**TRANSACTIONS OF SOCIETY OF ACTUARIES**
**1988 VOL. 40 PT 1**

# ANALYZING RECENT EXPERIENCE ON FHA INVESTOR LOANS

THOMAS N. HERZOG

ABSTRACT

This work describes a study undertaken to determine whether the Federal Housing Administration's (FHA) single-family-home mortgage insurance program for investor (nonoccupant) loans should be modified. Three probability samples of loans were drawn: one for each of endorsement years 1979, 1981, and 1983. The sample data were analyzed using both Bayesian and sample reuse procedures. The results should be of interest to private-sector financial actuaries considering this type of investment in their portfolios. The methodological approach should be of interest to a wider audience.

## I. INTRODUCTION

### 1.1 *Background*

The Federal Housing Administration (FHA) was created in 1934 to encourage improvements in housing standards and conditions, to provide an adequate home financing system, and to exert a stabilizing influence on the mortgage market in the United States. In general, FHA does not make loans or build houses, but instead operates various insurance programs under the National Housing Act. One such program, Section 203(b), provides insurance for private lenders against losses on mortgages financing single-family homes, that is one- to four-family dwellings. Thus, under Section 203(b), FHA insures such mortgages against the risk of foreclosure, which arises from the borrower's failure to continue to make his monthly mortgage payments.

When a lender causes an FHA-insured home to be foreclosed, and the home is not worth the amount still owed, the lender has the right to convey the property to FHA in exchange for insurance benefits equal to the sum of the outstanding balance on the mortgage at the time of foreclosure and expenses relating to the foreclosure and claim processes. Such a lender is said to have filed a claim (for insurance benefits) against FHA. A claim can also arise if the lender assigns the mortgage to FHA.

### 1.2 *Purpose*

This work was motivated by an earlier work [1] in which the claim rates of Section 203(b) single-family home mortgages were examined as a function of the loan-to-value ratio (the proportion of the purchase price that is financed). In that work, we found an unusually high claim rate on loans whose loan-to-value ratio was between 80.1 percent and 85.0 percent. Since this group of loans includes investor (nonoccupant borrower) loans with a minimum downpayment, we felt a study devoted to such loans was warranted. Such a study could help to determine whether FHA should modify its underwriting standards for investor loans on single-family homes and, if so, to what extent. In fact, in December 1987, the United States Congress passed legislation reducing the maximum loan-to-value ratio on FHA-insured single-family investor mortgages from 85 percent to 75 percent. President Reagan signed this legislation into law on February 5, 1988.

### 1.3 *The Loans Examined*

We restricted our attention to fully amortizing, level-payment loans having a term to maturity of 30 years, as in our earlier work [1]. Such loans include about 80 percent of FHA's single-family activity and probably an even larger percentage of its investor loans. Because we were not able to identify individual investor loans on our automated database, we needed to examine individual casebinders[1] to do so. As a result, we constructed a proxy definition for investor loans and restricted our attention to mortgages which satisfied this proxy definition. The proxy definition used was suggested by earlier studies [1, 6, and 7]. Finally, because it was expensive to go through each casebinder manually, we sampled about 6,000 casebinders on mortgages endorsed in 1979, 1981, and 1983. We believe this is the first published paper which describes a study of individual loans identified as investor loans.

### 1.4 *An Overview of Sections 2–7*

We first present the proxy definition, the sample design, and the procedure used to determine if the mortgagor is an investor. Next, we examine the claim rates and their distribution. Finally, we discuss some additional assumptions and limitations of our analysis.

[1]A casebinder is a file containing documents on the borrower's creditworthiness, the valuation of the insured property, and the endorsement of the insurance.

## 2. THE PROXY DEFINITION OF INVESTOR LOANS

The maximum loan-to-value ratio for FHA single-family home mortgages held by investors (those who are not owner-occupants) and insured prior to 1984 is 85 percent of the maximum loan-to-value ratio permitted for owner-occupants. The following table displays some typical maximum loan-to-value ratios:

| Purchase Price | Maximum Loan-to-Value Ratio | |
|---|---|---|
| | Owner-Occupant | Investor |
| $25,000 | 97% | 82.45% |
| $35,000 | 96.43% | 81.97% |
| $50,000 | 96% | 81.60% |
| $60,000 | 95.83% | 81.46% |

The proxy definition of investor loan used in this work encompasses the following loan-to-value ratio and mortgage amount combinations:

| Loan-to-Value Ratio | | Mortgage Amount |
|---|---|---|
| 80.1-83.0% | and | ≤ $35,000 |
| 80.1-82.0% | and | > $35,000 |

Of course, some mortgages on owner-occupied homes satisfy the proxy definition.

## 3. THE SAMPLE DESIGN

Three separate list frames—one for each of the endorsement years 1979, 1981, and 1983—were employed. These consisted of 7,946, 10,473, and 30,747 Section 203(b) 30-year term level-payment single-family home mortgages, respectively, that satisfied the proxy definition of an investor loan. Within each of the three frames, the mortgages were sorted in ascending order of their FHA case numbers.

Systematic probability samples were selected from each frame in the following fashion: For the 1979 endorsements, the 4th, 8th, . . ., and 7,944th mortgages were selected, yielding a total sample of 1,986 mortgages. For the 1981 endorsements, the 5th, 10th, . . ., and 10,470th mortgages were selected, yielding a total sample of 2,094 mortgages. Finally, for the 1983 endorsements, the 15th, 30th, . . ., and 30,735th mortgages were selected, giving us a total sample of 2,049 mortgages.

## 4. THE CASEBINDERS

The casebinders on each of the selected mortgages were sent to HUD headquarters from the Federal Records Center in Suitland, Maryland. Ms.

Jan Fogel of the Actuarial Branch then examined the HUD FORM 92900.1 in each casebinder to determine whether the mortgagor was an owner-occupant or an investor. The assumption was made that the mortgagor was an owner-occupant if the box labeled occupant was checked in Item 9B. The mortgagor was assumed to be an investor if one of the following boxes was checked: Landlord, Builder, Operative Builder, or Escrow Commitment. No box was checked on about 20 of the casebinders selected. These were then examined by members of the Underwriting Branch of FHA's Office of Single-Family to see if a firm decision could be made on the type of mortgagor. Using other information in the casebinders, the Underwriting Branch was able to classify all but two of the mortgagors whose casebinders were examined. Of the 6,124 cases in our three samples, 194 cases were missing according to the staff of the Federal Records Center. Of the missing 1981 endorsements, 11 were from Fresno, California and 12 were from Camden, New Jersey. Because such a large proportion of the missing cases were from two HUD area offices, we contacted the HUD staff in these locations to obtain information on these 23 cases. As a result, we were informed that all 11 of the Fresno cases were investor loans, that 8 of the Camden cases were owner-occupant loans, and that the four other Camden casebinders were missing. The remaining 175 missing cases plus the 2 unclassifiable mortgages were distributed by endorsement year as follows:

| Endorsement Year | Number of Casebinders Missing |
|---|---|
| 1979 | 39 |
| 1981 | 64 |
| 1983 | 74 |

## 5. RESULTS AND ANALYSIS

The results of the study are summarized in Tables 1 and 2. For endorsement years 1981 and 1983, actual investor loans constitute 54.4 percent and 53.7 percent of the casebinders examined, respectively. Thus, we estimate that 5,701 and 16,517 mortgages endorsed in 1981 and 1983, respectively, were actual investor loans with a loan-to-value ratio in excess of 80 percent. In both instances the claim rate on investor loans is higher than the corresponding claim rate on owner-occupant loans. Moreover, the results are consistent with our previous analysis (see Table 2) in that the observed claim rate on actual investor loans is in both instances higher than the corresponding claim rate obtained via the proxy definition of investor loan as well as the corresponding claim rate on all Section 203(b) 30-year term level payment loans.

TABLE 1

SAMPLED SECTION 203(b) 30-YEAR TERM LEVEL PAYMENT MORTGAGES SATISFYING PROXY DEFINITION OF INVESTOR LOAN

| | | Investor Loans | | | Owner-Occupant Loans | | | |
|---|---|---|---|---|---|---|---|---|
| Endorsement Year | Number of Casebinders In Sample | Number of Claims | Number of Endorsements | Claim Rate | Number of Claims | Number of Endorsements | Claim Rate | Number of Missing Casebinders |
| 1979 | 1,986 | 11 | 352 | 3.1% | 76 | 1,595 | 4.8% | 39 |
| 1981 | 2,094 | 281 | 1,105 | 25.4% | 140 | 925 | 15.1% | 64 |
| 1983 | 2,049 | 82 | 1,061 | 7.7% | 42 | 914 | 4.6% | 74 |

TABLE 2

CLAIM RATES ON SECTION 203(b) 30-YEAR TERM LEVEL PAYMENT LOANS

| Endorsement Year | Sampled Investor Loans* | Proxy Definition** | All Section 203(b) 30-Year Term Level Payment Loans |
|---|---|---|---|
| 1979 | 3.1% | 4.8% | 6.2% |
| 1981 | 25.4% | 23.1% | 16.2% |
| 1983 | 7.7% | 7.2% | 6.1% |

*From column 5 of Table 1 of this work.
**From Tables 10, 12, and 14 of Herzog and Stasulli [6].

For the 1979 endorsements, only about 18.1 percent of the sampled cases were determined to be investor loans. Thus, we estimate that only 1,437 of the 1979 endorsements are actual investor loans with a loan-to-value ratio above 80 percent. Thus, compared to the 1981 and 1983 results, few FHA investor loans were endorsed in 1979.

The 3.1 percent claim rate on investor loans endorsed in 1979 is less than both the 4.8 percent rate of the owner-occupants loans that satisfied the proxy definition and the 6.2 percent rate on all Section 203(b) 30-year term level payment loans. Why did investors do better than owner-occupants on 1979 endorsements, but worse on 1981 and 1983 endorsements? We have no definite answers, only some hypotheses and/or partial answers. First, relatively few FHA single-family investor loans were endorsed in 1979. In most parts of the country, single-family houses bought in 1979 experienced some appreciation during their first few policy years. In addition, the assumability of FHA mortgages increased their value substantially as interest rates rose sharply during the early 1980s. These factors probably helped investors more than owner-occupants, particularly poor owner-occupants in older, inner-city houses in Regions[2] 2 and 5 who may not be able to afford to maintain the major systems of their houses. Some of the houses bought in 1981 and 1983 may have been purchased because of what, in retrospect, was unfounded optimism about local housing markets. For example, the number of investor claims in Regions 6 and 9 increased from 13 and 58, respectively, on 1979 endorsements[3] to 375 and 764 on 1981 endorsements, while the number of investor loans increased by less than 75 percent in both regions. These regions included such overbuilt markets as Houston and Las Vegas. Also, in declining housing markets, investors are more likely to make rational economic decisions to default on their mortgages since they usually have no psychological attachment to the houses and may not have invested much money on decorating the house. Some unscrupulous investors in bad markets may resort to equity skimming (not making mortgage payments, but collecting rent on the property until foreclosure) to recoup some of their losses. Finally, investors are more likely to default on the mortgages of their rental properties before they default on the mortgages of their residences.

In summary, the results indicate an adverse selection problem with investors. In 1979, when house prices and interest rates were both rising, the number of investor loans and their claim rates were relatively low. However, in 1981 and 1983 when housing conditions deteriorated in many parts of the United States, the number of FHA investor loans increased as did their claim rates in comparison to those of owner-occupant loans.

## 5.1 *Regional Data*

In Tables 3 and 4, we summarize the 1981 and 1983 sample data by HUD Region. In other words, we examine the experience within the areas covered by each of HUD's 10 Regional Offices. Regions 1–3, covering the East Coast from Maine to Virginia, had relatively few claims and investor loans in both endorsement years. For the 1981 endorsements, Regions 5, 6, 9, and 10 all had investor claim rates in excess of 25 percent. In each instance,

[2]The Regions mentioned here consist of the following states:
Region 2 = New York, New Jersey, and Puerto Rico
Region 5 = Illinois, Ohio, Indiana, Michigan, Wisconsin, and Minnesota
Region 6 = Arkansas, Louisiana, Oklahoma, Texas, and New Mexico
Region 9 = Arizona, California, Hawaii, and Nevada
Region 10 = Alaska, Idaho, Oregon, and Washington.

[3]The claim experience through June 30, 1987, of 1979, 1981, and 1983 endorsements for all 10 HUD Regions and 41 HUD field offices is shown in Table 6 of Herzog and Stasulli [6].

the claim rate on investor loans was substantially higher than the corresponding claim rate on owner-occupant loans in the sample. For 1983 endorsements, investor loans seem to be doing particularly poorly in Regions 5 and 10.

TABLE 3

SAMPLED SECTION 203(b) 30-YEAR TERM LEVEL PAYMENT MORTGAGES SATISFYING PROXY DEFINITION OF INVESTOR LOAN FOR 1981

| Region | Number of Casebinders in Region | Investor Loans: Number of Endorsements | Investor Loans: Number of Claims | Investor Loans: Claim Rate | Owner-Occupant Loans: Number of Endorsements | Owner-Occupant Loans: Number of Claims | Owner-Occupant Loans: Claim Rate | Number of Missing Casebinders |
|---|---|---|---|---|---|---|---|---|
| 1 | 16 | 6 | 0 | 0.00 | 9 | 0 | 0.00 | 1 |
| 2 | 141 | 13 | 0 | 0.00 | 121 | 7 | 5.79% | 7 |
| 3 | 144 | 64 | 4 | 6.25% | 71 | 11 | 15.49 | 9 |
| 4 | 362 | 138 | 25 | 18.12 | 215 | 27 | 12.56 | 9 |
| 5 | 232 | 73 | 23 | 31.51 | 153 | 32 | 20.92 | 6 |
| 6 | 293 | 182 | 51 | 28.02 | 102 | 18 | 17.65 | 9 |
| 7 | 50 | 16 | 3 | 18.75 | 34 | 4 | 11.76 | 0 |
| 8 | 119 | 62 | 15 | 24.19 | 50 | 12 | 24.00 | 7 |
| 9 | 548 | 403 | 111 | 27.54 | 135 | 23 | 17.04 | 10 |
| 10 | 189 | 148 | 49 | 33.11 | 35 | 6 | 17.14 | 6 |
| Totals | 2,049 | 1,105 | 281 | 25.43 | 925 | 140 | 15.14 | 64 |

TABLE 4

SAMPLED SECTION 203(b) 30-YEAR TERM LEVEL PAYMENT MORTGAGES SATISFYING PROXY DEFINITION OF INVESTOR LOAN FOR 1983

| Region | Number of Casebinders in Region | Investor Loans: Number of Endorsements | Investor Loans: Number of Claims | Investor Loans: Claim Rate | Owner-Occupant Loans: Number of Endorsements | Owner-Occupant Loans: Number of Claims | Owner-Occupant Loans: Claim Rate | Number of Missing Casebinders |
|---|---|---|---|---|---|---|---|---|
| 1 | 23 | 5 | 0 | 0.00% | 18 | 0 | 0.00% | 0 |
| 2 | 93 | 34 | 0 | 0.00 | 56 | 0 | 0.00 | 3 |
| 3 | 193 | 95 | 4 | 4.21 | 94 | 0 | 0.00 | 4 |
| 4 | 388 | 228 | 14 | 6.14 | 136 | 7 | 5.15 | 24 |
| 5 | 288 | 95 | 12 | 12.63 | 186 | 11 | 5.91 | 7 |
| 6 | 238 | 140 | 15 | 10.71 | 89 | 7 | 7.87 | 9 |
| 7 | 75 | 27 | 1 | 3.70 | 47 | 0 | 0.00 | 1 |
| 8 | 167 | 85 | 8 | 9.41 | 79 | 7 | 8.86 | 3 |
| 9 | 437 | 259 | 15 | 5.79 | 164 | 8 | 4.88 | 14 |
| 10 | 147 | 93 | 13 | 13.98 | 45 | 2 | 4.44 | 9 |
| Totals | 2,049 | 1,061 | 82 | 7.73 | 914 | 42 | 4.60 | 74 |

## 6. ESTIMATING THE DISPERSION OF THE CLAIM RATE

We next estimate the dispersion of the claim rates using two distinct approaches. The first, and the one we prefer, is based on Bayes' Theorem.

The other is a frequentist approach[4] based on the jackknife statistic described in Mosteller and Tukey [10, pages 133–163]. Sample reuse methods, such as the jackknife, have recently been made popular by Efron [3]. Nevertheless, we feel that more insight into the problem at hand is gained by using the Bayesian approach of calculating the (posterior) distribution conditional on the observed data. Specifically,

1. it is instructive to think about the entire distribution,
2. we believe that the observed data are all we have to base our inferences on (in addition to our subjective prior opinions which may be quite diffuse) since we feel it does not make sense to draw repeated subsamples of our original sample as is done applying the bootstrap,
3. the Bayesian approach forces us to make explicit all of the assumptions used in our model, and
4. we can use the posterior distribution to perform a type of hypothesis testing which makes sense.

The last item is in contrast to the frequentist type of significance tests (based on the Neyman-Pearson Lemma) which Deming [2, page 272] says "have no application here or anywhere."

### 6.1 *The Bayesian Approach*

The usual assumption is that the data are realizations of a binomial distribution. Since the Beta distribution is the conjugate prior of the binomial distribution[5], we have assumed that the prior distribution is a member of the Beta family of distributions:

$$f(x;a,b) = \begin{cases} \dfrac{\Gamma(a+b+2)}{\Gamma(a+1)\,\Gamma(b+1)}\, x^a(1-x)^b & 0<x<1 \\ 0 & \text{elsewhere} \end{cases}$$

where $a > -1$ and $b > -1$. The noninformative prior of this family is obtained by setting $a = -1$ and $b = -1$. An alternative diffuse Beta prior distribution is obtained by letting $a$ be the observed cumulative claim rate of the mortgages satisfying the proxy definition of investor loan and $b = 1-a$. For example, for the 1983 endorsements $a = 0.072$ and $b = 0.928$.

[4]Hogg and Craig [8, page 2] calls this the "relative frequency approach."

[5]It follows from Bayes' Theorem that the posterior distribution is also a member of the Beta family of distributions. For a general discussion of this, see Lindley [9, pages 141–153]. For a discussion of this in an actuarial context, see Herzog [5].

Since there is little difference in the results when the noninformative prior is used in place of the above alternative prior, we restrict attention to the noninformative prior. The results are summarized in Tables 5 and 6 where we present the mode, mean, and standard deviation of the posterior (Beta) distribution for investor and owner-occupant loans, respectively. The mathematical expressions for the last three characteristics are:

$$\text{posterior mode: } (k-1)/(n-1)$$

$$\text{posterior mean: } k/n$$

$$\text{posterior standard deviation: } (k)(n-k)/(n^2)(n+1)$$

where $k$ is the number of claims observed from a sample of $n$ casebinders. In our opinion, the mode represents the best point estimate of the claim rate, although in most instances of interest here the values of the mode and mean are nearly equal.

TABLE 5

POSTERIOR DISTRIBUTION OF THE PROPORTION OF CLAIMS ON INVESTOR LOANS OF TABLE 1

| Endorsement Year | Mode | Mean | Standard Deviation |
|---|---|---|---|
| 1979 | 2.7% | 2.9% | 0.92% |
| 1981 | 25.3% | 25.4% | 1.31% |
| 1983 | 7.6% | 7.6% | 0.82% |

TABLE 6

POSTERIOR DISTRIBUTION OF THE PROPORTION OF CLAIMS ON OWNER-OCCUPANT LOANS OF TABLE 1

| Endorsement Year | Mode | Mean | Standard Deviation |
|---|---|---|---|
| 1979 | 4.7% | 4.7% | 0.53% |
| 1981 | 15.0% | 15.0% | 1.18% |
| 1983 | 4.4% | 4.5% | 0.69% |

We now illustrate how the above results could be used to perform hypothesis testing within a Bayesian framework. To test the null hypothesis, $H_0$, that the claim rate on 1981 investor loans, $I$, is greater than that on 1981 owner-occupant loans, $O$, versus the alternative hypothesis, $H_1$, that 1981 owner-occupant loans have higher claim rates than 1981 investor loans, we determine the probability that

$$I-O > 0,$$

where the Beta density function of $I$ has mean 25.4 percent and standard deviation 1.31 percent and that of $O$ has mean 15.0 percent and standard deviation 1.18 percent. Assuming $I$ and $O$ are stochastically independent and that both have approximately normal distributions, that is, that the central limit theorem applies, we find the desired probability to be almost 1.

Instead of using the normal approximation described here, we could alternatively use the approximation based on the $F$-statistic described in Chapter 7 of Lindley [9]. A third method is to carry out a stochastic simulation of the difference of two Beta random variables. Herzog [4] provides a discussion of stochastic simulation in an actuarial environment. Finally, a fourth method is discussed in Novick and Grizzle [11].

### 6.2 *The Jackknife Approach*

The second method of estimating the standard deviations of the claim rates of interest is based on the jackknife statistic described on page 135 of Mosteller and Tukey [10]. We calculated seven sets of estimates, using 7, 8, . . ., 13 (independent) replicates. For both investor and owner-occupant loans, the estimated mean claim rates were virtually identical to the values estimated under the Beta-binomial model of Section 6.1. The standard deviations estimated using 7, 8, . . ., 13 replicates are shown in Table 7. For both owner-occupant and investor loans, the jackknife estimates of the standard deviation show wide variations among themselves. For investor loans, the estimates range from 0.51 percent to 1.63 percent; for owner-occupant loans, the estimates range from 0.78 percent to 1.31 percent. We were somewhat surprised with this wide range of values. We thought that, given the wide use of the jackknife, it produced estimated standard deviations which were more stable, that is, less dependent on the number of replicates used, than were the results shown in Table 7.

TABLE 7

ESTIMATED STANDARD DEVIATIONS FOR 1981 MORTGAGES

| Type of Mortgagor | Jackknife Estimate Number of Replicates | | | | | | | Beta-Binomial Model |
|---|---|---|---|---|---|---|---|---|
| | 7 | 8 | 9 | 10 | 11 | 12 | 13 | |
| Investor | 1.14% | 0.51% | 1.63% | 1.34% | 1.07% | 1.23% | 1.22% | 1.31% |
| Owner-Occupant | 1.31% | 1.31% | 0.78% | 1.23% | 0.96% | 1.28% | 0.85% | 1.18% |

## 7. OTHER ASSUMPTIONS AND LIMITATIONS OF THE ANALYSIS

We have calculated the estimated claim rates for the various groups based only upon those casebinders which we have thus far obtained. Although the number of currently missing casebinders is low, such casebinders represent a disproportionately large number of claims:

| Endorsement Year | Number of Claims in Missing Casebinders | Number of Missing Casebinders | Claim Rate |
|---|---|---|---|
| 1979 | 5 | 39 | 12.8% |
| 1981 | 25 | 64 | 39.0% |
| 1983 | 17 | 74 | 23.0% |

As a result, the overall claim rate of each of the three samples is too low. Because the missing casebinders are not concentrated in any HUD area office, we do not have a good feel for how these cases would alter the relationship between investor and owner-occupant claim rates in any endorsement year.

In constructing the Beta distributions, we implicitly assumed that the claim rates of all the mortgages within a given endorsement year/occupant type grouping were mutually stochastically independent. In other words, we assumed that we had a simple random sample. We know that in practice this assumption is not true because claim rates are dependent on local, regional, and national economic conditions.

However, we do not believe this violation of the simple random sampling assumption is severe. In particular, using estimators $v_2$, $v_3$, and $v_4$ of Wolter [12], we believe that the standard deviation of the 1981 investor claim rate should be about 1.18 percent rather than 1.31 percent, while that of the 1981 owner-occupant claim rate should be about 1.15 percent instead of 1.18.

Since the proxy definition of an investor loan is based on the loan-to-value ratio and mortgage amount of individual mortgages, we are only able to analyze mortgages which had data on both these characteristics in our database. The percentages of mortgages in the database which could not be classified in this fashion are the following:

| Endorsement Year | Percentage of Cases Lacking Mortgage Amount and/or Loan-to-Value Ratio |
|---|---|
| 1979 | 15.4% |
| 1981 | 21.3% |
| 1983 | 18.9% |

We have assumed that these cases are missing at random, that is, that the proportion of investor loans, claims, and so on, is approximately the same among those missing as it is among the population as a whole. The estimates of the number of actual investor loans with loan-to-value ratios above 80 percent are based only on the mortgages having both characteristics present. Hence, the estimates of the number of actual investor loans should probably be increased to account for the investor loans whose loan-to-value ratios are not present in the database.

We have in the past encountered a small number of data entry problems with the database used in this study—FHA's A43 Single-Family Insurance System. We are unaware of any large-scale systematic study of the data quality of the A43 system. Our feeling is that there are a relatively low number of errors and that these have little or no impact on the results of this study. Nevertheless, this is a potential concern.

## REFERENCES


1. BAK, T.J., HERZOG, T.N., AND MIDDAUGH, D.A., "15 Million Mortgages: The FHA Experience," *Mortgage Banking*, November 1984.
2. DEMING, W.E. *Out of the Crisis*. Cambridge: M.I.T., 1986.
3. EFRON, B. "The Jackknife, the Bootstrap, and Other Resampling Plans," *SIAM*, Philadelphia, 1982.
4. HERZOG, T.N. *An Introduction to Stochastic Simulation*, Part 3 Study Note. Chicago: Society of Actuaries, 1984.
5. HERZOG, T.N. *An Introduction to Bayesian Credibility*, Part 4 Study Note. New York: Casualty Actuarial Society, 1985.
6. HERZOG, T.N. AND STASULLI, D.C., *Claims by HUD Field Office: Investor Loans vs. Non-Investor Loans*, Actuarial Branch, HUD, 1987.
7. HERZOG, T.N. AND FOGEL, J.R., *Mortgage Insurance Risk—Investor and Low-Downpayment Loans*, Actuarial Branch, HUD, 1987.
8. HOGG, R.V. AND CRAIG, A.T. *Introduction to Mathematical Statistics*. 4th ed. New York: Macmillan, 1978.
9. LINDLEY, D.V., *Introduction to Probability and Statistics from a Bayesian Viewpoint*, Part 2. New York: Cambridge University Press, 1970.
10. MOSTELLER, F. AND TUKEY, J.W., *Data Analysis and Regression*. Reading, Mass.: Addison-Wesley, 1977.
11. NOVICK, M.R. AND GRIZZLE, J.E. "A Bayesian Approach to the Analysis of Data from Clinical Trials," *Journal of the American Statistical Association* 60 (1965):81–96.
12. WOLTER, K.M. "An Investigation of Some Estimators of Variance for Systematic Sampling," *Journal of the American Statistical Association* 79 (1984): 781–90.

## DISCUSSION OF PRECEDING PAPER

ELIAS S. W. SHIU:

The dramatic growth in the mortgage-backed securities market over the past decade has stimulated enormous interest in these instruments. Currently, outstanding residential mortgage debt in the U.S.A. totals about $2 trillion, of which about $300 billion has been securitized. There is much potential for further growth in the mortgage-backed securities market. They have become a significant component in the investment portfolios of many insurance companies.

There are two kinds of risk in investing in mortgages—prepayment risk and default risk. However, mortgage-backed securities issued by the Government National Mortgage Association (Ginnie Mae) and, to a somewhat lesser extent, those issued by the Federal National Mortgage Association (Fannie Mae) and Federal Home Loan Mortgage Corporation (Freddie Mac) enjoy the full faith and credit of the U.S. Government. For such securities, default risk does not really exist as defaults become prepayments. It is my understanding that some models used on Wall Street for pricing mortgage-backed securities do not model defaults separately. One reason given is that there are no data on defaults, because investors are not told which prepayments are due to defaults. Dr. Herzog's paper has certainly illuminated the problem of default risk. His work will be welcomed by those interested in the pricing of mortgage-backed securities.

It is now widely recognized that all debt securities can be viewed as risk-free assets plus or minus various contingent claims, which can usually be modeled as options. In the case of mortgages, prepayment can be viewed as a call option, that is, an option to buy back or call the mortgage at par, and default can be treated as a put option, that is, an option to sell or put the house to the lender at a price equal to the value of the mortgage. Some researchers would use extensions of the Black-Scholes option-pricing theory [1] to price mortgage-backed securities. The Bibliography lists some articles related to the pricing of mortgage-backed securities and mortgage default risk.

My next comment is on the Bayesian approach. Lively discussions on the applicability of Bayesian methods to actuarial science can be found in Ryder [22]. Although I have no intention of criticizing Bayesian statistics, I wish to point out a theoretical difficulty in the use of conjugate families of distributions (a concept formalized by Raiffa and Schlaifer [21]). I have voiced such a criticism earlier with respect to graduation [23, p. 73]. Let me first illustrate my concern with the case that the sample observations are from a normal distribution; the variance is known; and the prior distribution for the mean is also normal. Then the parameters of the distributions are related by the formulas [17, p. 55, Theorem 2.3]

$$\frac{1}{\sigma^2_{Posterior}} = \frac{1}{\sigma^2_{Prior}} + \frac{1}{\sigma^2_{Likelihood}} \tag{1}$$

and

$$\frac{\mu_{Posterior}}{\sigma^2_{Posterior}} = \frac{\mu_{Prior}}{\sigma^2_{Prior}} + \frac{\mu_{Likelihood}}{\sigma^2_{Likelihood}} . \tag{2}$$

It follows from (1) that the variance of the posterior distribution is always smaller than each of the variances of the prior distribution and the likelihood distribution. Consequently, the posterior distribution is always more concentrated than the prior and likelihood distributions. Now, it seems reasonable that in those situations in which the prior distribution and the likelihood distribution are in conflict, with neither information source dominating the other, the posterior distribution should be fairly diffuse or, perhaps, bimodal with modes at both the prior mean and sample mean. A desirable property of a good inference model is that contradictory evidence should induce confusion. However, such posterior distributions will not arise no matter how severe the apparent conflict between the prior and likelihood distributions. The posterior variance does not depend on the prior mean or sample mean.

It is hypothesized in the paper that the data are from a Bernoulli distribution with the prior distribution of its parameter being a Beta distribution. If the prior density function is proportional to

$$x^a(1 - x)^b \tag{3}$$

and the likelihood function is proportional to

$$x^k(1 - x)^{n-k}, \tag{4}$$

then the posterior density function is proportional to

$$x^{a+k}(1 - x)^{b+n-k}. \tag{5}$$

In view of (3), (4) and (5), one might treat the prior information as if it were a previous sample [20, p. 186; 21, p. 64]. Again, the posterior distribution never distinguishes prior information from sample information, no matter how severe the apparent conflict between them; the posterior distribution is always more concentrated than the prior and likelihood distributions. Of course, such a conflict will not arise if the noninformative prior distribution is used as in the paper. For further discussion on the inappropriateness of conjugate families of distributions, see Leamer [17].

My last comment is motivated by the statement that "the mode represents the best point estimate of the claim rate." It reminded me of King's claim that the real object of graduation was "to get the most probable deaths" [15, p. 114]. Appealing to the law of large numbers, I would suggest that the mean rather than the mode is the appropriate statistic for setting insurance premiums.

## BIBLIOGRAPHY


1. BLACK, F. AND SCHOLES, M. "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy* 81 (1973):637–59. (Reprinted in BICKSLER, J.L. AND SAMUELSON, P.A., eds. *Investment Portfolio Decision-Making*. Lexington, Mass.: D.C. Heath, 1974, 181–98.)
2. CAMPBELL, T.S. AND DIETRICH, J.K. "The Determinants of Default on Insured Conventional Residential Mortgage Loans," *Journal of Finance* 38 (1983):1569–81.
3. CUNNINGHAM, D. AND HENDERSHOTT, P.H. "Pricing FHA Mortgage Default Insurance," *Housing Finance Review* 3 (1984):373–92.
4. EPPERSON, J.F., KAU, J.B., KENNAN, D.C. AND MULLER, W.J., III. "Pricing Default Risk in Mortgages," *Journal of the American Real Estate and Urban Economics Association* 13 (1985):261–72.
5. FABOZZI, F.J., ed. *The Handbook of Mortgage-Backed Securities*. Chicago: Probus Publishing, 1985.
6. FABOZZI, F.J., ed. *Mortgage-Backed Securities: New Strategies, Applications and Research*. Chicago: Probus Publishing, 1987.
7. FABOZZI, F.J. AND POLLACK, I.M., eds. *The Handbook of Fixed Income Securities*, 2nd ed. Homewood, Ill.: Dow Jones-Irwin, 1987.
8. FOSTER, C. AND VAN ORDER, R. "An Option-Based Model of Mortgage Default," *Housing Finance Review* 3 (1984):351–72.
9. FOSTER, C. AND VAN ORDER, R. "FHA Terminations: A Prelude to Rational Mortgage Pricing," *Journal of the American Real Estate and Urban Economics Association* 13 (1985):272–91.
10. HENDERSHOTT, P.H. "Mortgage Pricing: What Have We Learned So Far?" *Journal of the American Real Estate and Urban Economics Association* 14 (1986):497–509.
11. HENDERSHOTT, P.H. AND VAN ORDER, R. "Pricing Mortgages: An Interpretation of the Models and Results," *Journal of Financial Services Research* 1 (1987):19–55.

12. HERZOG, J.P. AND EARLEY, J.S. *Home Mortgage Delinquency and Foreclosure*. New York: National Bureau of Economic Research, 1970.
13. HERZOG, T.N., MIDDAUGH, D.A. AND STASULLI, D.C. "Some Recent FHA Experience on Section 203 30-Year Term and GPM Home Mortgages," *Actuarial Research Clearing House* 1986.2:15–44.
14. JACKSON, J.R. AND KASERMAN, D.L. "Default Risk on Home Mortgage Loans: A Test of Competing Hypotheses," *Journal of Risk and Insurance* 47 (1980):678–90.
15. KING, G. Discussion of SPRAGUE, T.B. "The Graphic Method of Adjusting Mortality Tables," *Journal of the Institute of Actuaries* 26 (1886):113–15.
16. KINNEY, J.M. AND GARRIGAN, R.T., eds. *The Handbook of Mortgage Banking: A Guide to the Secondary Mortgage Market*. Homewood, Ill.: Dow Jones-Irwin, 1985.
17. LEAMER, E.E. *Specification Searches: An Ad Hoc Inference with Nonexperimental Data*. New York: Wiley, 1978.
18. PETERS, H.F., PINKUS, S.M. AND ASKIN, D.J. "Default, the Last Resort." *Secondary Mortgage Markets* 1 (August, 1984):16–20.
19. PINKUS, S.M. "Mortgage-Backed Securities: An Analytical Framework." In *Controlling Interest Rate Risk: New Techniques and Applications for Money Management*, edited by R.B. Platt, 151–175. New York: Wiley, 1986.
20. RAIFFA, H. *Decision Analysis: Introductory Lectures on Choices under Uncertainty*. Reading, Mass.: Addison-Wesley, 1968.
21. RAIFFA, H. AND SCHLAIFER, R. *Applied Statistical Decision Theory*. Boston: Division of Research, Graduate School of Business Administration, Harvard University, 1961.
22. RYDER, J.M. "Subjectivism — A Reply in Defence of Classical Actuarial Methods," *Journal of the Institute of Actuaries* 103 (1976): 59–91; Discussion 92–112.
23. SHIU, E.S.W. " A Survey of Graduation Theory." In *Actuarial Mathematics,* Volume 35 of the *Proceedings of Symposia in Applied Mathematics*, edited by H.H. Panjer, 67–84. Providence, R.I.: American Mathematical Society, 1986.
24. VANDELL, K.D. AND THIBODEAU, T. "Estimation of Mortgage Defaults Using Disaggregate Loan History Data," *Journal of the American Real Estate and Urban Economics Association* 13 (1985):292–316.
25. VON FURSTENBERG, G.M. "Default Risk on FHA-Insured Mortgages as a Function of the Terms of Financing: A Quantitative Analysis," *Journal of Finance* 24 (1969):459–77.

## (AUTHOR'S REVIEW OF DISCUSSION)

THOMAS N. HERZOG:

I thank Professor Shiu for his thoughtful comments on my paper.

I agree with his suggestion that those using option-pricing models to price the yields on mortgage-backed securities should use separate models for claims and prepayments. One point of my paper is to suggest that separate claim models be developed for investor and owner-occupant loans as well.

The use of conjugate families of prior distributions is a device used to save computational efforts and costs. My philosophy on prior distributions is to give them little weight unless they are based on prior experimental results. This usually avoids the problem of having a severe conflict between the prior and sample information. Moreover, a good Bayesian analysis should demonstrate that the final results are relatively insensitive to a range of "reasonable" prior distributions.

I agree with Professor Shiu's last comment that the mean is the appropriate statistic for estimating net insurance premiums. However, in general, the optimal point estimate depends upon the purpose for which it is used. Consequently, it should be chosen by using the appropriate loss function. In particular, if the goal is to estimate the "most probable number of deaths," the mode should be used.