a strong pattern in which both GSEs appear to be providing less mortgage funding than non-GSE leaders to particular segments of the mortgage market (including low-income and minority borrowers in low-income, predominantly minority, central-city, and geographically targeted areas), the cross-tabulation procedure has one significant drawback: it enables us to control for only one characteristic at a time in evaluating loan purchase activity. Since the characteristics of interest tend to be highly correlated—that is, low-income borrowers tend to live in low-income, central-city, geographically targeted areas—the cross-tabulations might make it appear that all characteristics were important even if, as is possible, only one or two of them are truly determinative.

In this section we describe our use of logistic multiple regressions to evaluate spatial variation in GSE mortgage purchase activity in greater detail. Specifically, using data for all loans recorded in the merged PUDB-HMDA database in each of the 44 large MSAs,

we estimated a logistic regression equation describing the probability that a given loan was purchased by either of the GSEs:[19]

$$Prob(Y_{kt}=1 \mid X_{kt}) = f(X_{kt}, \varepsilon_{kt}),$$

where Y_{kt} =1 indicates that loan pool k (k = 1,...,N_k) recorded in the HMDA data set for year t (t = 1993,...,1996) was purchased by a GSE, while Y_{kt} = 0 indicates that loan pool k was not purchased by a GSE. X_{kt} represents a series of characteristics for that loan pool (including borrower, census tract, loan, and lender characteristics), and ε_{kt} is the disturbance term. This logistic regression is analogous to the cross-tabulations shown in the "All Purchasers" summary reports in exhibit A–1, and enables us to evaluate the performance of the GSE portion of the mortgage market as a whole by comparing it to the performance of the non-GSE portion of the mortgage market in purchasing loans in market segments characterized by different borrower, loan, lender, or neighborhood attributes.

It is important to keep in mind that the loans included in the PUDB and HMDA databases are not identified in any way that makes it possible to match individual loans across the databases. For this reason the observations in our merged PUDB-HMDA database are actually *pools* of loans sharing identical values for the eight common variables (State, county, MSA, census tract, origination/acquisition year, GSE purchaser, borrower income category, and borrower race) used in the matching and merging process. The dependent variable in each of the logit regression equations is a binary value indicating IF_GSE = 1 if pool was purchased by a GSE and IF_GSE = 0 otherwise.[20]

The independent variables included in the logistic regressions were those that we focused on in our cross-tabulations, along with a few additional variables reflecting additional borrower, neighborhood, and lender characteristics:

- **Borrower Income Category.** Because of the degree of collinearity among these variables, we collapsed the number of categories to five, as follows:

  BINC_C1X incomes less than 70 percent of MSA median.

  BINC_C2X incomes between 70 and 100 percent of MSA median.

  BINC_C3X incomes between 100 and 130 percent of MSA median.

  BINC_C4X incomes between 130 and 200 percent of MSA median.

  BINC_C5X incomes greater than 200 percent of MSA median (excluded category).

- **Borrower Race (BR_RACE).** 0 = White, 1 = minority.

- **Tract Income Category.** Same categories as in the cross-tabulations:

  TINC_C01 tract median incomes less than 60 percent of MSA median.

  TINC_C02 tract median incomes between 60 and 80 percent of MSA median.

  TINC_C03 tract median incomes between 80 and 100 percent of MSA median.

  TINC_C04 tract median incomes between 100 and 120 percent of MSA median.

  TINC_C05 tract median incomes greater than 120 percent of MSA median (excluded category).

- **Tract Percent Minority.** Same categories as in the cross-tabulations:

  TMIN_C01 tract minority concentration less than 10 percent (excluded category).

  TMIN_C02 tract minority concentration between 10 and 15 percent.

  TMIN_C03 tract minority concentration between 15 and 30 percent.

  TMIN_C04 tract minority concentration between 30 and 50 percent.

  TMIN_C05 tract minority concentration greater than 50 percent.

- **Central-City Versus Suburban Location (CNTRCITY).** Percentage of the census tract located in central-city areas of the MSA.
- **Geographically Targeted (GEOTARGT).** 1=targeted, 0=not targeted.[21]
- **Loan Balance Category.** Again, because of the degree of collinearity among these variables, we collapsed the number of categories to five, as follows:

  UPB_C1X average loan balance in lowest 20 percent of MSA distribution.

  UPB_C2X average loan balance between 20th and 40th percentile.

  UPB_C3X average loan balance between 40th and 60th percentile.

  UPB_C4X average loan balance between 60th and 80th percentile.

  UPB_C5X average loan balance in highest 20 percent of MSA distribution (excluded category).

- **Percent Female (PtPlFmL2).** Percentage of loans in the pool for which the borrower was identified as female.
- **Percent Owner-Occupied (PctOwnOc).** Percentage of properties in the census tract that are owner-occupied.
- **Percent Home Purchase (Pct_HmPr).** Percentage of loans in the pool that are for home purchase rather than refinance.

We also included as regressors variables indicating the percentage of loans in each pool identified in the HMDA data set as being of each loan type (conventional or Federal Housing Administration versus Farmers Home Administration/Veterans Administration) and the percentage identified in the HMDA data set as being originated by lenders regulated by each regulating agency (Federal Deposit Insurance Corporation, Federal Reserve Board, National Credit Union Administration, Office of the Comptroller of the Currency, or HUD versus Office of Thrift Supervision).

Exhibit 21 summarizes the empirical results for the logistic regressions conducted separately for each of the 44 MSAs included in our analysis, showing in particular the median, mean, and standard deviation of the regression coefficients estimated over the full group of 44 MSAs studied.[22] While these median and mean coefficients may mask significant variation in the coefficients estimated for individual MSAs, overall they help to reveal a pattern quite similar to the pattern observed in the cross-tabulations presented in this article's final section.

Exhibit 21

Summary of GSE/Non-GSE Logistic Regression Results

Independent Variable	Median Estimated Regression Coefficient	Average Estimated Regression Coefficient	Standard Deviation of Estimated Coefficients
BINC_C1X	0.115	0.061	0.769
BINC_C2X	0.960	0.900	0.610
BINC_C3X	1.047	1.023	0.412
BINC_C4X	1.016	1.077	0.313
BR_RACE	–0.914	–0.790	0.576
TINC_C01	–1.101	–1.110	0.719
TINC_C02	–0.267	–0.342	0.557
TINC_C03	0.163	0.046	0.369
TINC_C04	0.246	0.204	0.268
TMIN_C02	0.152	0.112	0.421
TMIN_C03	0.109	0.199	0.401
TMIN_C04	0.124	0.239	0.654
TMIN_C05	–0.190	–0.085	0.831
CNTRCITY	–0.261	–0.218	0.367
GEOTARGT	–0.130	–0.174	0.526
UPB_C1X	0.479	1.045	1.519
UPB_C2X	1.019	1.301	1.240
UPB_C3X	1.254	1.481	1.059
UPB_C4X	1.249	1.324	0.676
PtPlFmL2	–0.132	–0.150	0.195
PctOwnOc	1.940	2.331	1.965
Pct_HmPr	–0.471	–0.446	0.733
Pct_Cvtl	36.64	40.62	21.03
Pct_FHA	11.62	13.67	21.19
Pct_FDIC	–0.795	–0.758	2.633
Pct_FRB	2.552	2.934	2.309
Pct_HUD	2.572	2.645	2.210
Pct_NCUA	–11.14	–20.49	27.74
Pct_OCC	0.774	1.118	2.283
Predictive Power:			
Concordant	85.7%	85.2%	5.2%
Somers' *D*	0.719	0.710	0.107

Exhibit 21 also presents two summary measures of the predictive ability of the regression model, "Percent Concordant" and "Somers' *D*." The first measure reflects the percentage of pairs of observations with different outcomes—that is, one pool was purchased by a GSE and the other pool was not—for which the pool purchased by a GSE had a higher predicted probability of that outcome than did the other pool in the pair. As exhibit 21 shows, the median percent concordant for the 44 MSAs was 85.7 percent, indicating that

the regression models had very high predictive power. The second measure, "Somers' *D,*" is an index of the correlation between observed and predicted outcomes and is sometimes thought of as a measure for logistic regressions analogous to the R-squared measure of goodness-of-fit for ordinary least squares regressions.[23] As exhibit 21 shows, the median Somers' *D* for the 44 MSAs was 0.719, again indicating high predictive power for the regression models.[24]

Borrower Income Category

The median regression coefficients for the middle three borrower income categories (BINC_C2X, BINC_C3X, and BINC_C4X) are quite similar and strongly positive (median values of 0.960, 1.047, and 1.016), indicating that loans to borrowers with incomes between 70 percent and 200 percent of MSA median are more likely to be purchased by the GSEs than are loans to the highest income borrowers (the excluded category). On the other hand, the median coefficient on the lowest income borrower income category (BINC_C1X = 0.115) is smaller, suggesting that there is little difference between the lowest income and the highest income borrowers in terms of the probability that loans made to these borrowers would be purchased by either of the GSEs. This is not unexpected, since a large share of loans to borrowers in the higher income category can be expected to exceed the conforming loan limit and, therefore, are ineligible for purchase by the GSEs.

As noted, it is possible that the median regression coefficients shown in exhibit 21 may mask considerable variation in the coefficient values across individual MSAs. Therefore, in addition to the median coefficients shown in exhibit 21, we have also summarized the empirical results (in a tabular form similar to the exhibits presented to illustrate cross-tabulations of loan purchase activity in different market segments) to assist in discerning patterns across MSAs. Exhibit 22, for example, summarizes the results for Borrower Income Category, showing the cities in which each of the categorical variables indicating borrower income category had negative coefficients significantly different from 0 at the 95 percent level of statistical confidence and the cities in which the same variables had statistically significant positive coefficients.

The general pattern is the same as that revealed by the cross-tabulations: GSEs are less likely to purchase loans made to both the lowest income borrowers and the highest income borrowers, and are more likely to purchase loans made to middle- and upper middle-income borrowers. Loans made to higher income borrowers (those with incomes between 130 percent and 200 percent of MSA median) were significantly more likely to be purchased by the GSEs than were loans to borrowers in the excluded income category in all of the 44 MSAs studied.[25] Loans made to the next-lower category of borrowers (those with incomes between 100 and 130 percent of MSA median) were significantly more likely to be purchased by the GSEs in all but two MSAs; those in the lower middle-income category (70 to 100 percent of MSA median) were significantly more likely to be purchased by the GSEs in 38 of the 44 MSAs. Loans made to borrowers in the lowest income category (less than 70 percent of area median) were more likely to be purchased by GSEs in 21 MSAs, and less likely in 15.

Borrower Race

The median estimated regression coefficient on the Borrower Race variable (BR_RACE) is negative at –0.914, indicating that the GSEs are less likely (other characteristics being equal) to purchase loans made to minority borrowers than they are to purchase loans made to White borrowers. As exhibit 23 shows, this pattern is quite consistent across

Exhibit 22

Summary of GSE/Non-GSE Regression Results by Borrower Income Category

Borrower Income Category	Coefficient Negative and Significantly Different From Zero	Coefficient Not Significantly Different From Zero	Coefficient Positive and Significantly Different From Zero
0–0.7	ANA, BOS, BUF, LAN, MIL, NWO, NYC, NEW, PIT, POR,	BIR, CIN, COL, HOU, MEM, MIA, NOR, SAN	ATL, BAL, CHI, CLE, DAL, DEN, DET, FTW, HAR, IND, KAN, PRO, SBE, SDI, SFR, SJO, MIN, OKL, PHI, PHO, ROC, STL, SAL, SEA, TAM, WAS
0.7–1.0	LAN, NYC	ANA, NWO, SAN, SFR	ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, MEM, MIA, MIL, MIN, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SBE, SDI, SJO, SEA, TAM, WAS
1.0–1.3		NWO, NYC	ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, SFR, SJO, SEA, TAM, WAS
1.3–2.0			ANA, ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, LAN, MEM, MIA, MIL, MIN, NWO, NYC, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, SFR, SJO, SEA, TAM, WAS
2.0+		(Excluded category)	

MSAs: the coefficient on minority borrowers is significantly negative (at 95 percent confidence) in 36 of the 44 MSAs studied and significantly positive in 4.

Ratio of Tract Median Income to MSA Median Income

The coefficient on the lowest income category, TINC_C01, is negative with a median value of –1.101, indicating that loans originated in the lowest income census tracts, those where the median income is less than 60 percent of the MSA median, are less likely to be purchased by either of the GSEs than are those in the highest income census tracts (the excluded category). The coefficient on the next tract income category, TINC_C02, is also negative but less strongly so, suggesting that GSEs are also less likely to purchase loans originated in tracts where the median income is between 60 and 80 percent of the

Exhibit 23

Summary of GSE/Non-GSE Regression Results by Borrower Race

Borrower Race	Coefficient Negative and Significantly Different From Zero	Coefficient Not Significantly Different From Zero	Coefficient Positive and Significantly Different From Zero
Minority	ATL, BAL, BIR, BOS, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, MEM, MIL, MIN, NWO, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SDI, SEA, TAM	MIA, SBE, SFR, WAS	ANA, LAN, NYC, SJO
White		(Excluded category)	

Exhibit 24

Summary of GSE/Non-GSE Regression Results by Ratio of Tract to MSA Median Income

Tract/MSA Income Ratio	Coefficient Negative and Significantly Different From Zero	Coefficient Not Significantly Different From Zero	Coefficient Positive and Significantly Different From Zero
0–60	ANA, ATL, BAL, BIR, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HOU, IND, KAN, LAN, MEM, MIA, MIL, NWO, NEW, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SBE, SDI, TAM, WAS	BOS, HAR, MIN, NYC, SFR, SJO, SEA	
60–80	ANA, ATL, BAL, BIR, CIN, DAL, DEN, HOU, KAN, LAN, MEM, MIA, MIL, NWO, NEW, NOR, OKL, PIT, ROC, STL, SAN, SBE, SDI, TAM	BUF, CLE, COL, FTW, IND, MIN, PHI, PHO, POR, PRO, SAL, SFR, WAS	BOS, CHI, DET, HAR, NYC, SJO, SEA
80–100	ANA, BIR, HOU, KAN, MIA, NWO, NOR, OKL, PIT, SBE, SDI, TAM	BAL, DAL, FTW, LAN, MEM, MIL, SAN	ATL, BOS, BUF, CHI, CIN, CLE, COL, DEN, DET, HAR, IND, MIN, NYC, NEW, PHI, PHO, POR, PRO, ROC, STL, SAL, SFR, SJO, SEA, WAS
100–120	ANA, BIR, MIA, NWO, PIT, SBE	DAL, HOU, LAN, NOR, SAN, SDI, TAM	ATL, BAL, BOS, BUF, CHI, CIN, CLE, COL, DEN, DET, FTW, HAR, IND, KAN, MEM, MIL, MIN, NYC, NEW, OKL, PHI, PHO, POR, PRO, ROC, STL, SAL, SFR, SJO, SEA, WAS
>120		(Excluded category)	

MSA median. In contrast, both of the other tract income categories have positive coefficients, suggesting that loans to borrowers living in these medium-income tracts are more likely to be purchased by the GSEs than are loans to borrowers in the highest income tracts (the excluded category).[26]

Exhibit 24 shows that this pattern holds generally across all MSAs. For example, the coefficient on the lowest tract income category is significantly negative in 37 of the 44 MSAs studied, and is not significantly positive in any. The coefficients for the next-lower tract income category (tracts with median income between 60 and 80 percent of MSA median) are significantly negative in 24 MSAs and significantly positive in only 7. The coefficients on tracts with median incomes between 80 and 100 percent of MSA median are significantly positive in 25 MSAs, and significantly negative in 12; and the coefficients on the higher income tracts (those with median incomes between 100 and 120 percent of MSA median) were significantly positive in 31 MSAs and significantly negative in 6.

Tract Percentage Minority

The median coefficients shown in exhibit 21 for the census tract minority-concentration variables are relatively small and do not suggest the same pattern that is observed in the cross-tabulations. Moreover, as exhibit 25 shows, the estimated coefficients are quite

Exhibit 25

Summary of GSE/Non-GSE Regression Results by Percentage Minority of Tract Population

Tract Percent Minority	Coefficient Negative and Significantly Different From Zero	Coefficient Not Significantly Different From Zero	Coefficient Positive and Significantly Different From Zero
0–10		(Excluded category)	
10–15	ATL, BOS, CLE, DET, IND, MIA, MIL, NYC, PIT, PRO, SAL, SBE	BUF, CIN, HAR, KAN, MIN, NWO, NEW, ROC, SJO, SEA	ANA, BAL, BIR, CHI, COL, DAL, DEN, FTW, HOU, LAN, MEM, NOR, OKL, PHI, PHO, POR, STL, SAN, SDI, SFR, TAM, WAS
15–30	BOS, HAR, KAN, MIN, NYC, PIT, PRO	ATL, BIR, BUF, CIN, CLE, COL, DET, IND, MEM, MIL, NEW, OKL, PHI, ROC	ANA, BAL, CHI, DAL, DEN, FTW, HOU, LAN, MIA, NWO, NOR, PHO, POR, STL, SAL, SAN, SBE, SDI, SFR, SJO, SEA, TAM, WAS
30–50	BOS, COL, KAN, MIL, NYC, PIT, PRO, STL	ANA, ATL, BIR, BUF, CIN, CLE, DET, HAR, IND, MEM, MIN, NEW, OKL, PHI, ROC, SEA, WAS	BAL, CHI, DAL, DEN, FTW, HOU, LAN, MIA, NWO, NOR, PHO, POR, SAL, SAN, SBE, SDI, SFR, SJO, TAM
>50	BOS, BUF, CHI, CIN, CLE, COL, DET, IND, KAN, MIL, MIN, NYC, NOR, OKL, PHI, PIT, PRO, STL, SEA, TAM, WAS	ANA, ATL, BAL, BIR, FTW, HAR, MEM, NEW, PHO, POR, ROC, SAL, SBE	DAL, DEN, HOU, LAN, MIA, NWO, SAN, SDI, SFR, SJO

mixed across the MSAs studied. For example, the coefficient on the highest minority-concentration category (census tracts with greater than 50 percent minority population) was significantly negative in 21 MSAs but significantly positive in 10 MSAs and not significantly different from 0 in the remaining 13. The coefficients on the middle ranges of minority concentration—those for tracts with 10 to 15 percent, 15 to 30 percent, and 30 to 50 percent minority populations—were significantly positive in most of the MSAs studied. This result, indicating that GSEs are more likely to purchase loans in these racially mixed tracts than in predominantly White tracts (the excluded category), contrasts with the findings suggested by the cross-tabulations that indicate that the GSEs' market share is highest in predominantly White census tracts.[27]

Central-City Versus Suburban Tracts

The median estimated value for the central-city coefficient (–0.261 in exhibit 21) is negative, indicating that GSEs are less likely to purchase loans made to borrowers in central-city areas. Exhibit 26 suggests that this result is somewhat consistent across the MSAs studied, because the estimated central-city coefficient is significantly negative in 28 of the 44 MSAs and significantly positive in 9. This result agrees with the findings of the cross-tabulations that suggested that the GSEs' market share is lower in central-city areas of most MSAs.

Geographically Targeted Census Tracts

The median estimated value for the coefficient indicating geographic targeting (–0.130 in exhibit 21) is also negative, but not very strongly so, indicating that GSEs may be slightly less likely to purchase loans made to borrowers in geographically targeted census tracts. As exhibit 27 shows, the estimated geographic targeting coefficient is significantly negative in 19 MSAs, significantly positive in another 8, and not significantly different from 0 in the remaining 17. This result is in contrast to that seen in the cross-tabulations that pointed more uniformly to lower GSE market shares in targeted areas and is largely explainable by the collinearity between the geographic targeting variable and the tract income and tract percent minority variables.

In summary, the logistic regression results generally reinforce the conclusions suggested by the cross-tabulations discussed earlier. The GSEs appear to be more likely to purchase loans made to moderate- or higher income borrowers, White borrowers, borrowers in higher income census tracts, and borrowers in suburban areas, controlling for differences

Exhibit 26

Summary of GSE/Non-GSE Regression Results by Central-City Versus Suburban Location

Tract Location	Coefficient Negative and Significantly Different From Zero	Coefficient Not Significantly Different From Zero	Coefficient Positive and Significantly Different From Zero
Central City*	ANA, ATL, BAL, BOS, BUF, CHI, CIN, CLE, DAL, DEN, FTW, HAR, HOU, KAN, LAN, MEM, MIA, MIN, NWO, NYC, NEW, PIT, POR, SAL, SBE, SEA, TAM, WAS	BIR, DET, NOR, PRO, ROC, SDI, SFR	COL, IND, MIL, OKL, PHI, PHO, STL, SAN, SJO

*Variable indicates the percentage of the census tract located within central-city areas of the MSA.

Exhibit 27

Summary of GSE/Non-GSE Regression Results by Geographically Targeted Census Tract

Whether Targeted	Coefficient Negative and Significantly Different From Zero	Coefficient Not Significantly Different From Zero	Coefficient Positive and Significantly Different From Zero
Targeted	ATL, BAL, BOS, CHI, CLE, COL, DAL, DEN, DET, FTW, NYC, NEW, PHO, PIT, ROC, STL, SAN, SJO, TAM	ANA, BIR, BUF, CIN, HAR, HOU, IND, KAN, LAN, MEM, NWO, PHI, POR, PRO, SBE, SDI, SFR	MIA, MIL, MIN, NOR, OKL, SAL, SEA, WAS
Nontargeted		(Excluded category)	

in other borrowers, neighborhood, or loan characteristics. Conversely, the GSEs are less likely to purchase loans made to lower income borrowers, minority borrowers, borrowers in lower income census tracts, and borrowers in central-city areas. The results for geographically targeted census tracts also reinforce the findings of the cross-tabulations but are much less pronounced, while the results for tract percentage minority actually contradict somewhat the findings from the cross-tabulations. These results suggest that the large differences in lending patterns observed in the cross-tabulations for these neighborhood characteristics may actually be attributable to differences in borrower income, borrower race, tract income, and central-city/suburban location rather than to neighborhood minority concentration or geographic targeting. In short, the logistic regression results suggest that borrower income, borrower race, tract income, and central-city/suburban location all appear to be relevant in predicting whether a loan will be purchased by one of the GSEs.

Changes Over Time in Probability That Loans Are Purchased by a GSE

In addition to the logistic regressions presented above for the 44 large MSAs, we also estimated a second set of regressions that enabled us to focus on changes over time in the probability that loans from different market segments would be purchased by one of the GSEs. Specifically, for each MSA we estimated a logistic regression equation in which we included dummy variables indicating whether each loan originated in 1994, 1995, or 1996 (with 1993 representing the excluded category) and also interacted all four origination/acquisition-year variables with the market-segment variables included in the regressions described above. The results of these additional logistic regressions are summarized in exhibit 28 showing the median of the most important regression coefficients estimated over the full group of 44 MSAs studied.[28] As exhibit 28 shows, in general the regression equations fit the PUDB-HMDA data well: the median percentage of observation-pairs concordant was 86 percent, and the median value of the Somers' *D* statistic was 0.724, indicating generally high predictive power for the models.

Moreover, in order to develop a fuller picture of changes in nationwide GSE purchase activity in different market segments over the 1993–96 study period, we combined the PUDB-HMDA data sets for all 44 large MSAs studied and estimated the same logistic regression equation, allowing for annual changes in regression coefficients as well as for a different intercept for each MSA. The results of this nationwide analysis of GSE mortgage purchase activity in large MSAs is shown in exhibit 29. As this table shows, even though this regression imposes the implicit constraint that slope coefficients be equal

Exhibit 28

Summary of GSE/Non-GSE Logistic Regression Results With Annual Changes

Independent Variable	Median Estimated Regression Coefficient	Average Estimated Regression Coefficient	Standard Deviation of Estimated Coefficients
YEAR1994	0.020	–0.001	0.316
YEAR1995	–0.226	–0.204	0.488
YEAR1996	–0.109	–0.013	0.405
BINC1_93	–0.132	–0.045	0.643
BINC1_94	0.440	0.331	1.087
BINC1_95	0.122	0.178	0.975
BINC1_96	0.025	–0.013	0.784
BINC2_93	0.793	0.724	0.489
BINC2_94	1.335	1.325	0.918
BINC2_95	1.116	1.012	0.777
BINC2_96	0.982	0.923	0.744
BINC3_93	0.847	0.840	0.347
BINC3_94	1.398	1.368	0.611
BINC3_95	1.101	1.078	0.579
BINC3_96	1.072	1.134	0.535
BINC4_93	0.854	0.880	0.275
BINC4_94	1.197	1.272	0.446
BINC4_95	1.144	1.212	0.369
BINC4_96	1.095	1.229	0.375
BRRACE93	–0.659	–0.590	0.456
BRRACE94	–1.121	–1.041	0.725
BRRACE95	–1.076	–0.966	0.694
BRRACE96	–1.006	–0.813	0.648
TINC1_93	–0.783	–0.920	0.755
TINC1_94	–0.937	–1.151	0.810
TINC1_95	–1.355	–1.323	0.894
TINC1_96	–1.245	–1.241	0.895
TINC2_93	–0.094	–0.244	0.588
TINC2_94	–0.277	–0.304	0.667
TINC2_95	–0.219	–0.451	0.713
TINC2_96	–0.275	–0.466	0.737
TINC3_93	0.152	0.076	0.363
TINC3_94	0.206	0.152	0.380
TINC3_95	0.060	–0.035	0.495
TINC3_96	0.081	–0.020	0.505
TINC4_93	0.198	0.177	0.223
TINC4_94	0.351	0.309	0.367
TINC4_95	0.209	0.178	0.341
TINC4_96	0.267	0.211	0.371
TMIN0293	0.088	0.111	0.469
TMIN0294	0.104	0.122	0.506

Exhibit 28 (continued)

Summary of GSE/Non-GSE Logistic Regression Results With Annual Changes

Independent Variable	Median Estimated Regression Coefficient	Average Estimated Regression Coefficient	Standard Deviation of Estimated Coefficients
TMIN0295	0.119	0.126	0.514
TMIN0296	0.106	0.117	0.391
TMIN0393	0.135	0.187	0.365
TMIN0394	0.147	0.228	0.550
TMIN0395	0.080	0.191	0.469
TMIN0396	0.208	0.230	0.461
TMIN0493	0.108	0.172	0.578
TMIN0494	0.177	0.371	0.753
TMIN0495	0.211	0.263	0.867
TMIN0496	0.123	0.196	0.779
TMIN0593	–0.106	–0.154	0.818
TMIN0594	–0.011	0.071	0.988
TMIN0595	–0.141	–0.004	1.142
TMIN0596	–0.394	–0.252	0.901
CNCITY93	–0.247	–0.227	0.356
CNCITY94	–0.221	–0.204	0.459
CNCITY95	–0.237	–0.260	0.435
CNCITY96	–0.260	–0.180	0.374
GEOTGT93	–0.064	–0.089	0.431
GEOTGT94	–0.190	–0.281	0.707
GEOTGT95	–0.230	–0.288	0.748
GEOTGT96	–0.118	–0.167	0.505
Predictive Power:			
Concordant	86.0%	85.4%	5.2%
Somers' *D*	0.724	0.714	0.107

across all MSAs, the goodness-of-fit statistics remain quite high, with 81.1 percent of observation-pairs concordant and a Somers' *D* statistic of 0.625.

In this specification, the dummy variable coefficient estimated for each MSA indicates the relative likelihood that a loan originated in each city would be purchased by a GSE, all other attributes (loan, borrower, and neighborhood characteristics) being equal. For example, the estimated dummy coefficients are relatively low for cities such as Anaheim (the excluded MSA), Los Angeles, Newark, Pittsburgh, San Francisco, and San Jose, indicating that loans in those MSAs were less likely to be purchased by the GSEs. Conversely, the estimated dummy coefficients are relatively high for Norfolk, Oklahoma City, Rochester, and San Antonio, indicating that the GSEs are relatively more likely to purchase loans in those MSAs. These differences across MSAs will be discussed further in the next section.

Considering the borrower and neighborhood characteristics of greatest interest, exhibits 28 and 29 suggest that both ways of summarizing GSE mortgage purchase activity across the 44 MSAs studied yielded essentially identical results, so only the first set of results—

Exhibit 29

Results of GSE/Non-GSE Logistic Regression Estimated Over 44 MSAs

Independent Variable	Estimated Regression Coefficient	Standard Error	Independent Variable	Estimated Regression Coefficient	Standard Error
Intercept*	–27.29	12.2×10^{-5}			
ATL	0.966	1.8×10^{-5}	BINC1_93	–0.148	3.5×10^{-5}
BAL	1.013	2.1×10^{-5}	BINC1_94	0.019	4.3×10^{-5}
BIR	0.960	3.2×10^{-5}	BINC1_95	0.086	4.4×10^{-5}
BOS	0.941	1.8×10^{-5}	BINC1_96	0.153	3.1×10^{-5}
BUF	1.276	3.8×10^{-5}	BINC2_93	–0.247	2.5×10^{-5}
CHI	0.488	1.5×10^{-5}	BINC2_94	0.198	3.1×10^{-5}
CIN	0.791	2.3×10^{-5}	BINC2_95	–0.045	3.1×10^{-5}
CLE	0.201	2.3×10^{-5}	BINC2_96	0.864	2.2×10^{-5}
COL	1.321	2.4×10^{-5}	BINC3_93	–0.239	2.1×10^{-5}
DAL	1.002	2.1×10^{-5}	BINC3_94	0.110	2.6×10^{-5}
DEN	1.032	2.0×10^{-5}	BINC3_95	–0.103	2.6×10^{-5}
DET	0.958	1.7×10^{-5}	BINC3_96	1.002	1.8×10^{-5}
FTW	1.266	3.1×10^{-5}	BINC4_93	–0.253	1.5×10^{-5}
HAR	1.078	3.3×10^{-5}	BINC4_94	0.006	1.9×10^{-5}
HOU	0.965	2.0×10^{-5}	BINC4_95	–0.025	1.8×10^{-5}
IND	1.479	2.5×10^{-5}	BINC4_96	0.930	1.2×10^{-5}
KAN	1.314	2.6×10^{-5}	BRRACE93	0.219	2.0×10^{-5}
LAN	0.145	1.5×10^{-5}	BRRACE94	–0.079	2.5×10^{-5}
MEM	1.472	3.7×10^{-5}	BRRACE95	–0.072	2.4×10^{-5}
MIA	0.705	2.4×10^{-5}	BRRACE96	–0.154	1.6×10^{-5}
MIL	1.138	2.4×10^{-5}	TINC1_93	0.146	6.3×10^{-5}
MIN	1.152	2.0×10^{-5}	TINC1_94	0.148	7.5×10^{-5}
NWO	1.002	3.4×10^{-5}	TINC1_95	0.056	7.5×10^{-5}
NYC	0.286	1.6×10^{-5}	TINC1_96	–0.965	4.9×10^{-5}
NEW	–0.052	2.2×10^{-5}	TINC2_93	0.005	4.2×10^{-5}
NOR	2.104	4.0×10^{-5}	TINC2_94	0.110	5.1×10^{-5}
OKL	1.873	4.4×10^{-5}	TINC2_95	0.031	5.0×10^{-5}
PHI	0.799	1.8×10^{-5}	TINC2_96	–0.263	3.3×10^{-5}
PHO	1.212	2.1×10^{-5}	TINC3_93	–0.042	2.6×10^{-5}
PIT	–0.057	2.9×10^{-5}	TINC3_94	0.126	3.1×10^{-5}
POR	1.282	2.2×10^{-5}	TINC3_95	–0.022	3.0×10^{-5}
PRO	1.443	3.8×10^{-5}	TINC3_96	0.071	1.9×10^{-5}
ROC	1.618	3.7×10^{-5}	TINC4_93	–0.048	1.7×10^{-5}
STL	1.237	2.2×10^{-5}	TINC4_94	0.069	2.1×10^{-5}
SAL	1.332	2.6×10^{-5}	TINC4_95	–0.018	2.1×10^{-5}
SAN	2.038	4.1×10^{-5}	TINC4_96	0.184	1.3×10^{-5}
SBE	1.406	2.3×10^{-5}	TMIN0293	–0.063	1.9×10^{-5}
SDI	0.716	1.8×10^{-5}	TMIN0294	0.035	2.4×10^{-5}

Exhibit 29 (continued)

Results of GSE/Non-GSE Logistic Regression Estimated Over 44 MSAs

Independent Variable	Estimated Regression Coefficient	Standard Error	Independent Variable	Estimated Regression Coefficient	Standard Error
SFR	–0.887	$1.9x10^{-5}$	TMIN0295	0.016	$2.4x10^{-5}$
SJO	–0.549	$1.7x10^{-5}$	TMIN0296	0.041	$1.6x10^{-5}$
SEA	1.012	$1.9x10^{-5}$	TMIN0393	–0.031	$1.8x10^{-5}$
TAM	1.053	$2.3x10^{-5}$	TMIN0394	0.119	$2.3x10^{-5}$
WAS	0.639	$1.6x10^{-5}$	TMIN0395	–0.024	$2.2x10^{-5}$
Year1994	–0.080	$1.3x10^{-5}$	TMIN0396	0.151	$1.5x10^{-5}$
Year1995	0.149	$1.5x10^{-5}$	TMIN0493	–0.046	$2.9x10^{-5}$
Year1996	–0.069	$1.5x10^{-5}$	TMIN0494	0.102	$3.6x10^{-5}$
UPB_C1X	1.426	$1.6x10^{-5}$	TMIN0495	–0.001	$3.6x10^{-5}$
UPB_C2X	1.792	$1.1x10^{-5}$	TMIN0496	0.303	$2.4x10^{-5}$
UPB_C3X	1.981	$0.1x10^{-5}$	TMIN0593	–0.076	$3.3x10^{-5}$
UPB_C4X	1.630	$0.1x10^{-5}$	TMIN0594	0.217	$4.1x10^{-5}$
PtPlFml2	–0.154	$1.2x10^{-5}$	TMIN0595	0.045	$4.1x10^{-5}$
PctOwnOc	1.404	$2.1x10^{-5}$	TMIN0596	0.303	$2.8x10^{-5}$
Pct_HmPr	–0.484	$0.1x10^{-5}$	CNCITY93	0.053	$1.6x10^{-5}$
Pct_Cvtl	22.65	$11.8x10^{-5}$	CNCITY94	0.038	$2.0x10^{-5}$
Pct_FHA	1.084	$17.3x10^{-5}$	CNCITY95	–0.077	$2.0x10^{-5}$
Pct_FDIC	–0.928	$1.8x10^{-5}$	CNCITY96	–0.216	$1.3x10^{-5}$
Pct_FRB	2.428	$1.6x10^{-5}$	GEOTGT93	0.172	$3.2x10^{-5}$
Pct_HUD	2.158	$1.2x10^{-5}$	GEOTGT94	0.053	$3.9x10^{-5}$
Pct_NCUA	–8.037	$6.9x10^{-5}$	GEOTGT95	–0.082	$3.8x10^{-5}$
Pct_OCC	0.817	$1.3x10^{-5}$	GEOTGT96	–0.109	$2.5x10^{-5}$
Concordant: 81.1%			Somers' *D*: 0.625		

*ANA is the excluded MSA.

the median regression coefficients over the 44 separate logistic regression equations estimated—are discussed in detail below.

The regression coefficients shown in exhibits 28[29] and 29 can be used to illustrate changes over the study period from 1993 to 1996 in the estimated probability that loans originated in any specific market segment would be purchased by one of the GSEs. Specifically, we compare the probability of GSE purchase for any given loan to a *baseline* probability estimated for loans that fall in the excluded category on all discrete characteristics. The baseline represents loans originated/acquired in 1993 for which (1) the borrower is in the highest income category, (2) the borrower is White, (3) the property is in one of the highest income census tracts, (4) the property is in one of the most predominantly White census tracts, (5) the property is in a suburban area, and (6) the property is in a census tract that is not geographically targeted.[30]

For an identical loan originated/acquired in a year (1994, 1995, or 1996) other than 1993, the change in the probability of GSE purchase is indicated by adding the coefficient estimated for that year's dummy variable (Year1994, Year1995, or Year1996). Similarly, the

difference in GSE purchase probability between a baseline loan and any loan that differs according to the other discrete characteristics—borrower income category, borrower race, tract income ratio category, tract percent minority category, central-city/suburban location, and geographically targeted/nontargeted tract—can be estimated by adding the coefficient(s) estimated for the corresponding discrete variable(s).

For example, relative to the GSE purchase probability for a baseline loan originated/acquired in 1993, exhibit 28 indicates that an identical loan originated/acquired in 1994 would be slightly more likely to be purchased by a GSE (that is, the median estimated coefficient on Year1994 is positive at +0.020, so the incremental contribution to predicted GSE purchase probability is positive by 2 percent[31] at the median) while the same loan originated/acquired in 1995 or 1996 would be slightly less likely to be purchased by a GSE (the median estimated coefficients on Year1995 and Year1996 are negative at –0.226 and –0.109, so the median incremental contribution to predicted GSE purchase probability is negative for each year by 20 percent and 10 percent, respectively).[32]

For a loan that differs from the baseline loan only in that the borrower is a member of a minority racial group rather than White, the coefficients indicate that the probability of GSE purchase is sharply lower in all years. This is indicated by the negative median coefficient estimates for the variables indicating minority borrowers in all four years: BRRACE93 = –0.659 (48 percent less likely at the median), BRRACE94 = –1.121 (67 percent less likely), BRRACE95 = –1.076 (66 percent less likely), and BRRACE96 = –1.006 (63 percent less likely). The estimated GSE purchase probability, then, is lower for loans to minority borrowers than for loans to White borrowers in all 4 years and actually worsened sharply in 1994 before improving very slightly in 1995 and 1996.

The same process can be used to compare groups of loans to each other rather than merely to the baseline loans. For example, the median estimated coefficient on the variable TINCC4_93 is +0.198, indicating that the probability of GSE purchase in 1993 was 22 percent greater at the median for loans in these neighborhoods than in neighborhoods of the excluded tract income category. Conversely, the median estimated coefficient on the variable TINCC1_93 is –0.783, indicating that the probability of GSE purchase in 1993 was 54 percent smaller for loans in the lowest income neighborhoods than in neighborhoods of the excluded tract income category. The difference between these two median estimated coefficients, 0.981, reflects the difference in 1993 predicted GSE purchase probability (about 167 percent at the median) between loans in the high-income and low-income neighborhoods.[33] The difference is even greater in the other years of the analysis: 1.288 (241 percent) in 1994, 1.564 (378 percent) in 1995, and 1.512 (354 percent) in 1996, indicating that the difference in purchase probability between loans in higher income neighborhoods and loans in lower income neighborhoods seems to have increased over the study period. It is worth noting that the difference in incremental contribution to GSE purchase probability between the highest probability neighborhoods and the lowest probability neighborhoods is substantially greater than the difference in incremental contribution to GSE purchase probability between White and minority borrowers: that is, tract income ratio seems to be more important than borrower race in determining the probability that loans will be purchased by one of the GSEs.

Using this process, we can compute the incremental contribution to the predicted GSE purchase probability for loans represented by any combinations of values of the discrete variables and compare them to the baseline loans or to each other. For example, the median estimated regression coefficients suggest that we can identify a set of loans that can generally be predicted to have a relatively high probability of purchase by the GSEs. These *high-probability* loans are those that are originated to borrowers with incomes between 130

percent and 200 percent of the MSA median income (BINC4_93 = 1, BINC4_94 = 1, BINC4_95 = 1, BINC4_96 = 1) who are White (BRRACE93 = 0, BRRACE94 = 0, BRRACE95 = 0, BRRACE96 = 0), who live in census tracts with tract median incomes between 100 and 120 percent of the MSA median (TINC4_93 = 1, TINC4_94 = 1, TINC4_95 = 1, TINC4_96 = 1), and with minority concentrations between 15 and 30 percent of total tract population (TMIN0393 = 1, TMIN0394 = 1, TMIN0395 = 1, TMIN0396 = 1) that are located in suburban areas (CNCITY93 = 0, CNCITY94 = 0, CNCITY95 = 0, CNCITY96 = 0) and are not geographically targeted (GEOTGT93 = 0, GEOTGT94 = 0, GEOTGT95 = 0, GEOTGT96 = 0). The median coefficient estimates shown in exhibit 28 indicate that the estimated incremental contribution to predicted GSE purchase probability for these loans in 1993 is +1.187, computed by summing 0.854 (BINC4_93), 0.198 (TINC4_93), and 0.135 (TMIN0393). The estimated incremental contribution for the same loans in 1994 is +1.715, in 1995 it is +1.207, and in 1996 it is +1.461.

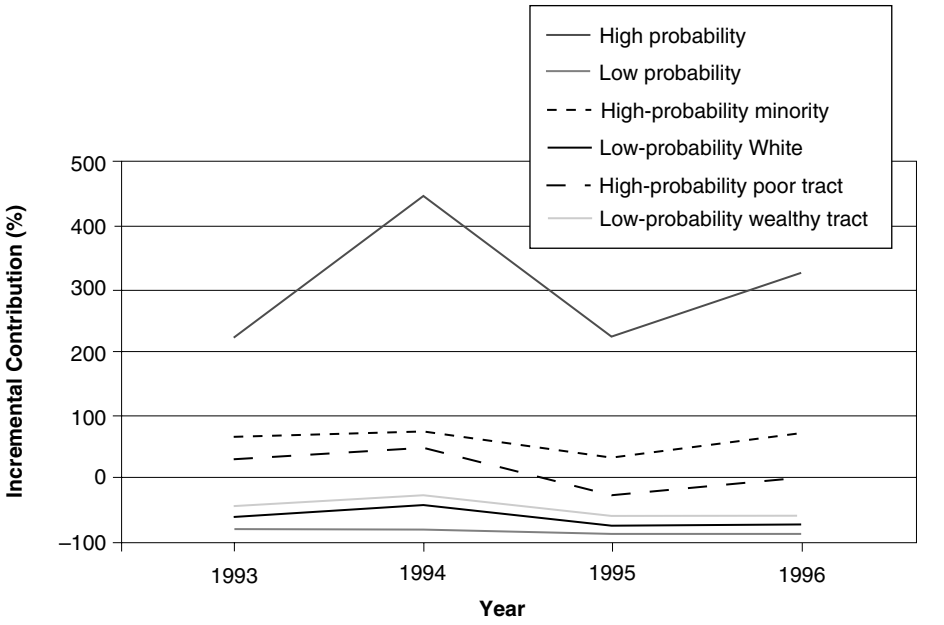
It is interesting to compare the estimated incremental contribution to the predicted value of the dependent variable for these high-probability loans to the incremental contribution for a group of loans that appear to have a relatively low probability of GSE purchase. These *low-probability* loans are those that are originated to borrowers with incomes less than 70 percent of the MSA median income (BINC1_93 = 1, BINC1_94 = 1, BINC1_95 = 1, BINC1_96 = 1), are members of minority racial groups (BRRACE93 = 1, BRRACE94 = 1, BRRACE95 = 1, BRRACE96 = 1), and who live in census tracts with tract median incomes less than 60 percent of the MSA median (TINC1_93 = 1, TINC1_94 = 1, TINC1_95 = 1, TINC1_96 = 1) and with minority concentrations greater than 50 percent of total tract population (TMIN0593 = 1, TMIN0594 = 1, TMIN0595 = 1, TMIN0596 = 1) that are located in central-city areas (CNCITY93 = 1, CNCITY94 = 1, CNCITY95 = 1, CNCITY96 = 1) and are geographically targeted (GEOTGT93 = 1, GEOTGT94 = 1, GEOTGT95 = 1, GEOTGT96 = 1). The median coefficient estimates shown in exhibit 28 indicate that the estimated incremental contribution to predicted GSE purchase probability for these loans in 1993 is –1.991, computed by summing –0.132 (BINC1_93), –0.659 (BRRACE93), –0.783 (TINC1_93), –0.106 (TMIN0593), –0.247 (CNCITY93), and –0.064 (GEOTGT93). The estimated incremental contribution for the same loans in 1994 is –2.020, in 1995 it is –3.143, and in 1996 it is –3.107.

The difference in the estimated incremental contribution to the predicted value of the dependent variable for high-probability and low-probability loans was 3.178 in 1993, 3.735 in 1994, 4.350 in 1995, and 4.568 in 1996. These figures suggest that, at the median, high-probability loans were much more likely to be purchased by the GSEs than were low-probability loans: 24 times as likely in 1993, 42 times in 1994, 77 times in 1995, and 96 times as likely in 1996.[34] They also indicate that the difference in GSE purchase probability between the two groups of loans seems to have widened sharply during the study period.

Exhibit 30 shows the estimated incremental contribution to the predicted value of the dependent variable in each year of the analysis for these two groups of loans (high-probability and low-probability) along with two groups that differ from them only in the race of the borrower (low-probability White and high-probability minority) and two others that differ only in the median income of the census tract (high-probability poor tract and low-probability wealthy tract). As this exhibit suggests, loans originated to high-income White borrowers in high-income, predominantly White suburban, nontargeted census tracts (high-probability) are much more likely to be purchased by the GSEs in each year of the analysis than are loans originated to low-income minority borrowers in low-income, predominantly minority, central-city, geographically targeted census tracts

Exhibit 30

Loans With High Versus Low Purchase Probability



(low-probability). Tract income has a significant effect: loans in high-income neighborhoods are substantially more likely to be purchased by the GSEs than are loans in low-income neighborhoods, regardless of other characteristics of the neighborhood or borrower. Borrower race, however, has an even greater effect: loans to White borrowers are more likely to be purchased by GSEs than are loans to minority borrowers, regardless of borrower income or neighborhood characteristics, and the difference in GSE purchase probability between loans to White and minority borrowers is even greater than the difference in GSE purchase probability between high- and low-income neighborhoods.

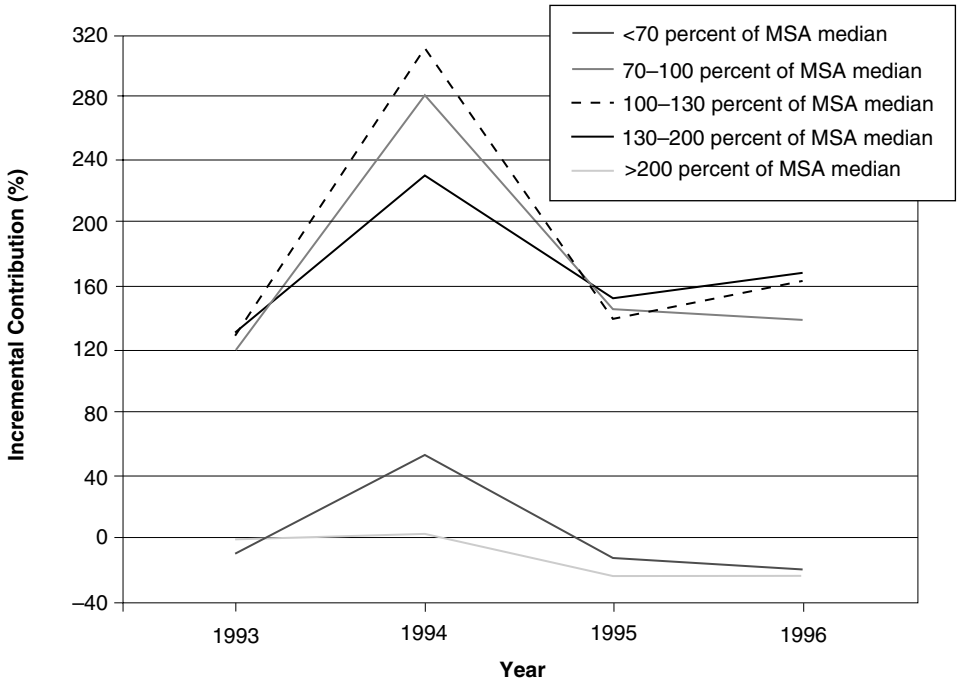
Although this type of comparison of specific groups of loans is useful, the major advantage of the multiple regression approach is that it enables us to focus on the effect of changes in one borrower or neighborhood characteristic on the probability of GSE purchase, holding all other characteristics constant. Exhibits 31 through 36 show the incremental contribution to the predicted value of the dependent variable for each borrower or neighborhood characteristic considered in this discussion over the period 1993 to 1996. In each case, the incremental contribution is computed by adding the median estimated coefficient from exhibit 28 for the dummy variables indicating origination/acquisition year (Year1993 = 0, Year1994 = +0.020, Year1995 = –0.226, and Year1996 = –0.109) to the median estimated coefficient for each borrower or neighborhood characteristic for the same year. In each case, the excluded category is represented by the median estimated coefficients for the year dummies.

Borrower Income Category

Exhibit 31 depicts the incremental contribution to the predicted value of the dependent variable for borrower income category in the 44 large MSAs included in the analysis

Exhibit 31

Borrower Income Category



over the 1993–96 time period. As this exhibit and exhibits 28 and 29 show, other things being equal, the median estimated probability that loans to the highest income borrowers (those with incomes greater than twice the MSA median) are purchased by a GSE remained essentially unchanged from 1993 to 1994, declined from 1994 to 1995, and then recovered somewhat from 1995 to 1996. The median estimated GSE purchase probabilities for loans to borrowers in the other four income categories move essentially parallel to each other: increasing sharply in 1994, then declining in 1995 almost to their 1993 level, and remaining essentially unchanged in 1996.

Borrower Race

Exhibit 32 depicts the incremental contribution for borrower race in the 44 large MSAs from 1993 to 1996. The incremental contribution to predicted GSE purchase probability in 1993 is zero for Whites compared to –0.659 for minorities; in 1994 it is +0.020 for Whites and –1.101 (= 0.020 – 1.121) for minorities; in 1995 it is –0.226 for Whites and –1.302 (= – 0.226 – 1.076) for minorities; and in 1996 it is –0.109 for Whites and –1.115 (= – 0.109 – 1.006) for minorities. These figures indicate that the median estimated probability that loans to White borrowers would be purchased by either of the GSEs (other things being equal) seems to increase very slightly by 2 percent from 1993 to 1994, then drop sharply by about 22 percent in 1995 before recovering somewhat by about 12 percent in 1996. The median estimated probability that loans to minority borrowers would be purchased by the GSEs, however, declines sharply by about 36 percent from 1993 to 1994 and by about another 18 percent from 1994 to 1995 before again recovering somewhat by about 21 percent in 1996.[35]

Exhibit 32

Borrower Race

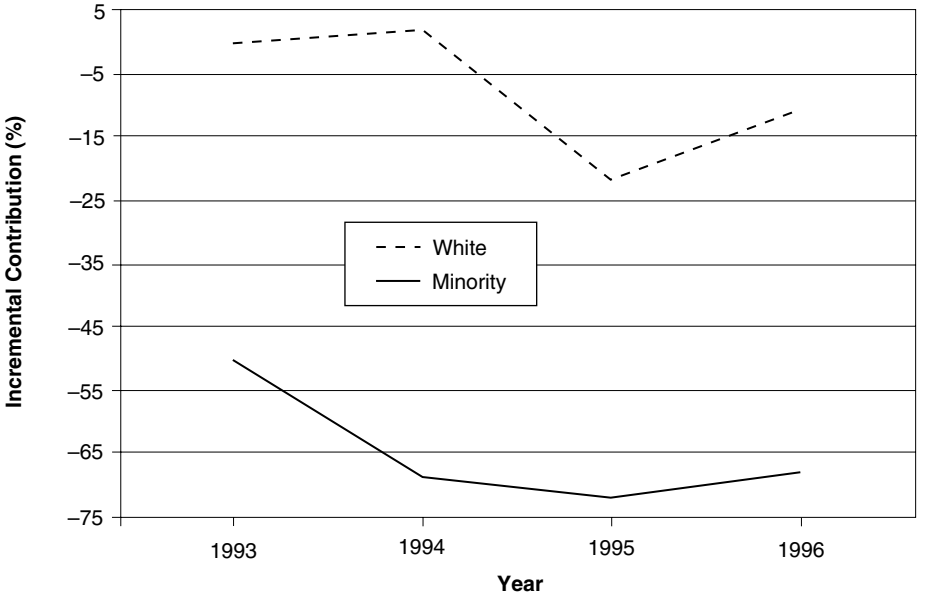
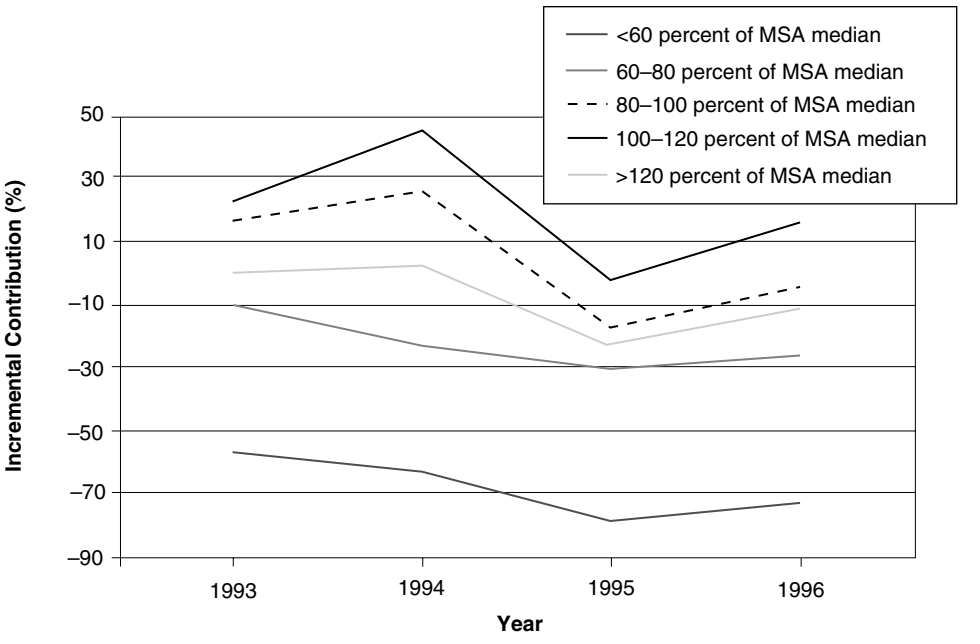


Exhibit 33

Census Tract Median Income Category



Ratio of Tract Median Income to MSA Median Income

Exhibit 33 depicts the incremental contribution according to census tract median income in the 44 large MSAs over the period 1993–96. As this exhibit and exhibits 28 and 29 show, the estimated probabilities that loans to borrowers in each census tract income category would be purchased by a GSE move generally parallel to each other over the time period but (except for the highest income tracts) increase steadily with tract income. Loans in the lowest income census tracts decline sharply enough from 1994 to 1995 so that they remain less likely to be purchased by GSEs in 1996 than they were in 1993.

Tract Percentage Minority

Exhibit 34 depicts the incremental contribution for tracts with different levels of minority concentration in the 44 MSAs over the 1993–96 period of study. As this exhibit and exhibits 28 and 29 show, the probability that loans in all types of census tracts would be purchased by either of the GSEs increases slightly from 1993 to 1994, declines sharply from 1994 to 1995, and then recovers somewhat from 1995 to 1996—except for loans to borrowers in predominantly minority census tracts. Loans to borrowers in tracts with greater than 50 percent minority concentration dropped from 1995 to 1996.

Central-City Versus Suburban Tracts

Exhibit 35 depicts the incremental contribution for central-city versus suburban tract location in the 44 large MSAs from 1993 to 1996. As this exhibit and exhibits 28 and 29 show, the median estimated probability that loans to central-city borrowers would be purchased by either of the GSEs, other things being equal, parallels the median estimated probability for GSE purchase of loans to suburban borrowers, with a sharp decline in 1995 greater than the modest gains in 1994 and 1996.

Exhibit 34

Census Tract Minority Concentration Category

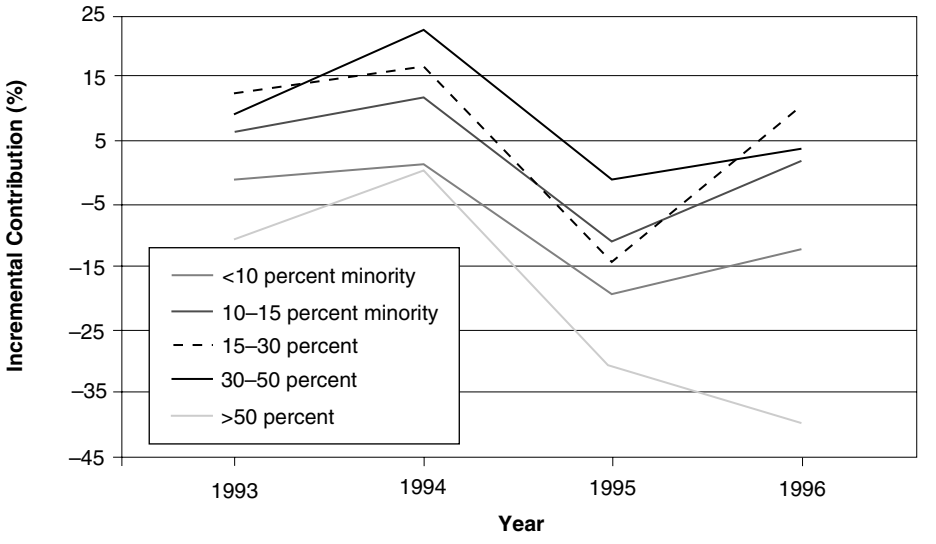


Exhibit 35

Census Tract Location

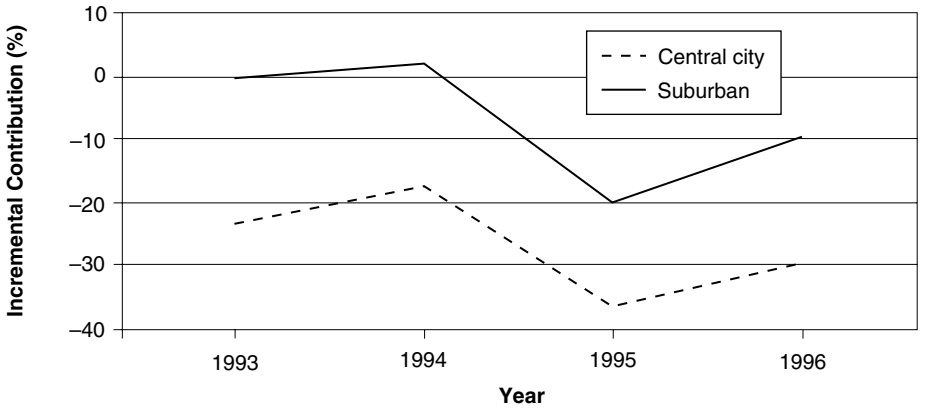
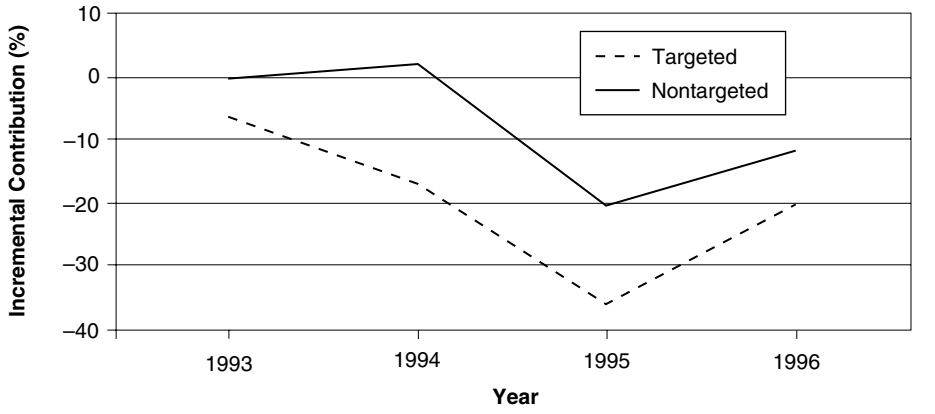


Exhibit 36

Geographically Targeted Versus Nontargeted Census Tracts



Geographically Targeted Census Tracts

Exhibit 36 shows the incremental contribution for loans to borrowers in geographically targeted versus nontargeted census tracts. As this exhibit and exhibits 28 and 29 show, the median probability that loans to borrowers in geographically targeted tracts would be purchased by either of the GSEs, other things being equal, declined from 1993 to 1994 and again more sharply from 1994 to 1995, before recovering in 1996.

In summary, the time-interacted logistic regression results suggest that, while the probability that loans originated to particular types of borrowers changed from year to year, there is no systematic time trend visible in any of the findings. The differences from year to year can perhaps be explained by changes in mortgage origination activity (such as a surge in

mortgage refinances that occurred in 1993) or changes in the reporting of mortgage originations under HMDA. At any rate, there does not appear to be any systematic change over time in the probability that loans originated to particular types of borrowers, including those with limited access to mortgage credit, would be purchased by either of the GSEs.

Systematic Variation in GSE Mortgage Purchase Activity by MSA Type

Although the conclusions that we drew from our analysis of cross-tabulations and logistic regressions were primarily based on consistent patterns that we observed across a large number of the MSAs studied, the considerable variation in both sets of empirical results raises the question whether there are some types of metropolitan areas where the GSEs tend to be more active in purchasing mortgage loans. Several of the cross-tabulations—particularly those focusing on differences in market share by borrower race (exhibits 5 and 6) but also several others as well—seem to suggest that the MSAs in California (Anaheim, Los Angeles, San Bernardino, San Diego, San Francisco, and San Jose) are consistently different from those in other parts of the country. New York, Washington, and a few other MSAs also seem to appear frequently among the exceptions to the general patterns suggested by the cross-tabulations.

To explore whether there are any consistent differences among MSAs in terms of GSE mortgage purchase activity, we used the estimated coefficients from the logistic regressions summarized in exhibit 28 as descriptive indicators of GSE mortgage purchase patterns. We then used a *clustering* method to explore whether there were any natural groupings of MSAs such that the MSAs within a given group all share relatively similar values for the estimated regression coefficients, whereas MSAs in different groups would have relatively dissimilar values for the estimated coefficients.[36]

Exhibit 37 summarizes the results of this clustering analysis. As this table shows, the 44 MSAs included in this study group naturally into 2 clusters according to the patterns of GSE mortgage purchase activity indicated by the estimated coefficients from the logistic regressions. The first cluster consists of the California metropolitan areas—Anaheim, Los Angeles, San Bernardino, San Diego, San Francisco, and San Jose—along with Boston, Newark, New York, and Washington; the second cluster encompasses all 34 other MSAs included in the study.

Using these clusters enables us to explore more carefully the similarities in logit regression coefficients within the two groups and the differences in regression coefficients across the two groups. Exhibit 38 summarizes this view of the regression results, showing the median estimated regression coefficient for MSAs in each of the two groups.[37]

Exhibit 37

Groups of MSAs Identified by Cluster Analysis on Logit Regression Coefficients

Cluster 1: California Plus	Cluster 2: Other MSAs
ANA, LAN, SBE, SDI, SFR, SJO, BOS, NYC, NEW, WAS	ATL, BAL, BIR, BUF, CHI, CIN, CLE, COL, DAL, DEN, DET, FTW, HAR, HOU, IND, KAN, MEM, MIA, MIL, MIN, NWO, NOR, OKL, PHI, PHO, PIT, POR, PRO, ROC, STL, SAL, SAN, SEA, TAM

Exhibit 38

Median Logistic Regression Coefficients of MSAs in Two Clusters

Independent Variable	Median Regression Coefficient: California Plus	Median Regression Coefficient: Other MSAs
BINC_C1X	–0.83	0.27
BINC_C2X	0.23	1.08
BINC_C3X	0.73	1.14
BINC_C4X	0.77	1.08
BR_RACE	0.02	–1.02
TINC_C01	–0.42	–1.21
TINC_C02	–0.12	–0.27
TINC_C03	0.22	0.15
TINC_C04	0.24	0.25
TMIN_C02	0.16	0.15
TMIN_C03	0.31	0.09
TMIN_C04	0.29	0.12
TMIN_C05	–0.03	–0.30
CNTRCITY	–0.20	–0.27
GEOTARGT	–0.07	–0.15
UPB_C1X	3.35	0.31
UPB_C2X	3.09	0.76
UPB_C3X	3.03	0.99
UPB_C4X	1.94	1.12
PtPlFml2	–0.18	–0.13
PctOwnOc	0.97	2.16
Pct_HmPr	–0.97	–0.45
Pct_Cvtl	17.4	13.8
Pct_FHA	–8.19	17.9
Pct_FDIC	–0.76	–0.85
Pct_FRB	2.35	3.15
Pct_HUD	1.54	2.63
Pct_NCUA	–14.9	–9.52
Pct_OCC	0.03	0.95

As this table shows, the greatest difference between the two groups is in the variables indicating the effect of loan balance on GSE market share. Among the 10 MSAs in the *California-plus* cluster, the median coefficients for the loan balance variables are all very strongly positive (ranging from 1.94 to 3.35) and the largest median coefficient applies to loans in the smallest average loan balance category, indicating that GSEs are more likely to purchase smaller loans than larger loans in these cities. In contrast, among the 34 *Other MSAs,* the median coefficients for the loan balance variables are smaller (ranging from 0.31 to 1.12) and the largest median coefficient is for the variable UPB_C4X, indicating that GSEs are more likely to purchase loans in the fourth quintile of the distribution in these cities—that is, that they are more likely to purchase larger loans than smaller loans,

except for the largest loans, which as noted before, are most likely to exceed the conforming loan limit.

It is interesting to note that the 10 MSAs in the California-plus group all consist of metropolitan areas with relatively high housing costs, and the 34 MSAs in the other group have relatively low housing costs. For example, in 1995 the 10 most expensive MSAs in the country[38] were Anaheim, Boston, Honolulu (not included in this study), Los Angeles, Newark, New York, San Diego, San Francisco (including San Jose), Seattle, and Washington, D.C. All of these except Honolulu and Seattle are among the California-plus group of MSAs, along with San Bernardino. Thus, it perhaps should not be surprising that this group of 10 MSAs would differ from the other 34 with respect to the logistic regression coefficient on loan balance category, since more of the loans in these MSAs can be expected to exceed the conforming loan limit and thus be ineligible for purchase by the GSEs.

The second-largest difference between the two groups of MSAs is in the coefficient for borrower race that is near zero for the California-plus group of MSAs, but strongly negative for the Other MSAs at –1.02. As we noted in the section of this article titled "Cross-Tabulations of Loan Purchase Activity in Different Market Segments" and also earlier in this section, the results of our cross-tabulations and logit regressions suggest that the GSEs are systematically more likely to purchase loans made to White borrowers and less likely to purchase loans made to minority borrowers. When seen through the perspective of this cluster analysis, however, it appears that borrower race is not nearly as great a factor in the California-plus group of metropolitan areas as it is in the Other MSAs.

An interesting difference can be noted between the California-plus group of MSAs and the Other MSAs with respect to race and ethnicity: the 10 MSAs in the California-plus group generally have higher proportions of Asians among their minority populations, while the remaining MSAs generally have low Asian populations. Because the analysis does not distinguish among minority borrowers of different racial/ethnic groups, it is possible that loans are more likely to be purchased by the GSEs if they are originated to Asian borrowers rather than to Black or Hispanic borrowers, but this difference is masked in the cross-tabulation and regression results by the fact that all minority borrowers are grouped together for the purposes of the analysis.

The estimated coefficients on the borrower income variables also appear to differ markedly between the two groups of metropolitan areas. The median coefficients on these variables are lower among the California-plus group of MSAs (ranging from –0.83 to 0.77) than among the Other MSAs (ranging from 0.27 to 1.14). In both cases, the lowest median coefficient (aside from the excluded category) applies to borrowers in the lowest income category (BINC_C1X), which suggests that the GSEs are less likely to purchase loans made to the lowest income borrowers than to purchase loans made to borrowers in the higher income categories. The fact that the median coefficients range more widely among the California-plus group of MSAs suggests that the overall pattern—in which GSEs tend to fund mortgage lending less to the lowest income borrowers than to the higher income groups—is particularly strong in the California-plus metropolitan areas.

The median coefficient on the home-purchase variable (Pct_HmPr) is more strongly negative in the California-plus group of MSAs than among the Other MSAs, suggesting that home-purchase loans tend to be somewhat less likely to be purchased by GSEs relative to refinance loans in these cities than in the rest of the country. In contrast, the median coefficient on the owner-occupancy variable (PctOwnOc) is more strongly positive among the Other MSAs than among the California-plus group of MSAs, suggesting that loans to

borrowers in areas with low rates of owner-occupancy are less likely to be funded by the GSEs in the rest of the country than they are in the California-plus group of metropolitan areas, relative to borrowers in areas with higher rates of owner-occupancy.

The differences between the two groups of MSAs do not appear to be nearly as striking among most of the remaining regression coefficients.[39] In particular, there is very little difference among the two groups of metropolitan areas in the median regression coefficients for the geographically targeted census tract variable, the central-city location variable, the tract minority concentration variables, or the tract income variables. This result suggests that, while loans to borrowers in geographically targeted, central-city, predominantly minority, or low-income areas may well be less likely to be purchased by the GSEs (as the results and earlier in this article suggested), there does not appear to be any substantial difference among the two groups of cities in the extent to which this is so.

Summary

The purpose of this study was to combine the new Public Use Database (PUDB) of mortgage acquisitions by Fannie Mae and Freddie Mac with data on mortgage originations collected under the Home Mortgage Disclosure Act (HMDA) to form as complete as possible a picture of the activities of the two GSEs in the national mortgage market. The combined data set makes it possible to estimate the market shares of Fannie Mae and Freddie Mac in several segments of the mortgage market defined by borrower and neighborhood characteristics such as borrower income, borrower race, census tract median income, census tract minority concentration, central-city versus suburban location, and geographic targeting. These estimated GSE market shares, in turn, enable us to evaluate the probability that any loan originated to a borrower with limited access to mortgage credit would be purchased by one of GSEs. The answers to this question—the probability that loans to different types of borrowers would be purchased by the GSEs—in turn should assist HUD in evaluating the extent to which GSEs are succeeding in promoting the public purposes specified in their charters.

The PUDB and HMDA databases were combined by grouping loans recorded in each database into pools defined by a unique permutation of variables common to both databases, including geographic location (State, county, MSA, and census tract), borrower race (White/minority), borrower income category, GSE purchaser (Fannie Mae/Freddie Mac), and loan origination/acquisition year. These loan pools, the majority of which consisted of just one loan, were then matched across the two databases. Problems with missing data, implausible data values, inconsistent data collection, inconsistent variable definitions, and undetectable incorrect data made this pooling and matching process extremely difficult and make it impossible to ascertain precisely the accuracy of the loan matches. Nevertheless, the procedure results in a database that represents substantially all of the nationwide mortgage market, including nonconforming loans and subprime loans as well as conforming loans in the primary market. We are confident that the procedure resulted in the most complete and accurate possible nationwide database of loan origination and purchasing activity.

Using this combined PUDB-HMDA database, we estimated the market shares of Fannie Mae and Freddie Mac, relative to each other and to the non-GSE portion of the mortgage market, for several different segments of the market identified by borrower and neighborhood characteristics. These cross-tabulations of market share by market segment suggested a very consistent set of conclusions. First, the GSEs' market shares were relatively low among those borrowers who may have limited access to mortgage credit: lower income borrowers, minority borrowers, borrowers in lower income neighborhoods,

borrowers in neighborhoods with high minority concentrations, borrowers in central-city (versus suburban) neighborhoods, and borrowers in geographically targeted census tracts. Conversely, the GSEs' market shares were relatively high among those borrowers who are believed to have the greatest access to mortgage credit: higher income borrowers, White borrowers, borrowers in higher income neighborhoods, borrowers in predominantly White neighborhoods, borrowers in suburban neighborhoods, and borrowers in nontargeted census tracts. Second, while this pattern appears to hold for both GSEs, it appears to be significantly more pronounced for Freddie Mac than for Fannie Mae.

Cross-tabulations, however, enable the researcher to control for only one borrower or neighborhood characteristic at a time, and the high degree of correlation among the characteristics addressed makes it possible that the results of the cross-tabulations are misleading. Because of this, we conducted a further analysis by estimating logistic regression equations that enable us to control for all of the borrower and neighborhood characteristics simultaneously. In general, the logistic regression results reinforced the findings from the cross-tabulations: the GSEs appear to be significantly less likely to purchase loans made to lower income borrowers, minority borrowers, borrowers in lower income neighborhoods, and borrowers in central-city neighborhoods; and significantly more likely to purchase loans made to higher income borrowers, White borrowers, borrowers in higher income neighborhoods, and suburban borrowers, even after controlling for other characteristics. On the other hand, the logistic regression results imply that the differences attributable to geographically targeted census tracts are less pronounced than the cross-tabulations suggested, and that there is essentially no systematic difference attributable to neighborhood minority concentration.

The logistic regression was also extended to test for changes over time (1993–96) in GSE mortgage purchase activity, but no systematic time trends were detected. The differences from year to year can perhaps be explained by changes in mortgage origination activity, such as a surge in mortgage refinances that occurred in 1993, or changes in the reporting of mortgage originations under HMDA. At any rate, there does not appear to be any systematic change over time in the probability that loans originated to particular types of borrowers, including those with limited access to mortgage credit, would be purchased by either of the GSEs.

Further analysis indicated that the 44 MSAs included in this analysis can be divided into 2 groups with somewhat different logistic regression results: the first group includes 6 California cities along with Boston, Newark, New York, and Washington, and the second group includes the remaining 34 MSAs, none of them in California. The California-plus MSAs seem to be differentiated from the others by very high housing costs, implying that a higher proportion of mortgage loans exceed the conforming loan limit and therefore cannot be purchased by GSEs. They also have a higher concentration of Asians among their minority populations, suggesting that perhaps the GSEs are more likely to purchase loans made to Asian borrowers than to other minority borrowers.

Appendix

Exhibit A–1

Summary Report for Philadelphia, All Purchasers, 1993–96

	Fannie Mae Aggregate Loan Balance		Freddie Mac Aggregate Loan Balance		Non-GSE Aggregate Loan Balance		Total Market Aggregate Loan Balance	
	$Millions	%	$Millions	%	$Millions	%	$Millions	%
Borrower income category								
0–0.5	127.7	12	53.5	5	871.0	83	1,052.3	2
0.5–0.6	150.7	16	98.1	10	704.3	74	953.1	2
0.6–0.7	267.1	22	183.7	15	775.4	63	1,226.3	2
0.7–0.8	400.1	23	306.6	17	1,052.1	60	1,758.9	3
0.8–0.9	558.0	25	418.2	18	1,300.4	57	2,276.7	4
0.9–1.0	662.9	27	520.0	21	1,303.8	52	2,486.7	4
1.0–1.1	765.7	25	614.7	20	1,715.1	55	3,095.6	5
1.1–1.2	803.1	27	680.0	23	1,471.7	50	2,954.9	5
1.2–1.3	815.6	28	744.6	26	1,315.9	46	2,876.2	5
1.3–1.4	809.4	28	751.4	26	1,281.3	45	2,842.2	5
1.4–1.5	831.6	29	748.9	26	1,308.7	45	2,889.3	5
1.5–2.0	3,146.4	28	2,946.5	27	5,005.2	45	11,098.3	18
2.0–2.5	1,598.3	25	1,518.0	24	3,286.4	51	6,402.9	11
2.5–3.0	679.8	19	694.4	20	2,185.3	61	3,559.7	6
3.5	751.5	10	733.9	9	6,291.0	81	7,776.6	13
Missing	1,309.8	17	477.3	6	5,770.7	76	7,557.9	12
Total	**13,678.6**	**22**	**11,490.6**	**19**	**35,639.1**	**59**	**60,808.3**	**100**
Borrower race								
White	10,887.4	24	9,527.6	21	25,874.7	56	46,289.9	76
Minority	951.3	20	686.5	14	3,110.6	66	4,748.5	8
Missing	1,839.8	19	1,276.3	13	6,653.7	68	9,769.9	16
Total	**13,678.6**	**22**	**11,490.6**	**19**	**35,639.1**	**59**	**60,808.3**	**100**
Tract median income ratio (%)								
0–60	128.3	19	52.2	8	503.9	74	684.6	1
60–80	457.2	21	271.8	12	1,472.9	67	2,202.1	4
80–100	2,028.1	23	1,448.1	16	5,488.5	61	8,964.9	15
100–120	3,454.6	24	2,816.2	20	7,831.5	56	14,102.5	23
>120	7,590.2	22	6,829.6	20	20,262.8	58	34,682.7	57
Missing	19.8	12	72.4	42	79.2	46	171.5	0
Total	**13,678.6**	**22**	**11,490.6**	**19**	**35,639.1**	**59**	**$60,808.3**	**100**

Exhibit A–1 (continued)

Summary Report for Philadelphia, All Purchasers, 1993–96

	Fannie Mae Aggregate Loan Balance		Freddie Mac Aggregate Loan Balance		Non-GSE Aggregate Loan Balance		Total Market Aggregate Loan Balance	
	$Millions	**%**	**$Millions**	**%**	**$Millions**	**%**	**$Millions**	**%**
Tract percent minority (%)								
0–10	10,142.4	22	8,847.4	20	26,172.5	58%	45,162.4	74
10–15	1,672.3	25	1,275.9	19	3,761.3	56	6,709.7	11
15–30	1,043.9	22	811.3	17	2,973.7	62	4,829.0	8
30–50	477.5	22	334.0	16	1,324.4	62	2,136.0	4
>50	342.2	18	161.6	9	1,371.1	73	1,875.0	3
Missing	1.0	0	60.1	63	35.9	37	96.1	0
Total	**13,678.6**	**22**	**11,490.6**	**19**	**35,639.1**	**59**	**60,808.3**	**100**
Central-city versus suburban tract								
Suburban	12,148.8	23	10,511.0	19	31,325.9	58	53,985.9	89
Central City	1,478.9	23	875.8	14	4,009.9	63	6,364.7	10
Missing	50.7	11	103.6	23	303.2	66	457.7	1
Total	**13,678.6**	**22**	**11,490.6**	**19**	**35,639.1**	**59**	**60,808.3**	**100**
Geographically targeted tract								
Non-targeted	12,021.3	23	10,404.4	20	30,432.7	58	52,858.4	87
Targeted	1,637.4	21	1,013.7	13	5,127.2	66	7,778.4	13
Missing	19.8	12	72.4	42	79.2	46	171.5	0
Total	**13,678.6**	**22**	**11,490.6**	**19**	**35,639.1**	**59**	**60,808.3**	**100**

Note: Components may not add to totals because of rounding.

Exhibit A–2

Summary Report for Philadelphia, GSE Purchasers Only, 1993–96

	Fannie Mae Aggregate Loan Balance		Freddie Mac Aggregate Loan Balance		Total Market Aggregate Loan Balance	
	$Millions	%	$Millions	%	$Millions	%
Borrower income category						
0–0.5	127.7	70	53.5	30	181.3	1
0.5–0.6	150.7	61	98.1	39	248.8	1
0.6–0.7	267.1	59	183.7	41	450.8	2
0.7–0.8	400.1	57	306.6	43	706.8	3
0.8–0.9	558.0	57	418.2	43	976.2	4
0.9–1.0	662.9	56	520.0	44	1,182.9	5
1.0–1.1	765.7	55	614.7	45	1,380.4	5
1.1–1.2	803.1	54	680.0	46	1,483.1	6
1.2–1.3	815.6	52	744.6	48	1,560.2	6
1.3–1.4	809.4	52	751.4	48	1,560.8	6
1.4–1.5	831.6	53	748.9	47	1,580.6	6
1.5–2.0	3,146.4	52	2,946.5	48	6,093.0	24
2.0–2.5	1,598.3	51	1,518.0	49	3,116.4	12
2.5–3.0	679.8	49	694.4	51	1,374.3	5
3.5	751.5	51	733.9	49	1,485.5	6
Missing	1,309.8	73	477.3	27	1,787.2	7
Total	**13,678.6**	**54**	**11,490.6**	**46**	**25,169.2**	**100**
Borrower race						
White	10,887.4	53	9,527.6	47	20,415.1	81
Minority	951.3	58	686.5	42	1,637.8	7
Missing	1,839.8	59	1,276.3	41	3,116.1	12
Total	**13,678.6**	**54**	**11,490.6**	**46**	**25,169.2**	**100**
Tract median income ratio (%)						
0–60	128.3	71	52.2	29	$180.6	1
60–80	457.2	63	271.8	37	729.1	3
80–100	2,028.1	58	1,448.1	42	3,476.3	14
100–120	3,454.6	55	2,816.2	45	6,270.9	25
>120	7,590.2	53	6,829.6	47	14,419.8	57
Missing	19.8	22	72.4	78	92.2	0
Total	**13,678.6**	**54**	**11,490.6**	**46**	**25,169.2**	**100**

Exhibit A–2 (continued)

Summary Report for Philadelphia, GSE Purchasers Only, 1993–96

	Fannie Mae Aggregate Loan Balance		Freddie Mac Aggregate Loan Balance		Total Market Aggregate Loan Balance	
	$Millions	%	$Millions	%	$Millions	%
Tract percent minority						
0–10	10,142.4	53	8,847.4	47	18,989.8	75
10–15	1,672.3	57	1,275.9	43	2,948.3	12
15–30	1,043.9	56	811.3	44	1,855.2	7
30–50	477.5	59	334.0	41	811.6	3
>50	342.2	68	161.6	32	503.9	2
Missing	1	0	60.1	100	60.2	0
Total	**13,678.6**	**54**	**11,490.6**	**46**	**25,169.2**	**100**
Central-city versus suburban tract						
Suburbs	12,148.8	54	10,511.0	46	22,659.9	90
Central City	1,478.9	63	875.8	37	2,354.8	9
Missing	50.7	33	103.6	67	154.4	1
Total	**13,678.6**	**54**	**11,490.6**	**46**	**25,169.2**	**100**
Geographically targeted tract						
Non-targeted	12,021.3	54	10,404.4	46	22,425.7	89
Targeted	1,637.4	62	1,013.7	38	2,651.1	11
Missing	19.8	22	72.4	78	92.2	0
Total	**13,678.6**	**54**	**11,490.6**	**46**	**25,169.2**	**100**

Note: Components may not add to totals because of rounding.

Authors

Bradford Case is an economist with the Board of Governors of the Federal Reserve System. He earned his B.A. from Williams College, his M.P.P. from the University of California at Berkeley, and his Ph.D. in economics from Yale University in 2000. He conducted this research while he was a candidate at Yale. His research has been published in Review of Economics and Statistics, Real Estate Economics, Journal of Real Estate Finance and Economics, *and other journals. He has extensively researched the measurement of house price appreciation as an input into models of mortgage default and prepayment and is currently investigating default risk in mortgage portfolios in support of the negotiation of the New Basel Capital Accord.*

Kevin Gillen holds a B.A. in Economics and Mathematics from Canisius College. After completing postgraduate work at Cambridge University in Great Britain, he worked as a research assistant for the Federal Reserve in Washington, D.C., and the Office of Federal Housing Enterprise Oversight in the U.S. Department of Housing and Urban Development. He is currently in the Ph.D. program in Applied Economics at the Wharton School of the University of Pennsylvania, where he has coauthored several articles published in academic journals. Gillen also holds a patent on a valuation algorithm that he developed while at Wharton.

Susan M. Wachter is professor of real estate and finance at the Wharton School of the University of Pennsylvania and is a former chairperson of the Wharton Real Estate Department. From November 1999 through January 2001 she served as Assistant Secretary for Policy Development and Research at the U.S. Department of Housing and Urban Development. She founded and has served as director of Wharton's geographical information systems laboratory. She is a former president of the American Real Estate and Urban Economics Association, and was editor of Real Estate Economics *from 1997 to 1999.*

The views expressed in this paper are those of the authors and do not necessarily represent the views of the Board of Governors of the Federal Reserve System or its staff.

Notes

1. However, Scheessele (1998) found that "HMDA coverage of Federal Housing Administration (FHA) loans and GSE acquisitions has increased since 1993. In 1996, HMDA covered...81.6 percent of GSE acquisitions."

2. The 44 MSAs selected for the analysis include all of those in which data are collected for the metropolitan sample of the American Housing Survey (AHS). Because the MSAs included in the AHS were selected some time ago, the 44 MSAs do not include several metropolitan areas that have grown rapidly and have recently taken places among the largest in the country, such as Las Vegas, Fort Lauderdale, Orlando, and Charlotte. Conversely, the 44 MSAs included in this analysis include several metropolitan areas that were previously more prominent but have grown more slowly, such as Buffalo, Hartford, and Providence. On balance, however, the MSAs included in this analysis represent all regions of the country and include all of the largest metropolitan areas in the country.

3. It is not clear why our conclusions differ from those of Berkovec and Zorn, but it is worth noting that their data covered the time period 1992–93, whereas ours covered the later time period 1993–96. It is possible that the problem that they cited—systematic underreporting of loans in higher income areas—was addressed during our time period to eliminate the reporting discrepancies that they observed. In addition, Berkovec and Zorn noted that a small number of large lenders who underreported their loans in the HMDA database were responsible for a large portion of the underreporting problem that they observed. It is possible, therefore, that these particular lenders eliminated their underreporting during the later 1993–96 time period. Berkovec and Zorn conducted their study using loans in all metropolitan areas in which Freddie Mac conducted business; since our analysis is restricted to 44 large MSAs, it is possible that the extent and nature of underreporting differs in these 44 large MSAs compared with the remaining metropolitan areas that are not included in our analysis.

4. Scheessele (1998), for example, found that HMDA coverage of FHA loan originations had increased in each year from 1993 to 1996.

5. Loan origination/acquisition year is not, in fact, a variable common to the two data sets. The HMDA data set includes the year in which the loan was originated, while the PUDB includes the year in which the loan was acquired by one of the GSEs. We formed loan pools and merged pools across the two data sets using certain assumptions, such as that most loans, with certain exceptions, are acquired by a GSE in the same year that they were originated, that enable us to infer that a pool of loans in the HMDA data set that were originated in a given year corresponds with a pool of loans

in the GSE data set that were acquired in the same (or the next) year. Therefore, we treat loan origination year and loan acquisition year as if it was one variable common to the two data sets.

6. This means that the average loan pool contained 2.65 loans; it is important to note, however, that the large majority of these pools actually consisted of just one loan, with a few pools containing a much larger number of loans.

7. Again, the loan pools contained an average of 2.97 pools, but most of them consisted of only one loan.

8. In practice, we did not use the full set of 12 common variables to conduct the data set merge, because both the PUDB and HMDA databases contained several types of data problems—missing data, implausible data values, undetectable incorrect data, inconsistent data collection, and inconsistent variable definitions—that made it impossible to match loan pools as precisely as could be accomplished if both data sets were perfectly clean. Because of these data problems, we reduced the set of common variables used to pool the data to eight and decided not to pool on the basis of borrower gender, coborrower race, coborrower gender, or loan balance.

9. Recently, however, both GSEs have started purchasing subprime loans, particularly those with a credit risk rating of "A–" or just below prime. It is difficult to identify subprime loans because the interest rate on the mortgage is not recorded on either the PUDB or the HMDA database. It is possible, however, to infer that some loans are subprime by identifying particular lenders as actively primarily in the subprime and not in the prime segment of the mortgage market. Randall Scheessele of HUD's Office of Policy Development and Research has performed this exercise, and the lists of prime and subprime lenders could in principle be used to exclude subprime loans from the analysis. Lenders are identified in the HMDA database but not in the PUDB data, however, so the identification of subprime loans would be approximate at best.

10. A benefit of combining the data sets is that missing data from either of the two data sets can often be filled in using the other data set; thus, a key step in our analysis is the combining of the two databases at the loan level. This step also enables the linking of fields: for example, the "Center City Location" field in the PUDB database is attached to the HMDA data as well.

11. We use all HMDA and PUDB data for home purchase loans and refinances combined. We do so because the goals themselves are measured using data on both types of transactions.

12. Many census tracts are located partly in central-city areas and partly in suburban areas; these are assigned to central-city or suburban areas based on the location of the majority of the census tract.

13. As noted in the previous section, the two data sets provide two different types of loan balance data. The HMDA data set includes the initial loan amount (that is, the unpaid balance as of the date of origination of the loan), while the PUDB includes the acquisition unpaid principal balance, which is the unpaid balance as of the date that the loan was acquired by one of the GSEs. We use the term *loan balance* generically to stand for both concepts. Also, we present results for the total amount (aggregate loan balance) of loans rather than the number of loans because loan balance better reflects the distribution among market segments of the aggregate amount of mortgage finance being funded by the GSEs and by the non-GSE portion of the mortgage

market. We also performed the analysis separately according to number of loans rather than aggregate loan balance, and the results were essentially identical to those reported.

14. Sound financial condition is, of course, a concern for all market participants, not just the GSEs. Furthermore, the mere fact that loans were originated suggests that the originators found them to present no more than some maximum level of risk. However, Federal sponsorship of the GSEs demands that they maintain a level of financial safety and soundness that may well exceed that sought by many mortgage market participants that have no relationship with the Federal Government.

15. Market shares were considered *approximately equal* if they rounded to the same percentage with no decimal places. Note that the exhibits present (as closely as possible within data constraints) population figures rather than sampling results, so that confidence intervals are not appropriate.

16. It is useful to point out that it is possible for a GSE to have a relatively high (or relatively low) market share in one segment while its market share in the other segment is approximately the same as its overall market share; indeed, it is possible for a GSE to have a relatively high (or low) market share in both market segments, as is the case for New York, San Diego, and Seattle in exhibit 6. This is so for two reasons. First, the relative market shares are rounded, so that rounding error could explain the seeming paradox, especially if (as with race) the market segments are of very different sizes. Second, and more important, loans in which the market-segment variable was missing are included in the denominator in calculating market segments, and this could cause market shares for both market segments to be more (or less) than the overall market share. For example, in New York City, loans totaling 21 percent of the market were missing data on borrower race. Freddie Mac's overall market share in New York was 13 percent, but its share of the market for loans with missing data for borrower race was just 9 percent. Because of this, Freddie Mac had relatively high market shares for both market segments at 14 percent for both White and minority borrowers.

17. Geographically targeted census tracts comprise those tracts with median family income not more than 90 percent of area median income or with minority concentration of at least 30 percent and tract median income not more than 120 percent of area median income.

18. The difference appears most clearly for Fannie Mae in the lower income market segments of borrowers with incomes between 50 and 110 percent of area median, and in the higher income segments of borrowers with incomes at least 150 percent of area median. For Freddie Mac, the difference appears most clearly in the lower and middle-income segments of borrowers with incomes between 70 and 150 percent of area median and in the higher income segments of borrowers with incomes at least 250 percent of area median.

19. In estimating the regression equation we apply a weight for each observation equal to the aggregate balance of loans in that loan pool. As we did for the cross-tabulations, we also estimated the regressions using the number of loans in each pool instead of aggregate loan balance as the weighting variable; the results were virtually the same as those reported.

20. Since the pools are of different sizes, in conducting our regressions we weighted the pools by their aggregate loan balance (CONPLAMT) so that larger pools would receive more weight in the empirical results. We also conducted a second set of

regressions, identical to the first except that we weighted the pools according to the number of loans in the pool (HMDACONT) rather than the aggregate loan balance. The empirical results were essentially identical and this discussion focuses only on the results weighted by loan balance.

21. Because census tracts are defined as geographically targeted or not on the basis of their tract median income and tract percentage minority, we expect the variable indicating geographic targeting to be highly correlated with the tract income and tract percent minority variables. We nevertheless choose to include geographic targeting as an additional variable for two reasons. First, the large number of observations available for our empirical analysis suggests that any relationship between the dependent variable and geographic targeting that remains after controlling for tract income and tract percent minority may be estimated with enough precision to yield parameter estimates that are statistically significantly different from zero even if the relationship is slight. Second, it is relevant for policy purposes to determine the effect of geographic targeting after controlling for the other two types of variables, since if there is no remaining relationship then the use of geographic targeting might be considered superfluous.

22. The median estimated regression coefficient is defined as that value for which half of the coefficients for that variable estimated across all regression equations are greater than that value, and half are less than that value. (For an even number of coefficients such as the 44 that we estimate, the median is computed as the average of the coefficients just above and below it.) The mean coefficient value is the average of the coefficient values estimated for that variable across all regression equations. Both the median and the mean can be used to discuss the "typical" coefficient value estimated for each variable across the equations, but the median may be more useful since the mean is sensitive to large deviations. For example, if the coefficient estimated for a particular variable is positive in the majority of regression equations, then the median will also be positive; the mean, however, may be negative if the coefficient estimated for one or a few of the regression equations is sufficiently negative to offset the positive values estimated in the majority of equations. The standard deviation is the average of the squared differences between the actual coefficient values for that variable estimated across all regression equations and the mean coefficient value for that variable; it is a measure of the degree of variation among estimated coefficient values.

23. The formula for Somers' D is $D = (nc - nd) / t$ where nc = the number of concordant pairs, nd = number of discordant pairs, and t = total number of pairs with differing outcomes. D can take on values from –1 to +1 inclusive.

24. Exhibit 21 also shows that the standard deviation of the percent concordant figures was just 5.2 percent and the standard deviation of the Somers' D statistics was just 0.107, indicating relatively little variation in the predictive ability of the model across MSAs.

25. Loans to borrowers in the highest income category (the excluded category, those with incomes exceeding twice the MSA median) are expected to mainly consist of jumbo loans exceeding the conforming loan limit; therefore, we expect the GSE share of these loans to be relatively small.

26. Again, it can be expected that loans in the highest income tracts are less likely to be purchased by the GSEs than are loans in other tracts, because loans in the highest income tracts are more likely to exceed the conforming loan limit. Thus, it is not

surprising that loans in medium-income tracts are more likely to be purchased by the GSEs.

27. This disagreement between the results of the cross-tabulations presented in "Cross-Tabulations of Loan Purchase Activity in Different Market Segments" and the logistic regression results presented in this section may be attributable to correlation among the characteristics used to define market segments. Specifically, if two characteristics are highly correlated, such as tract percentage minority and borrower race (White/minority), then cross-tabulations performed separately using each characteristic will tend to suggest the same conclusion: even if only one of the correlated characteristics is associated with actual differences in lending patterns, the cross-tabulations will suggest an association with both characteristics. The multiple regression technique, however, attempts to attribute to each characteristic only that portion of the total variation in lending patterns associated with variations in that characteristic and not with variations in any other characteristic. The presence of correlations among the explanatory characteristics makes it important to use caution in interpreting the results, since the regression technique may have difficulty apportioning correctly the total variation in lending patterns to the two correlated characteristics. With this caveat, however, the presence of correlations also reinforces the importance of using regression analysis to control simultaneously for several characteristics, in order to evaluate the contribution of each characteristic separately in "explaining" the overall variation in lending patterns.

28. The logistic regression estimated for each MSA also included the other variables included in the previous logistic regressions: average loan balance categories (UPB_C1X, UPB_C2X, UPB_C3X, UPB_C4X, PtPlFml2, PctOwnOc, Pct_HmPr, Pct_Cvtl, Pct_FHA, Pct_FDIC, Pct_FRB, Pct_HUD, Pct_NCUA, and Pct_OCC).

29. Strictly speaking, exhibit 28 does not present coefficient estimates, but rather presents the median of estimated coefficients across all 44 MSAs included in the analysis. If we understand the medians to be representative of the actual estimated regression coefficients, we can treat them as estimated coefficients for the purpose of interpreting the results.

30. For the logit regression estimated across MSAs shown in exhibit 29, the baseline is also defined to be loans in the Anaheim MSA, and the probability of GSE purchase for loans in any other MSA is computed by adding the coefficient estimated for that MSA. The remainder of the discussion applies, however, equally to the median regression coefficients shown in exhibit 28, so the contribution of the MSA dummies is ignored in the discussion. To ease the comparison of GSE purchase probabilities across loans with different characteristics, we normalize the value of the dependent variable to zero for the baseline loans. This does not mean that the probability of GSE purchase for the baseline loans is zero; rather, it means that the incremental contribution of the discrete variables of greatest interest to the predicted GSE purchase probability is zero for the baseline loans. That is, the probability of GSE purchase for these loans can be estimated using the estimated intercept and the estimated coefficients of the variables (UPB_C1X, UPB_C2X, UPB_C3X, UPB_C4X, PtPlFml2, PctOwnOc, Pct_HmPr, Pct_Cvtl, Pct_FHA, Pct_FDIC, Pct_FRB, Pct_HUD, Pct_NCUA, and Pct_OCC) that are of less interest alone. The probability of GSE purchase for any other group of loans can then be estimated using the same coefficients plus one or more additional coefficients from the discrete variables of greater interest. Thus, the incremental contribution of the nonexcluded variables to the predicted GSE purchase probability is greater than or less than zero, depending on whether the sum of the appropriate coefficient estimates is positive or negative.

31. The incremental contribution is computed as $e^{0.020}-1 = 0.02$, or 2 percent.

32. The incremental contributions are computed as $e^{-0.226}-1 = -0.202$ and $e^{-0.109}-1 = -0.103$, respectively.

33. More precisely, it represents the difference in the predicted value of the dependent variable that in combination with values for the other variables included in the regression can be used to compute the difference in predicted GSE purchase probability.

34. $e^{3.178} = 24.00$, $e^{3.735} = 41.89$, $e^{4.350} = 77.48$, and $e^{4.568} = 96.35$.

35. Note that the 1996 recovery does not bring the probability back to the level of 1994 because the 21-percent increase from 1995 to 1996 is applied to a much smaller base than the 18-percent decline from 1994 to 1995.

36. The clustering analysis described here is an application of the k-means method used by Jesse Abraham, William Goetzmann, and Susan Wachter (1994) and Goetzmann and Wachter (1995) in studying residential house price indexes and commercial rent indexes. Specifically, we formed a matrix of estimated regression coefficients in which each row corresponded to 1 of the variables included in the regression analysis, and each column represented 1 of the 44 MSAs included in the analysis. We then randomly selected cluster *seeds* and used the *k-means* clustering algorithm to allocate each MSA to the cluster that would minimize the distance between the MSA and the cluster centroid, where the distance is measured along dimensions equal to the number of logistic regression coefficients estimated for each MSA. Because the cluster assignment is sensitive to the seed selected, we repeated this process 10,000 times using different randomly selected seeds, and then allocated MSAs to clusters based on the frequency with which each pair of MSAs clustered together across all iterations.

37. The estimated regression coefficients summarized in exhibit 38 are the same that were summarized in the left (median) column of exhibit 21, but the medians in exhibit 38 are computed only across the MSAs in each cluster rather than across all 44 MSAs included in the analysis.

38. Data on median sales price of existing single-family homes provided by National Association of Realtors, August 12, 1998.

39. Except for those indicating whether the loans are conventional or FHA-insured (Pct_Cvtl and Pct_FHA). These two variables, however, are better considered as *nuisance variables* than as variables of interest. Indeed, the differences between the median values of the two variables are almost identical across the two groups of MSAs, suggesting that there is actually no systematic difference among cities in the two clusters with respect to the tendency of the GSEs to purchase conventional versus FHA loans.

References

Abraham, Jesse M., William N. Goetzmann, and Susan M. Wachter. 1994. "Homogeneous Groupings of Metropolitan Housing Markets," *Journal of Housing Economics* 3:186–206.

Berkovec, James, and Peter Zorn. 1996. "Measuring the Market: Easier Said than Done," *Secondary Mortgage Markets* (Winter):18–21.

Bunce, Harold L., and Randall M. Scheessele. 1998. *The GSEs' Funding of Affordable Loans: A 1996 Update.* Working Paper No. HF–005. Washington, D.C.: U.S. Department of Housing and Urban Development, Office of Economic Affairs.

———. 1996. *The GSEs' Funding of Affordable Loans.* Working Paper No. HF–001. Washington, D.C.: U.S. Department of Housing and Urban Development, Office of Economic Affairs.

Canner, Glenn B., Wayne Passmore, and Dolores S. Smith. 1994. "Residential Lending to Low-Income and Minority Families: Evidence From the 1992 HMDA Data," *Federal Reserve Bulletin* 80:79–108.

Goetzmann, William N., and Susan M. Wachter. 1995. "Clustering Methods for Real Estate Portfolios," *Real Estate Economics* (Fall): 271–310.

Manchester, Paul B. 1998. *Characteristics of Mortgages Purchased by Fannie Mae and Freddie Mac: 1996–97 Update.* Housing Finance Working Paper No. HF–006. Washington, D.C.: U.S. Department of Housing and Urban Development, Office of Policy Development and Research.

Manchester, Paul B., Sue George Neal, and Harold L. Bunce. 1998. *Characteristics of Mortgages Purchased by Fannie Mae and Freddie Mac, 1993–95.* Housing Finance Working Paper No. HF–003. Washington, D.C.: U.S. Department of Housing and Urban Development, Office of Policy Development and Research.

Scheessele, Randall M. 1998. *HMDA Coverage of the Mortgage Market.* Housing Finance Working Paper No. HF–007. Washington, D.C.: U.S. Department of Housing and Urban Development, Office of Policy Development and Research.

Segal, William, and Edward J. Szymanoski. 1997. *The Multifamily Secondary Mortgage Market: The Role of Government-Sponsored Enterprises.* Housing Finance Working Paper No. HF–002. Washington, D.C.: U.S. Department of Housing and Urban Development, Office of Policy Development and Research.