

Assessing Problems of Default in Local Mortgage Markets

Prepared for:
U.S. Department of Housing and Urban Development
Office of Policy Development and Research

Prepared by:
Robert F. Cotterman
Unicon Research Corporation
Santa Monica, CA

Under Contract: C-OPC-18484

March 2001

Revised September 30, 2000

The contents of this report are the view of the contractor and do not necessarily reflect the views or policies of the U.S. Department of Housing and Urban Development or the U.S. Government.

Preface

This study was undertaken to address concerns raised about concentrations of FHA defaults in neighborhoods and among lenders in a 1997 study by the National Training Information Center (NTIC). This report was completed in two phases between December 1997 and September 2000. The first phase culminated with a March 1998 report which examined whether FHA defaults were concentrated among a group of high-default neighborhoods and high-default lenders. The report used statistical analysis to distinguish between patterns caused by chance from those attributable to specific factors such as the loan-to-value characteristics of the loan. After controlling for these factors, the statistical analysis found evidence of non-random default concentrations, at a substantially lower scale than that found in the more limited NTIC analysis. Moreover, the particular neighborhoods and lenders identified as high default changed from year to year, suggesting transitory causes which might not be amenable to policy changes. Finally, it was thought that control for differences in applicant credit histories, which was not possible for the first phase of the analysis, might explain the remaining non-random concentrations.

The second phase of the study replicated a number of the initial analyses with the addition of credit history data. This report combine the first phase of findings with the new second phase of findings, which are presented separately in Appendix C. As expected, the updated analysis shows that high default neighborhoods and lenders have more borrowers with poor credit. Including credit data in the analysis reduces the differential default rate among neighborhoods and lenders, but not as much as might be expected. However, non-random concentrations of default remain even after controlling for differential credit histories, and thus the findings of the original study still hold, albeit at a smaller scale.

TABLE OF CONTENTS

PREFACE TO REVISED REPORT	iii
.....	
SUMMARY OF FINDINGS	iv
.....	
SECTION 1	
INTRODUCTION	1
1.1. Purpose and Methodological Approach	1
1.2. Organization of the Report	2
SECTION 2	
DATA AND DEFINITIONS	4
2.1. Data Construction and Data Sources	4
2.2. Preliminary Matters: The Identification of Defaults and Observation Intervals	5
2.2.1. Timing Issues	5
2.2.2. Definitions of Default	6
SECTION 3	
DISPARITIES IN DEFAULT RATES ACROSS AREAS AND LENDERS	9
3.1. The Geographic Concentration of Defaults	9
3.1.1. Tract Default Rates Relative to MSA Default Rates	9
3.1.2. Variation Across Definitions and Across Years	13
3.2. Characteristics of Tracts with High Relative Default Rates	14
3.2.1. Attributes of Borrowers and Loans in High-default Tracts	14
3.2.2. FHA vs. Conventional Lending in High-Default Tracts	16
3.2.3. Other Attributes of High-default Tracts	17
3.2.4. Default Rates in Low Income and Minority Tracts	18
3.3. Are Defaults Concentrated in Particular Lenders?	19
3.4. Characteristics of High-default Lenders	21
3.4.1. Does Lender Performance Vary Across Areas?	21
3.4.2. Attributes of Borrowers, Loans, and the Population Served by High-default Lenders	23
3.5. Conclusions	25
SECTION 4	
PRELIMINARY STATISTICAL ANALYSIS OF DISPARITIES IN DEFAULT RATES	26
.....	
4.1. Statistical Tests of Geographical Differences	26
4.1.1. Chi-Square Tests of Geographical Differences	27

4.1.2. Exact Probability Calculations	28
4.2. Statistical Tests of Differences Across Lenders	32
4.2.1. Chi-Square Tests of Differences	32
4.2.2. Exact Probability Calculations	32
4.3. Comparisons with NTIC Designations	33
 SECTION 5	
STATISTICAL ANALYSIS OF DATA ON INDIVIDUAL LOANS	39
5.1. Analysis of Defaults	39
5.1.1. Explanatory Variables	40
5.1.2. Logit Estimates of Pooled Data	41
5.1.3. Logit Results for Individual MSAs	42
5.1.4. Possible Reasons for Area and Lender Effects	43
5.1.5. Additional Comparisons with the NTIC Findings	45
5.2. A Tentative Analysis of Lender Responses to Delinquencies	47
 SECTION 6	
A SUMMARY OF THE METHODOLOGY AND GENERAL FINDINGS	49
6.1. Choosing Time Periods, Default Definitions, and Areas	49
6.2. Statistical Analysis of Default Probabilities	50
6.3. Statistical Analysis of Lending or Servicing Practices	52
 REFERENCES	54
 APPENDIX A	
LIST OF TABLES	A-1
 APPENDIX B (Available from HUD)	
MSA-SPECIFIC RESULTS	B-1
B.1	B-2
B.2	B-9
B.3	B-40
B.4	B-71
 APPENDIX C	
A SEQUEL	C-1

PREFACE TO REVISED REPORT

The vast majority of this paper was completed in March 1998 using then-available data on individual FHA-insured loans. After data on credit scores for many of these loans subsequently became available, selected analyses were rerun to incorporate these credit scores. Appendix C has been added to the original paper to present the findings obtained after including the credit scores. Nothing substantive other than this appendix has been changed. In particular, the original Summary of Findings and the complete body of the paper have been left in their original form.

SUMMARY OF FINDINGS

The purpose of this study is to help FHA understand the geographical dimension of default behavior by examining concentrations of defaults of 1992 and 1994 loan originations in 22 urban areas. More specifically, the paper asks, first, whether defaults on FHA-insured loans are concentrated within a distinct set of high-default neighborhoods and, second, whether defaults on FHA-insured loans are concentrated within a set of high-default lenders. A heavy concentration of defaults in certain areas or among certain lenders could reveal problems that are amenable to policy solutions. Such problems might include overly generous underwriting standards, lax or fraudulent application of existing underwriting criteria, or inappropriate servicing of delinquent borrowers. Alternatively, heavy concentrations of defaults may occur because of chance alone, suggesting that the reasons may be fleeting and not amenable to policy changes. The findings in this study regarding the extent and implications of default concentrations differ from those presented in a recent study by the National Training and Information Center (NTIC), which has received much public attention.

Given the numerous possible causes of concentrations of default among areas and lenders, along with correspondingly divergent remedies, it is important to identify the reasons for concentrations of defaults, and it is especially important to identify causes that are likely to respond to policy changes, and those that are not. Because statistical analysis permits the calculation of the probability that chance alone (which is typically beyond the reach of any remedial policy) could be responsible for observed levels of default activity, this study adopts statistical analysis as the primary tool for deciding whether concentrations of defaults in specific areas or among certain lenders should be a source of concern.

Although three different measures of default are entertained in this study, a primary measure includes both claims paid and 90-day delinquencies in progress that are not observed to cure by the end of the observation period. Most of the latter delinquencies go uncured for at least nine months following the recording of the 90-day delinquency; allowing this much time to pass without observing a cure helps ensure that the delinquency is on the way to claim. The latter delinquencies thus tend to be more serious than the set of all 90-day delinquencies, most of which cure rather than proceed to claim.

MAIN CONCLUSIONS

The study reaches the following five principal conclusions:

- **By serving less affluent borrowers, FHA extends home ownership to those who are less well served by the conventional market. In neighborhoods where less affluent borrowers predominate, FHA assumes an especially important role, but default activity is more common as well. Putting further restrictions on FHA borrowers will reduce default rates but will also work against extending homeownership.**

- **Some of the differences in default rates across neighborhoods and lenders are plausibly traceable to characteristics of the borrowers and loans. Borrowers in neighborhoods and among lenders with high default rates are more frequently first time homebuyers and are more often black, have higher loan-to-value ratios, lower incomes, and smaller values of assets after closing than do borrowers in neighborhoods and among lenders with low default rates.**
- **Although low incomes are associated with higher default rates, income does not completely determine default behavior. Many neighborhoods with low incomes or substantial minority representation have default rates that are below the metropolitan area average.**
- **Simple statistical analysis identifies a set of high-default neighborhoods and a set of high-default lenders, though far fewer neighborhoods and different lenders than are identified using the NTIC methods. Removing the influence of a variety of default-related factors with more sophisticated techniques generally reduces the estimated effect on default of residence in a high-default neighborhood and origination by a high-default lender, but there still appear to be some high-default neighborhoods and high-default lenders in most of the urban areas examined in this study. It is unclear what factors are responsible for these differences in default rates, but differences in credit history may play a role.**
- **The identification of high-default neighborhoods and high-default lenders varies with the loan origination year, indicating that some problems generating high default rates are temporary. Transitory causes of high default rates are less important to treat and are less amenable to remedial action.**

The NTIC study uses data on twenty Metropolitan Statistical Areas (MSAs), ten of which are included here as well, to address some of the same basic issues. The NTIC study draws different conclusions and employs a different methodology than that used here. For example, the NTIC study identifies high-default neighborhoods solely by comparing the default rate of the neighborhood to that of the metropolitan area as a whole, and it identifies poorly performing lenders as those with the largest number of defaults in the metropolitan area, regardless of loan volume. These non-statistical methods lead to improper identification of high-default lenders, to labeling of neighborhoods as high-default even when causes appear to be transitory, and to overzealous labeling of neighborhoods as high-default (about 7 percent of areas are so identified using statistical methods in this study, as compared with 24 percent of areas using the NTIC method). Not only does the NTIC study fail to adhere to commonly accepted statistical practices in its analysis of raw default rates, it fails to consider whether other default-related factors might vary with, and perhaps account for, the default rate of the area or lender. The methodology and findings of the current study are contrasted to those presented in the NTIC report at various points in the text.

SPECIFIC FINDINGS

Through both independent study and comparisons with the NTIC approach, this paper reaches the following conclusions:

- **A comparison of census tract-level default rates to MSA-level default rates, without regard to the level of lending activity within the tract, can be misleading and, if used to guide the application of remedial activity, unproductive as well.**

By ignoring the role of randomness, simple comparisons of census tract-level default rates to those at the MSA level can mislead in either direction. On the one hand, a tract with a high default rate on very few loans may receive attention even though it is quite likely that the observed level of defaults is due to chance alone. On the other hand, tracts with many loans and a default rate that exceeds the metropolitan area rate by a moderate amount may go unnoticed even though there is little chance that such a level of defaults could be traceable to randomness. For these reasons, using such simple comparisons to prescribe further investigation or intervention will result in an incorrect focus. In addition, tracts embedded in metropolitan areas with very low default rates (*e.g.*, Denver) may be singled out as high-default tracts even though they have default rates that are, as a practical matter, too low to be of real concern. Moreover, the tendency to give disproportionate attention to tracts with small loan volume also results in a focus on tracts where the potential gain from remedial action is small as well. That is, all else the same, low-volume tracts offer smaller possible gains from a given reduction in the default rate as a consequence of remedial action.

- **The NTIC method of classifying lenders as poor performers by looking only at the volume of defaults unfairly penalizes large lenders and misses potential problems in smaller lenders.**

A high volume of defaults could be traceable to high loan volume alone and, in and of itself, implies nothing about the selectivity of the lender or its policies in handling delinquencies. In the data used in this study, applying the NTIC method tends clearly to select large lenders, some of which have default rates that are lower than the rate in the metropolitan area as a whole.

- **Default rates vary substantially across tracts and lenders within an MSA. For certain census tracts and lenders, rates are high enough that chance alone is unlikely to be the explanation; systematic factors are probably at work.**

Although we emphasize that one should not draw conclusions on the basis of a simple comparison of the default rate for a tract or lender with the default rate for the MSA as a whole, default rates do in fact vary across tracts and across lenders. The first row in the summary table¹ below gives some indication of the variation in default rates across tracts (Panel A) and lenders (Panel B). Here tracts

¹ Both the summary table and all other references to specific numbers within this section refer to a default measure that includes claims paid and uncured delinquencies in progress at two years following origination. The text considers two other definitions of default as well.

SUMMARY TABLE

CHARACTERISTICS OF TRACTS AND LENDERS IN VARIOUS RELATIVE DEFAULT RATE CLASSES
(TRACTS OR LENDERS WITH MORE THAN 30 LOANS)

PANEL A: TRACTS

Row Number	Characteristic	Default Rate of Tract Relative to MSA Rate				
		0 to < 0.5	0.5 to < 1.0	1.0 to < 1.5	1.5 to < 3.0	3.0+
1	% of All Tracts in Default Rate Class	40%	20%	17%	17%	6%
2	% of All Loans in Default Rate Class	34%	24%	20%	18%	4%
3	% FHA Black	6	10	15	24	34
4	% First Time	42	44	45	48	52
5	% LTV .97 +	23	23	25	31	39
6	Income-\$ Difference from MSA average	123	90	-10	-237	-657
7	Assets-\$ Difference from MSA average	881	508	-253	-1497	-3633

PANEL B: LENDERS

Row Number	Characteristic	Default Rate of Lender Relative to MSA Rate				
		0 to < 0.5	0.5 to < 1.0	1.0 to < 1.5	1.5 to < 3.0	3.0+
1	% of All Lenders in Default Rate Class	30%	28%	21%	17%	4%
2	% of All Loans in Default Rate Class	18%	39%	28%	13%	1%
3	% FHA Black	10	11	14	20	33
4	% First Time	40	43	48	51	60
5	% LTV .97 +	22	25	25	31	39
6	Income-\$ Difference from MSA average	53	54	-19	-163	-451
7	Assets-\$ Difference from MSA average	562	342	-145	-1091	-3287

or lenders in each MSA are classified according to the ratio of the tract (or lender) default rate relative to the default rate in the MSA as a whole. The first row of Panel A shows the percentage distribution of tracts across relative default rate classes, while Panel B presents the corresponding distribution of lenders across relative default rate classes. The second row of each panel shows how loans are divided up among the corresponding groups of tracts or lenders.

In the MSAs examined here, standard statistical tests show that, using both origination years together, about 5.2 percent of tracts and about 5.7 percent of lenders can be classified as “high-default” according to conventional standards.² These percentages vary widely across MSAs and across origination years. For example, using both origination years together, 9.6 percent of tracts in the Memphis MSA are labeled as high-default by the statistical methodology employed here, while only 2.4 percent of the Sacramento, CA PMSA tracts are so labeled. About 9.3 percent of lenders in the Fort Worth-Arlington, TX PMSA are labeled as high-default, but only 3.1 percent of the lenders in the Sacramento, CA PMSA.

- **The identification of high-default census tracts and high-default lenders varies with the loan origination year, indicating that some problems generating high default rates are temporary.**

Whether one identifies high-default tracts and lenders by simply looking at the default rate relative to the metropolitan area average, or by conducting formal statistical tests, the identification of particular tracts or lenders as high-default depends on the loan origination year; and this is true even if defaults are recorded at a given number of years following loan origination. For some MSAs, there is virtually no overlap in the tracts or lenders identified as “high-default” in the two origination years (1992 and 1994) used in this study. For most MSAs, less than one percent of tracts are labeled as high-default in both origination years, while for the vast majority of MSAs, less than three percent of the lenders are labeled as high-default in both years. This finding suggests that some of whatever is captured in identifying a “high-default” tract or lender is transitory. Not only are truly transitory fluctuations presumably less important to remedy, they may also be less susceptible to remedial action since they may require anticipation on the part of monitoring agencies. That is, to identify and cure a problem that appears only sporadically, one may need to be able to predict when the problem will arise. For these reasons, it is not at all clear that policy should be altered in an attempt to contend with these transitory problems or that policy could successfully do so.

- **Borrowers in tracts and lenders with high default rates are more frequently first time homebuyers and are more often black, have higher loan-to-value ratios, lower incomes, and smaller values of assets after closing than do borrowers in tracts and among lenders with low default rates.**

² These calculations pertain to tracts or lenders with two or more loans in the two origination years together. When restricted to tracts and lenders with more than 30 loans in the two years together, 7.2 percent of tracts and 9.6 percent of lenders are classified as high-default.

Rows 3 through 7 of each panel in the summary table illustrate this point by showing the attributes of loans among those tracts or lenders falling into each of the relative default rate categories. These rows show, in order, that the percentages of borrowers who are black, who are first-time homebuyers, and who have a loan-to-value ratio exceeding 97 percent, are all higher in tracts with higher relative default rates, as well as among lenders with higher relative default rates. For example, while 23 percent of borrowers in tracts in the lowest default rate category have loan-to-value ratios of at least 97 percent, the figure rises to 39 percent of borrowers in tracts with the highest default rates. The fraction of borrowers who are black is only 6 percent for tracts in the lowest default rate category, but rises to 34 percent for tracts in the top default rate category. The bottom two rows (rows 6 and 7) show that borrowers in tracts or lenders with higher relative default rates have lower incomes and smaller asset levels when compared with MSA averages. For example, average monthly incomes are \$123 above the MSA average for borrowers in tracts that are in the lowest default rate category, but average monthly incomes for borrowers in tracts in the highest default rate category are \$657 below the MSA average.

Taken as a whole, these findings suggest that the observed differences in default rates across tracts or lenders should perhaps not be surprising. FHA promotes homeownership by serving those who are not well served by the conventional market. This mission places FHA in a position in which it would be expected to attract borrowers who have higher default probabilities, and we should not be surprised to find differences in the distribution of these borrowers across areas and lenders.

- **Although tracts with high default rates tend to have borrowers with lower incomes than in the MSA as a whole, many low income or high minority tracts have default rates that are below the MSA average.**

It is worth emphasizing that while there appear to be relationships between default rates of tracts (and lenders), on the one hand, and borrower income and related characteristics, on the other hand, it is not at all unusual to find low income or minority tracts with relatively low default rates. We find, for example, that among tracts³ that are 30 to 50 percent minority, about 45 percent have default rates that are below the MSA average. Among tracts with median family incomes that are no more than 80 percent of the MSA median, we find that 40 percent of such tracts have default rates that are below the MSA average. Thus, many tracts with substantial minority populations or low incomes, which are traditionally viewed as portions of the underserved population that FHA attempts to aid, still have relatively low default rates.

- **The fraction of loans that are FHA-insured is greater in tracts with higher default rates, but even in tracts with high default rates the FHA share of the market is under 50 percent.**

³ Figures in this paragraph refer to tracts with more than 30 loans, which in turn contain over 90 percent of the FHA loans in these MSAs.

By a variety of measures, tracts with higher default rates tend to be poorer, and it is not surprising to find that FHA-insured loans have a more substantial market share within such tracts. The FHA is not intended to displace conventional lending but is instead intended to expand opportunities for home ownership. Even within tracts and lenders that exhibit relatively high default rates, however, the FHA share of the market is under 50 percent. Thus, even in areas where there are relatively high default rates on FHA loans, conventional lending has not been driven out. The evidence is that conventional lenders find acceptable risks even in these areas.

- **Allowing for the influence of a variety of default-related factors generally reduces the estimated impact of residence in a high-default tract and origination by a high-default lender. Even after controlling for the impact of these other factors, however, there still appear to be a set of high-default tracts and a set of high-default lenders in most of the MSAs examined in this study. It is unclear what factors are responsible for these differences in default rates, but differences in credit history may play a role.**

The more sophisticated statistical analysis presented in Section 5 shows that once we account for the influence of those default-related factors that can be measured in our data, there is typically a marked decline in the effect that can be attributed to residence in a high-default tract or origination by a high-default lender. Effects remain, however, even after these statistical adjustments are made. We cannot be certain why these effects persist, but one probable ingredient is our inability to control for differences in credit history among borrowers, which stems from a simple lack of data on credit history. The reason that this omission is likely to be important is that the quality of a borrower's credit may vary, on average, across tracts and lenders; and the result may well be that differences in default rates across tracts and lenders remain even after adjusting for factors that we can observe. Other subtle statistical influences may reinforce this tendency.

The lack of data on underwriting factors like credit history also makes it impossible to ascertain whether or not lenders are following FHA underwriting guidelines by simply looking at default statistics or even by performing sophisticated statistical analyses. Underwriting guidelines permit underwriters to trade off weakness in one area for strength in another. This practice makes it impossible to tell whether an unfavorable value for one underwriting criterion that we might observe is offset by a very favorable rating in another area, like credit history, that we do not observe.

- **When compared with non-high-default lenders, high-default lenders do not appear to intervene more quickly in a delinquency, nor do they more often institute foreclosure proceedings when contending with a delinquency.**

The FHA delinquency data permit us to perform a rather limited investigation of two possible avenues by which default rates could be affected by lender servicing behavior. We look at the possibility, suggested in the NTIC study, that high-default lenders intervene more quickly in delinquencies in progress than do non-high-default lenders, and that such intervention more often takes the form of a movement toward foreclosure. The evidence on the first of these points is

entirely ambiguous, sometimes showing high-default lenders intervene more quickly, sometimes less quickly, depending on the definition of default. We next group lender interventions into two categories: either as a movement to foreclose, on the one hand, or as providing help to avoid foreclosure (through offering forbearance, for example), on the other hand. The evidence on differences across lenders is again weak and ambiguous, with no convincing evidence of any differences in the path chosen by high-default lenders versus non-high-default lenders. Thus, the possibility of overly aggressive pursuit of foreclosure on the part of high-default lenders, as suggested in the NTIC report, receives little support in the FHA data examined here.

- **The non-statistical methods employed by NTIC lead to misclassification of tracts and lenders and substantial overstatement of potential problems. That is, these methods lead to overzealous labeling of tracts as high-default and to improper identification of high-default lenders. In addition, the NTIC study does not attempt to unravel the effects of other factors on the default rates of tracts and lenders, making it impossible to judge whether there are problems that do warrant attention.**

When the statistical methods used in this study are applied to the ten MSAs that also appear in the NTIC study, we find that about 7 percent of the tracts⁴ are labeled as high-default tracts. In contrast, the NTIC method labels about 24 percent of such tracts as high-default. About 70 percent of the tracts labeled as high-default under the NTIC methodology are labeled as non-high-default in this study.

While both this study and the NTIC method single out about 10 percent of lenders in these ten MSAs as high-default lenders, the identities of the lenders so labeled are very different. The reason is that lenders with high numbers of defaults do not necessarily have default rates well above the MSA average. Sixty-three percent of the lenders identified as high-default under the NTIC methodology are labeled as non-high-default in this study. In addition, the NTIC methodology fails to identify 60 percent of the lenders labeled as high-default in the current study.

We again emphasize that the problem is not simply that the NTIC criteria select too many or too few tracts or lenders as high-default entities; even if the percentage identified is the same under the two methods, the particular tracts or lenders will generally be different. The NTIC method will single out some tracts or lenders for which chance alone is a plausible explanation for size of the default rate, but it will ignore others for which default activity is very unlikely to be a consequence of chance alone.

⁴ Calculations in this discussion use tracts or lenders with more than 30 loans in the two origination years together.

- **FHA serves less affluent borrowers, thus extending home ownership to those who are less well served by the conventional market. In neighborhoods where less affluent borrowers predominate, and thus FHA lending assumes an especially important role, one can anticipate heavier default activity. Putting further restrictions on FHA borrowers will reduce default rates but will also work against extending homeownership.**

FHA plays an especially important role in supporting the home ownership opportunities of less affluent borrowers. The extension of homeownership to such groups fosters neighborhood stability. Given the role of FHA, however, less affluent neighborhoods will tend to have a stronger FHA presence, as well as a higher level of default activity. Reducing the risk in FHA lending by raising loan qualification standards can be expected to reduce default rates, but it can also be expected to reduce FHA's ability to support the market that it has historically served. The result may be a lower default rate, but also reduced homeownership rates and thus reduced neighborhood stability. Hence, there is a tradeoff inherent in policy choices. What is important is for FHA to monitor and understand the causes for defaults so that the appropriate tradeoffs can be made in an informed manner.

Although the approach in this paper is essentially statistical, the paper opens with a purely descriptive section that examines how default rates vary across census tracts and across lenders within each of the 22 MSAs that we study. Digging deeper, we examine the characteristics of loans originated within high-default tracts and by high-default lenders, as well as the characteristics of the corresponding borrowers and of the economic environment. Following this descriptive work, attention shifts to more formal statistical tests applied to simple counts of defaults and loans. In this way we isolate tracts and lenders that may be labeled as "high-default" according to standard statistical criteria. Because this simple analysis does not account for differences among the borrowers who live in the various tracts or are serviced by the various lenders, we perform a more sophisticated statistical analysis that removes the effects of observable characteristics of loans and borrowers, thus permitting us to isolate the effects of neighborhoods and lenders.