New Estimates of the Jumbo-Conforming Mortgage Spread

Alan S. Blinder
Department of Economics
Princeton University

Mark J. Flannery
Warrington College of Business
University of Florida

G. Brandon Lockhart
Warrington College of Business
University of Florida

Abstract

We use Monthly Interest Rate Survey (MIRS) data for April 1997 through May 2003 to estimate the effect of "conforming" status on the effective interest rate for 30-year fixed-rate mortgages. We show that plausible econometric refinements materially affect the "jumbo-conforming spread" as measured in the existing literature, and that the treatment of loan size is particularly important. We borrow from the discrimination literature to derive a new way to estimate this effect and conclude that the jumbo-conforming spread is about 25 basis points, with some evidence of decline since late 2001.

First Draft: September, 2005
This draft: January 18, 2006

We thank Jim Kamihachi for help in obtaining the MIRS data and Ron Ward for econometric advice. Brent Ambrose, Wayne Archer, David Brown, Scott Frame, Michael Fratantoni and Michael LaCour-Little provided helpful comments and suggestions. Remaining errors are our own.

Beginning in 1932, Congress created a small set of "government sponsored enterprises" (GSEs) designed to encourage home-ownership by making mortgage loans available to a wider range of citizens. The Federal Home Loan Bank System, the Federal National Mortgage Association (now Fannie Mae), and the Federal Home Loan Mortgage Corporation (now Freddie Mac) were all designed to increase the availability of mortgage loans and to reduce their cost. Partly in response to the troubles some homeowners had with balloon loans during the Depression, these institutions have emphasized the provision of long-term (now typically thirty year), fully-amortizing instruments with a fixed rate of interest and no prepayment penalty.

One major way in which the GSEs fulfill their central mission was originally to create, and now to maintain, a highly-liquid secondary market in what are called "conforming" mortgages—meaning loans eligible for sale to Fannie Mae or Freddie Mac.[1] Most researchers who have attempted to assess the effectiveness of the housing GSEs begin by estimating the difference between the effective interest rates charged on conforming versus jumbo loans, since the latter cannot be sold to the GSEs.[2] The lower interest rate on conforming loans is a starting point for measuring how much the GSEs have improved the (conforming) mortgage market. Yet estimating this premium has proven to be surprisingly difficult. McKenzie [2002] reports that previous researchers have estimated the "jumbo-conforming spread" to be as high as 60 basis points (bps) or as low as 8 bps, depending on the methodology and time period studied.

We show in this paper that plausible econometric refinements materially affect the estimated effect of GSEs on mortgage rates. Controlling for the impact of loan size on the contract rate turns out to be particularly important. Using Monthly Interest Rate Survey (MIRS) data for the period from April 1997 to May 2003, our estimates of the jumbo-conforming spread vary between 23 and 29 bps, depending on the regression specification. In our preferred

[1] The GSEs also have "affordable housing" goals, which are much more difficult to measure in market data.
[2] See for example, Ambrose *et al.* [2004], Hendershott and Shilling [1989], McKenzie [2002], Naranjo and Toevs [2002], Passmore [2003], Passmore *et al.* [2002], Passmore *et al.* [forthcoming], and Pearce and Miller [2001].

specifications, the estimated spread is at least 24 bps. We also divide the sample into two sub-periods, separated by a November 2001 regulatory change that made it less costly for banks to securitize private pools of jumbo mortgages. Logic suggests that this regulatory innovation should have reduced the jumbo-conforming spread, and we find that it did. Prior to the change, the estimated spread was 26-27 bps; afterward, it fell to 21-20 bps.

The paper is organized as follows. Section I briefly reviews the past literature and motivates the study. Section II discusses estimation issues. Section III describes the MIRS data for our selected time period. Results are presented in Sections IV through VI. The last section concludes and briefly summarizes the implications of our findings.

I. Literature Review

Two U.S. housing GSEs (Fannie Mae and Freddie Mac) have become integral to the mortgage market via their willingness to securitize 1-4 family mortgage loans below a certain principal amount, the "conforming loan limit." A government agent annually adjusts this limit to reflect changes in the average price of houses sold in the U.S.[3] Anecdotal evidence suggests that mortgagors near the conforming limit often take out a second mortgage in order to keep the bulk of their funds under the conforming limit. This phenomenon can be seen clearly in our sample of nearly one million new, fixed-rate mortgages originated between April 1997 and May 2003. Figure 1 is a histogram of these mortgage loans ranked by their size *relative to each year's conforming loan limit*. Naturally, the number of loans declines as loan size increases. But notice the spike at 100% of the conforming loan limit and the "dip" just after that. Figure 2 employs a finer scale to make both the spike and the dip more visible. Clearly, many borrowers take the maximum conforming loan, and there are few loans "just over" the limit.

[3] Prior to 1981 the conforming loan limit was established by statute. Between 1981 and 2003, the housing GSEs calculated each new conforming limit using the MIRS data assembled by FHFB. OFHEO began to calculate the new conforming limit in 2004.

The most common explanation for the concentration of loans at the conforming limit is that the GSEs create a more liquid secondary market for loans they can purchase. The GSEs provide credit enhancement for pools of qualifying mortgages, and they may add to the demand for conforming loans by purchasing the associated mortgage-backed securities (MBS) for their retained portfolios.[4] In a competitive market, the ability to sell conforming (but not jumbo) mortgages into a more liquid secondary market should reduce the effective interest rate on newly-originated conforming mortgages (Cotterman and Pearce [1996]). If so, one is left wondering about the (few) borrowers who take out mortgages just slightly above the conforming limit. Did credit problems prevent them from securing additional funds? For this reason, previous researchers have omitted such borrowers from their samples (Passmore *et al*. [2002], Passmore [2003]). We return to this issue below.

Many researchers have tried to estimate the pure, *ceteris paribus* rate premium required for non-conforming 1-4 family mortgage loans: How much more does a large borrower have to pay, after controlling for all other measurable features of her loan? Estimating this premium is the first, but not the last, step in assessing the social benefits provided by the GSEs.[5] McKenzie's [2002] excellent summary of past studies (especially his Table 1) permits our own review to be brief. With only one exception (Ambrose *et al.* [2004]), prior researchers have studied the Federal Housing Finance Board's Monthly Interest Rate Survey data (MIRS). This monthly data set contains information from a set of voluntary reporters about mortgage loans closed during the last five days of each month. The MIRS data provide no information about borrowers' creditworthiness, which complicates the identification of conforming loans. The housing GSEs purchase only high-quality loans below a certain size. Because the MIRS data do not identify "high quality" loans, we cannot observe the borrower's credit quality and hence cannot definitively identify which loans are truly conforming. The data permit us only to differentiate

[4] Critics have focused on the social value of these retained portfolios. See, for example, Jaffee [2003, 2005]).

[5] The other steps lie outside the scope of this paper. But see Blinder *et al.* [2004].

between "jumbo" and "not jumbo" loans. We take the liberty of referring to this differential as the "jumbo-conforming spread."

Hendershott and Shilling [1989] (HS) were the first to estimate the jumbo-conforming spread with MIRS data, and their methodology has influenced subsequent research. They observed that the GSEs' mortgage pool guarantees grew from "under 5 percent to over 50 percent" of the market between the early 1980s and 1986, and sought to examine the impact of this development on the jumbo-conforming spread. Using data for three months (May, June, July) in 1978 and 1986, they estimated separate pricing equations for 1978 and 1986 of the form:

$$RM_i = \alpha_0 + \alpha_1 J_i + \alpha_2 X_i + \varepsilon_i \,, \tag{1}$$

where RM_i is the i^{th} loan's effective interest rate (defined as the annual contract rate of interest plus an amortized amount (linear, over ten years) of any fees charged at closing),

J_i is a dummy variable equal to unity if the loan's principal amount exceeds the current conforming loan limit (a "jumbo"), and zero otherwise,

X_i is a vector of loan characteristics such as loan size, loan-to-value (LTV) ratio, month dummies, and a dummy for loans secured by a new home.

Within each time period, α_1 measures the jumbo-conforming spread. HS estimate this coefficient as 5 bps in 1978 and 29–39 bps in 1986. The estimated effect of the housing GSEs on the spread is therefore 24–34 bps in 1986. Many other researchers have used the same basic methodology to measure the value of conforming status (e.g., Cotterman and Pearce [1996], ICF [1990] McKenzie [2002], Naranjo and Toevs [2002], Passmore [2003], Passmore *et al*. [forthcoming], Pearce [2000], and U.S. CBO [2001]). This research provides some evidence that the spread has declined over time.[6]

Ambrose *et al.* [2004] utilize a unique private data set provided by a major mortgage originator. Unlike the MIRS data, these 1995-1997 data include an important indicator of credit

[6] Passmore *et al.* [forthcoming] estimate the jumbo-conforming difference as only 17 basis points, which is lower than most previous estimates. They "attribute [their] lower estimate to the later sample (April 1997 through May 2003), which is consistent with trends in the established literature…" (footnote 18, page 14). Our estimated spread is considerably higher even though we (deliberately) use the same time period as Passmore *et al.* More on this below.

quality, namely the borrower's FICO score, which permit the authors to identify better those observations most likely to be truly conforming mortgage loans.[7] They estimate that the spread between jumbo and non-jumbo loans is about 18 bps, and that non-jumbo loans are priced about 9 bps lower if they qualify for sale to the GSEs. Together, these imply a (conventionally-measured) jumbo-conforming spread of 27 bps—right in line with the original HS estimate.[8] However, Ambrose *et al*.'s regression specification seems to have a major flaw: Their measure of interest rate volatility consistently gets a negative coefficient, whereas theory predicts that the value of the prepayment option embedded in fixed-rate loans rises with rate volatility.[9] Without understanding the reason(s) for this result, the regression's estimated spread cannot be taken at face value.

II. Methodology

The standard model for estimating the jumbo-conforming spread has resembled Hendershott and Shilling's [1989] specification (1), which suffers from three potential flaws. First, if the variables included in the vector $\boldsymbol{X_i}$ omit determinants of the loan rate that are correlated with J, the estimate of α_1 may be biased. Subsequent research has identified additional variables that affect mortgage rates. Second, if (1) is estimated over several time periods, the explanatory variables must control for intertemporal changes in capital market conditions. Finally, the same loan characteristics might be valued differently in the markets for conforming versus jumbo loans. If so, equation (1) imposes inappropriate constraints on the coefficients.

[7] The private data set identifies which non-jumbo loans were sold to GSEs, which is a clear improvement over the MIRS. However, the data do *not* indicate whether the non-jumbo loans not sold to GSEs were eligible for purchase by a GSE. To control for this data problem, the authors code non-jumbo observations not sold to GSEs as GSE-eligible if the borrower's FICO score exceeded 620.

[8] Theoretically, an estimate like Ambrose *et al*.'s should exceed one obtained from the MIRS data, because the latter data include some low-quality loans, which presumably would carry higher effective rates.

[9] The authors use mortgage spreads (effective yield less the then-current 10-year U.S. Treasury yield) as their dependent variable. Not only does the specification impose a unit coefficient for the 10-year Treasury, it also assumes identical ten-year interest rate sensitivities for conforming and jumbo loans. Neither constraint seems justified.

Recent researchers using the MIRS data have included the following loan features in the vector ***X***:

LTV_i: A set of three dummy variables, equal to unity if the loan-to-value ratio at takedown is:
- greater than 75% and less than or equal to 80% (*LTV_1*)
- greater than 80% and less than or equal to 90% (*LTV_2*)
- greater than 90% and less than or equal to 100% (*LTV_3*).

MTGCO: A dummy variable equal to unity if the loan is originated by a mortgage company (as opposed to a bank or a thrift).

THRIFT: A dummy variable equal to unity if the loan is originated by a thrift institution (as opposed to a bank or a mortgage company).

NEW: A dummy variable equal to unity if the loan is associated with a newly-built house.

FEES: A dummy variable equal to unity if the loan's effective rate (the dependent variable) includes positive fees and charges amortized over a 10-year period.

PRIN: The loan's principle amount (divided by $1,000,000).

The ***LTV*** dummy variables presumably capture objective default probabilities. Since we omit *LTV_0* (LTV $\leq$ 0.75), we expect all three of these variables to get positive coefficients. The two institutional dummies (*MTGCO* and *THRIFT*) capture any systematic differences between loans originated by these firms, relative to those originated by commercial banks. The *NEW* variable captures any pricing differential associated with financing new homes, e.g., if the builder buys down the mortgage rate as part of the sales negotiation, or if mortgages secured by new homes realize lower default rates. The *FEES* dummy controls for any biases in the effective rate (*RM*) resulting from the simple amortization of fees over a ten-year period, which is how *RM* is constructed. Finally, the loan's principal amount (*PRIN*) should get a negative coefficient if there are economies of scale in loan origination or servicing. This effect should not be linear, and it is not clear how loan size should be measured. One common way to capture a nonlinear effect is to express the loan's principal amount as the reciprocal of *PRIN*, which we denote by *INV_PRIN*. We show in Section VI that it is important to include these scale effects in the regression specification, and that our main results are only slightly different if we substitute the (natural) logarithm of *PRIN* for *INV_PRIN*.

When (1) is estimated with data from different time periods, statistical controls for time-series variation in financial market conditions are imperative. But they have not always been used.[10] Not only should the current level of market rates affect *RM*, but a fixed-rate mortgage contains embedded interest rate and default options, which are implicitly priced.[11] These option values should vary with the yield curve's slope and/or the volatility of market rates, and hence so should the competitive mortgage rate. Another potential financial market effect on *RM* was identified by Ambrose *et al.* [2001], who found that *RM* is positively related to the volatility of house prices, presumably as an indicator of the default option's value. Some authors also include month or year dummies in (1), to capture seasonal variations in homeowners' closing date preferences or intertemporal variations in the market price of mortgage risk (e.g. Ambrose *et al.* [2004]). All this implies that a more appropriate method for estimating α_1 with a panel specification is:

$$RM_{it} = \alpha_0 + \alpha_1 J_{it} + \boldsymbol{\alpha_2} \boldsymbol{X_{it}} + \boldsymbol{\alpha_3} \boldsymbol{Z_t} + \varepsilon_{it} \ , \tag{2}$$

where $\boldsymbol{Z_t}$ is a vector of variables describing market conditions in period t. This specification recognizes the close integration of the mortgage market into other U.S. capital markets.

However, (2) retains any biases that may arise from constraining the coefficients to be identical for all mortgages, regardless of size. Both theory and empirical evidence deny that such restrictions are appropriate. For example, the more expensive homes that secure jumbo mortgages have greater price uncertainty (Ambrose *et al.* [2001]), and the GSEs require private mortgage

---

[10] As in Fama-MacBeth [1973], some researchers estimate (1) by combining the coefficient estimates from cross-sectional regressions estimated at multiple points in time. The mean estimated coefficients are then taken to indicate the true, time-invariant coefficients' values. Although we know that the mortgage market is well integrated into the capital markets, a cross-sectional regression cannot incorporate any descriptors of market-wide conditions. The estimated values of α_0 and α_1 will therefore reflect the average effect of contemporaneous market conditions on *RM*, in addition to the value of conforming status.

[11] Kau and Keenan [1995] and Ambrose *et al.* [2004] show that the value of a mortgage contract depends on two stochastic processes, interest rates and collateral values.

insurance for all purchased loans with LTV ratios greater than 80%. Both of these facts are consistent with the findings of McKenzie [2002] and Blinder *et al.* [2004] that jumbo loan rates are more sensitive to the loan-to-value ratio than are conforming rates. Jumbo loan borrowers may also be more aggressive in exercising their prepayment options, although an initial principal amount just slightly above the conforming limit may indicate credit or liquidity problems that curtail refinancing options. In short, there are many reasons to believe that the slope coefficients of a pricing equation like (2) might differ for conforming versus jumbo loans.

The impact of conforming status on *RM* is analytically similar to a question that has been studied in the literature on discrimination for thirty years: Does education have differential effects on the wage rate of Blacks versus Whites?[12] Adopting that methodology to our problem, we can allow for differential sensitivity of mortgage rates to the $\boldsymbol{X_{it}}$ vector by modifying specification (2) to get:

$$RM_{it} = \alpha_0 + \alpha_1 J_{it} + \boldsymbol{\alpha_2 X_{it}} + \boldsymbol{\beta} J_{it}\boldsymbol{X_{it}} + \boldsymbol{\alpha_3 Z_t} + \varepsilon_{it} \ . \qquad (3)$$

Here the new coefficient vector $\boldsymbol{\beta}$ measures the differential sensitivity of the pricing of jumbo loans to the various features in $\boldsymbol{X_{it}}$, and α_1 measures the effect of jumbo status *per se*--beyond any effects attributable to differential pricing. Following Blinder [1973], a still more complete separation of jumbo and conforming loans' features can be obtained by estimating separate pricing regressions for each group:

$$RM^c_{it} = \alpha_0 + \boldsymbol{\alpha_2 X^c_{it}} + \boldsymbol{\alpha_3 Z_t} + \varepsilon^c_{it} \quad \text{for conforming loans} \qquad (4a)$$

$$RM^j_{it} = \beta_0 + \boldsymbol{\beta_2 X^j_{it}} + \boldsymbol{\beta_3 Z_t} + \varepsilon^j_{it} \quad \text{for jumbo loans,} \qquad (4b)$$

The superscript "c" denotes conforming, "j" denotes jumbo. This more general specification removes the remaining constraints imposed by (3), namely that the effects of $\boldsymbol{Z_t}$ are the same in both markets and that all loans have the same error structure.

[12] The original references are Blinder [1973] and Oaxaca [1973]. For this reason, the decomposition discussed just below is often called a "Blinder decomposition" or a "Blinder-Oaxaca" decomposition.

Just as in the discrimination literature, specification (4) can be used to *decompose* the difference between jumbo and conforming interest rates into a portion attributable to the groups' *characteristics* (differences in $\boldsymbol{X}$) and a portion attributable to the differential *valuations* of the same characteristics across the two markets (differences in coefficients). This second component measures the effect of conforming status. Rewrite (4a) and (4b) more compactly as

$$RM^{c}_{it} = \boldsymbol{A}\boldsymbol{W}^{c}_{it} + \varepsilon^{c}_{it} \quad \text{for conforming loans} \qquad (5a)$$

$$RM^{j}_{it} = \boldsymbol{B}\boldsymbol{W}^{j}_{it} + \varepsilon^{j}_{it} \quad \text{for jumbo loans} \qquad (5b)$$

where the composite column vector of variables, $\boldsymbol{W}$, is $[1, \boldsymbol{X}, \boldsymbol{Z}]'$, and the composite row vectors of coefficients, $\boldsymbol{A}$ and $\boldsymbol{B}$, are respectively $[\alpha_0, \boldsymbol{\alpha_2}, \boldsymbol{\alpha_3}]$ and $[\beta_0, \boldsymbol{\beta_2}, \boldsymbol{\beta_3}]$.

Now using lower case symbols to denote means and estimated coefficients, we can subtract the expected value of (5a) from that of (5b) to obtain:

$$rm^{j}_{it} - rm^{c}_{it} = \boldsymbol{b}\boldsymbol{w}^{j}_{it} - \boldsymbol{a}\boldsymbol{w}^{c}_{it} = \boldsymbol{a}(\boldsymbol{w}^{j}_{it} - \boldsymbol{w}^{c}_{it}) + (\boldsymbol{b} - \boldsymbol{a})\boldsymbol{w}^{j}_{it} \qquad (6)$$

This equation decomposes the average jumbo-conforming spread into two parts. The first measures the portion of the mean spread that is due to differences in the typical loan characteristics of jumbo versus conforming loans, $\boldsymbol{w}^{j}_{it} - \boldsymbol{w}^{c}_{it}$, valuing those characteristics at estimated prices from the market for conforming loans ($\boldsymbol{a}$). This part of the rate differential would presumably exist even in the absence of any GSE effect. The second term in (6) is more interesting. It measures the portion of the higher yield on a jumbo loan that arises *because it is priced differently* than an otherwise-identical conforming loan. If conforming status reduces loan rates *ceteris paribus*, we expect this term to be positive.

Of course, using the price vector from the *conforming* loan market ($\boldsymbol{a}$) to value the differences in mean characteristics is arbitrary; we could just as well have used the prices from the *jumbo* market ($\boldsymbol{b}$). Doing so leads to a different, but equally valid, decomposition:

$$rm^{j}_{it} - rm^{c}_{it} = \boldsymbol{b}\boldsymbol{w}^{j}_{it} - \boldsymbol{a}\boldsymbol{w}^{c}_{it} = \boldsymbol{b}(\boldsymbol{w}^{j}_{it} - \boldsymbol{w}^{c}_{it}) + (\boldsymbol{b} - \boldsymbol{a})\boldsymbol{w}^{c}_{it} \qquad (7)$$

The critical second term of this decomposition now asks the question: How much more would the average conforming loan cost if it were priced as a jumbo instead? Since neither (6) nor (7) has any inherent claim to superiority, we will use both decompositions in what follows. It turns out that the choice does not matter much.

III. Data

We follow many previous researchers in using the Federal Housing Finance Board's Monthly Interest Rate Survey (MIRS) to estimate the effect of conforming status on mortgage loan rates. Each month, the FHFB collects a sample of conventional, fully-amortizing, single-family, non-farm, purchase-money mortgage loans closed during the last five business days of the previous month. We study the 74-month period from April 1997 to May 2003, during which the MIRS data set has information on more than one million fixed-rate loans originated by a sample of U.S. mortgage lenders.[13] The loan characteristics listed earlier are all taken from the MIRS data base. We identify jumbo loans by a dummy variable (called J above) equal to unity if the loan's principal exceeds the conforming loan limit for the year in which the loan closed.

We matched each mortgage with variables describing conditions in the fixed-income markets at the time the loan was closed—the vector $\boldsymbol{Z_{it}}$ above.[14] RM_{it} should vary with these conditions. For example, we control for the height and slope of the yield curve by including both a short-term and a long-term interest rate among the explanatory variables:

CM_1 = the monthly average of daily yields on the constant-maturity Treasury bond with one year to maturity as reported on the Federal Reserve's H.15 report,

CM_10 = the monthly average yield on the constant-maturity Treasury bond with ten years to maturity as reported on the Federal Reserve's H.15 report.

The value of a mortgage prepayment (call) option depends on the expected future volatility of interest rates, which we measure as:

[13] This time period was used by Passmore [2003] and by Passmore, *et al.* [forthcoming], which facilitates comparison.
[14] Each month's mortgage data apply to the last five business days of the month, but loan rates on fixed-rate mortgages are often locked in much earlier. We use monthly averages to capture the level of market rates during the (un-specified) time these mortgages were being priced. We discuss the robustness of our results to these measurement decisions in Section VI below.

VOL_10 = the standard deviation of the daily 10-year Treasury rate over the sixty trading days preceding the month-end at which loans are closed.

The market's aversion to default risk varies substantially over time. To control for such variations, we include the spread between Baa and Aaa bond yields:

SPREAD = the monthly average yield spread (Baa – Aaa) for U.S. industrial bonds with ten years to maturity, obtained from Bloomberg.[15]

Finally, Ambrose *et al*. [2001, 2004] show that house price volatility significantly affects mortgage pricing, and that jumbo loans are more sensitive to this uncertainty. We therefore follow Ambrose *et al.* [2004] in constructing a measure of house price volatility to capture variations in a mortgage's default option:

HPI_STD = the standard deviation of OFHEO's state-level house price index over the preceding eight quarters for the state in which the mortgage property is located.

All previous studies trimmed the MIRS data sets, as described in Table 1 of McKenzie [2002]. Following previous researchers, we analyze only comparable 30-year, fixed-rate mortgages. The MIRS data indicate whether each loan carries an adjustable or fixed rate. However, some of the allegedly fixed-rate mortgages report rates that are suspiciously below Freddie Mac's reported mean monthly rates for such instruments. For example, while the Freddie Mac monthly rate series varied between 5.48% and 8.52% over our sample period, some MIRS observations contain allegedly "fixed-rate" effective loan rates as low as 2.96%. Such low rates seem more likely to describe the initial "teaser" rate on an adjustable-rate loan than the true fixed rate on a 30-year loan. Other data points report rates that are substantially *above* the Freddie Mac monthly average, which might reflect credit problems that make the loan nonconforming irrespective of its principal amount. Following McKenzie [2002], we remove such suspicious observations from the sample by excluding loans with an effective rate more than 50 basis points

[15] We also tried to measure the debt market's aversion to risk with the difference between the yields on indices of Aaa and Baa corporate bonds from the Federal Reserve's H.15 data. These rates reflect unknown call provisions and maturities, arguably making them a noisier measure of market risk aversion. Even so, our regressions yield very similar estimated jumbo-conforming spreads when we use this alternative measure of *SPREAD*.

below, or more than 200 bps above, the corresponding Freddie Mac monthly average for 30-year fixed-rate mortgages.

We also follow previous researchers by eliminating loans whose principal amounts are less than 25% or more than 200% of the contemporaneous conforming loan limit. We do this to limit the impact of loan scale on our estimates. If underwriting or servicing mortgages entails any fixed cost components, smaller loans require higher interest rates to recover the lender's out-of-pocket costs. In principle, this effect can be incorporated into a regression, and we try to do that. But the proper specification is not obvious (more on this below), so we limit the range of included loan sizes. We also omit from our regressions the set of jumbo loans that exceed the conforming loan limit by a small amount. HS were the first to document the unusual behavior of these "near-conforming" loans, and McKenzie [2002] later paid them considerable attention. Such loans might be unusual because (a) the originator can sell them to the GSEs after holding them until year-end (which should reduce their rates if conforming status is valuable), or (b) borrowers who cannot obtain additional financing to get their primary mortgage under the conforming loan limit may have unobserved (to the econometrician) credit problems, which should raise their rates. So we exclude near-conforming loans from our primary sample. Specifically, we omit from the sample all jumbo loans whose principal amount lies below the *following year's* conforming loan limit.[16]

Application of all of these filters left a sample of 964,634 observations. Table 1 presents summary statistics for the variables used in our primary analysis. Out of nearly one million loans, about 5.5% exceeded the conforming limit for the year in which they were originated. The monthly average effective rate of interest (*RM*) ranges from 4.98% to 10.52% in our data, with a mean of 7.22%. The mean (median) *RM* on jumbo loans exceeds that of conforming loans by 4 (12) basis points. The average spread over 10-year Treasuries is 2.03%, with a relatively large

[16] The conforming loan limit increased in every year of our sample, with increases ranging 5.3% to 9.3%. This screen removes only 5,941 loans, constituting 11.3% of the jumbos. Note that Figure 2 includes these near conformers, even though we omit them from the regressions.

standard deviation (0.42%). The mean and median loan-to-value (LTV) ratios are both very close to 80%, although jumbos are slightly (but significantly) less leveraged: About 83% of jumbos had an LTV ratio of 80% or less, compared to 61% for non-jumbos. Higher LTV mortgages expose the investor to more potential loss in the event of default, so we expect them to be priced higher. However, *conforming* loans with an LTV ratio greater than 80% are required to obtain private mortgage insurance to protect the investor from loss in the event of default. For this reason, we expect differential pricing of this default risk between jumbos and conformers.

Jumbos are more likely to be originated by thrifts (32.8% vs. 26.0%), but less likely to be originated by mortgage companies (61.5% vs. 68.0%). Loans originated by mortgage companies account for 67.6% of the sample, and mortgages funded to purchase new construction make up 17.3%. California mortgages are known to be the most common in MIRS; they comprise 16.4% of our sample and 39.4% of our jumbos. We follow Passmore [2003] by including dummy variables for four states with relatively high volume of jumbo loans: *CA*, *MD*, *NJ*, and *VA*. *MD*, *NJ*, and *VA* together account for 10.5% of our sample and 16.4% of our jumbos.

IV. Results from Single-Equation Regression Models

We begin our analysis with the simplest specification, equation (1), reported in the first column of Table 2.[17] This pooled regression yields a very high estimate (38 bps) for the jumbo-conforming spread, but its $\overline{R}^2$ is quite low (0.16) and the coefficients on the *LTV* variables do not increase monotonically, which is odd. Moreover, the estimated scale economies seem implausibly large: The *INV_PRIN* coefficient of 0.056 implies that a $100,000 loan would cost 28 bps more than a $200,000 loan, *ceteris paribus*. By comparison, the estimated scale coefficient in the next column (0.021) implies only an 11 bps price advantage.

[17] The reader should take note of the last line of the Notes to Table 2. In a regression with almost one million observations, significance beyond the 1% level is the norm. Asterisks connote the few coefficients that are *not* significant at this level.

a. Controlling for Changing Market Conditions over Times

Column (1) includes no controls for time-series variation in financial market conditions—the ***Z*** vector discussed above. Merely adding dummy variables to identify calendar years (in column (2)), increases the $\overline{R}^2$ statistic to 0.70, reduces the estimated scale economies by nearly two thirds, and produces *LTV* coefficients that rise monotonically. The estimated coefficient on the jumbo loan dummy falls to 22 bps, which is more in line with previous estimates. Including year dummies also substantially affects the estimated coefficients on *FEES*, the *MTGCO* and *THRIFT* originator dummies, and the mortgage premium paid on a *NEW* house. Clearly, something that changes over time matters.

A more economically satisfying way to control for time variation focuses on conditions in financial markets. Column (3) replaces the year dummies with a set of five financial market variables pertaining to the month of origination: the short-term (*CM_1*) and long-term (*CM_10*) Treasury rates, rate volatility (*VOL_10*), the industrial bond spread (*SPREAD*), and house-price volatility (*HPI_STD*). The significantly positive coefficient on *SPREAD* indicates that mortgage rates rise with the market price of default risk. Both *VOL_10* and *HPI_STD* get significantly positive coefficients, consistent with the hypotheses that more volatility raises a mortgage's embedded prepayment and default options, respectively. This specification produces a slightly smaller estimate of the value of conforming status: 20 bps. Column (4) combines the year dummies (which now matter much less) with the financial-conditions variables, which boosts the $\overline{R}^2$ to 0.80 and moves the jumbo-conforming spread estimate back up to 22 bps.

Seasonality is another issue. Figure 3, which plots the monthly pattern of loan closings, reveals that both conforming and jumbo loans close disproportionately during the spring and summer months, when people tend to move.[18] If resources could flow into and out of the

[18] Figure 3 describes loans closed during 1998-2002, the years for which we have data on all twelve months. The December spike in "near-conforming" loan closings reinforces our suspicion that these loans deserve special treatment. See below.

mortgage underwriting business freely and costlessly, there should be no seasonal pattern in mortgage rates despite the obvious seasonality in demand. But supply-side frictions in the underwriting process may let some of this seasonality flow through to rates. We therefore add eleven month dummies (excluding *JAN*) in column (5) of Table 2. These dummies get coefficients as high as 19.5 bps and the $\overline{R}^2$ increases slightly to 0.81. The seasonal dummies reveal a pattern that is consistent with anecdotal evidence that families are anxious to change houses before the next school year begins. Relative to January, mortgage rates are low through May, rise steadily through September, and then generally decline through year end. We explore seasonal patterns further below. But the month dummies do not affect our other conclusions. In particular, the jumbo-conforming spread estimate increases only slightly (to 23 bps), and the coefficients of the LTV dummies remain monotone increasing. The specification in column (5) provides the basis for our subsequent regression analysis.

Table 3 adds additional explanatory variables to our base regression specification. The first column repeats column (5) of Table 2, but does not report the coefficients for the month, year, state, and originator dummies to save space. Column (2) allows the LTV measures of default risk to have separate coefficients for conforming and jumbo loans. We see that the slope is steeper for jumbos. This means that two loans with the same LTV, one jumbo and one conforming, would be priced differently, perhaps due to the requirement that conforming loans with LTV > 80% must provide private mortgage insurance.

b. Using Information on "Near Conforming" Loans

The specification in column (2) of Table 3 omits an important phenomenon initially documented by HS and explored further by McKenzie [2002]: the tendency of "near conforming" loans to carry lower rates near the end of a calendar year. The rationale is clear enough. The conforming loan limit is raised each January 1, based on the national average house price appreciation implied by MIRS data for the preceding October-October period. Competitive mortgage lenders should therefore anticipate the new limit, leading to lower rates for near-conforming jumbos late in the year. That said, an investor holding a near-conforming loan until the following January does run some risk. First, the housing GSEs will purchase seasoned mortgages only if they have not had any negative payment events. Second, the underwriter does not know the subsequent year's conforming loan limit with certainty until it is announced in late November. During the year, both of these sources of risk decline; and they are basically gone by late December because the new conforming limit is known and the borrower is scheduled to make no payments until the end of January. This suggests that December near-conforming loans may be special.

We define near-conforming loans as those with principal between the current and the subsequent year's conforming loan limits. Figure 3 exhibits a very sharp spike in near-conforming loan originations in December. Not only do 30% of the year's near-conforming loans close in December, but 35% of December's near-conforming loans are for *exactly* the following year's (already announced) conforming limit. This is clearly not a coincidence.

We investigate the impact of near-conforming status on loan pricing in column (3) of Table 3, which adds six new dummy variables to identify near-conforming loans closed in each of the year's last six months. (The sample for this regression must now include the near-conformers, which were omitted previously.) Interestingly, we find no significant pricing effect through August. Near conformers are 5 to 8 bps more expensive in September and October (perhaps because credit quality issues dominate), but they are priced substantially lower in November (-9 bps) and December (-24 bps). Combining the coefficient on the jumbo dummy, *J*, with the one on *DEC_NC* indicates that December near-conforming loans are priced with *no premium* over conforming loans.

In fact, the coefficient on *DEC_NC* can be interpreted as another measure of the value of conforming status because these loans can be sold just a few days later with no risk and little cost of carry. It is striking that this coefficient almost exactly matches the estimated coefficients on *J* in columns (1) and (2). To confirm this finding, we ran another (unreported) regression on a sample that included *only* December's near-conforming loans and January's conforming loans. In this specification, the dummy variable that identifies the near-conformers carries a negligible coefficient (0.017) that is statistically indistinguishable from zero (t-statistic of 0.371).

c. Differences by loan size

It is tempting to think that the jumbo-conforming spread can be estimated well by comparing two similar-sized loans, one on either side of the conforming loan limit. But that will not be true if the characteristics of these two types of loans are priced differently in the conforming versus non-conforming markets. Thus Table 4 looks in more detail at how loans of different sizes are priced.

Column (1) includes all conforming-sized loans in our sample, and column (6) includes all the jumbo loans (except the near-conformers). The intervening columns estimate regressions over subsets of the conforming (columns (2) and (3)) and jumbo (columns (4) and (5)) loans that increase in size as we move from left to right. By comparing the coefficient estimates across loan sizes, we can obtain further indications of whether our specification is reasonable. In addition, these regressions will be used in the decomposition exercises that follow.

What do we find? First, as we move up to larger loan classes, the sensitivity of the mortgage rate to the LTV ratio clearly rises--presumably reflecting a risk premium. For example, moving from *LTV_2* to *LTV_3* costs 2.7 bps in column (3) (large conformers) but 8.6 bp in column (4) (small jumbos). Second, jumbo loan rates are more sensitive to market risk aversion, as indicated by their higher coefficient estimates on *SPREAD_10*. In contrast to some previous speculation, however, jumbo rates are not substantially more sensitive to rate volatility (*VOL_10*) or housing price fluctuations (*HPI_STD*). Third, while all mortgage rates load more heavily on the 10-year Treasury rate *(CM_10)* than on the short rate *(CM_1)*, jumbo loans are *relatively* more sensitive to changes in the short rate.[19] Finally, the monthly seasonal patterns are substantially stronger for conformers than for jumbo loans. In sum, Table 4 clearly indicates that similar loan characteristics are priced differently in conforming versus jumbo loan markets. This is the main idea underlying our final method for computing the jumbo-conforming spread.

V. Estimating the Jumbo-Conforming Spread by the Decomposition Method

Our last and, we would argue, best method of estimating the jumbo-conforming spread is via the decomposition method explained earlier, since it controls for any systematic differences in jumbo versus conforming loans in *either* characteristics *or* pricing. Recall the basic idea. Using separate regressions (equations (5a) and (5b) above) for conforming and jumbo loans, which are reported in Table 4, we can compute two decompositions, (6) and (7), which answer two similar but distinct questions:

- What would be the interest rate on the average *conforming* loan, controlling for all the factors in the regression, if it were priced like a *jumbo* loan?
- What would be the interest rate on the average *jumbo* loan, controlling for all the factors in the regression, if it were priced like a *conforming* loan?

The results are presented in Table 5 for the four non-overlapping size groups listed in Table 4. Consider first the entries below the diagonal. For each group of loans, we compare its actual mean interest rate (*rm*) with a hypothetical fitted *rm* using a *different* group's estimated pricing coefficients to value the loan characteristics. The differences between the two estimates are then reported in the various cells. For example, the bottom entry in the first column (0.226) indicates that the average loan between 120 and 200% of the conforming loan limit would have been priced 22.6 bps *lower* had it been priced like a loan between 25 and 80% of the limit. Likewise, the top right cell (-0.264) indicates that loans between 25 and 80 percent of the

[19] We use the mortgage rate level as our dependent variable in all regressions, in large part because the estimated coefficients on *CM_10* are substantially less than unity. This suggests that using a yield spread as the dependent variable (e.g. *RM-CM_10)* would impose an unwarranted restriction on the data.

conforming limit would have been priced 26.4 bps *higher* had they been valued like loans from the 120-200 size group.

Notice that when we compare either small versus large *conforming* loans or small versus large *jumbo* loans, the estimated differences are small; three of the four estimates are under a basis point. In contrast, when jumbo loans are valued according to the estimated equation for conforming loans, the differences are large. For example, loans between 100 and 120% of the limit (but not near-conformers) would be priced 22.2 bps lower if they were valued according to the 25-80% coefficients, and 23.5 bps lower if valued as a large conforming (80-100%) loan. The four lower-left estimates of jumbo premium vary in a narrow range between 22 and 25 basis points, with an average of 23.4. Alternatively, we can value conforming loans using the estimated coefficients for jumbo loans. This is the expression $(\boldsymbol{\beta} - \boldsymbol{\alpha})\boldsymbol{w}^{c}_{it}$ in (7), whose values are reported above the diagonal in Table 5. For example, the upper right cell (-0.264) indicates that the mean fitted *rm* for 25-80% loans using 120-200% coefficients exceeds its actual mean value by 26.4 bps. The upper-right estimated spreads range from 25 to 29 bps, with a mean of 26.3.

Our preferred estimates are the ones that compare loan pricing relatively close to (though both above and below) the conforming limit: these estimated conforming spreads are very close to one another in Table 5: 23.5 and 25 bps.

VI. Robustness Checks

Our results are quite robust to alternative trims of the data, to various ways of measuring market conditions, and to the addition of dummy variables for state-level bankruptcy regimes. We also find evidence consistent with the hypothesis that the cost of private MBS originations affects quoted loan rates. However, the way we handle loan size econometrically does seem to matter. So we turn to that issue first.

a. <u>Loan-scale Effects</u>

Most previous researchers have sought to capture scale effects in the cost of underwriting mortgage loans by including either the inverse of principal (our *INV_PRIN*) or *ln(PRIN)*. However, Passmore *et al.* [2003] use only a dummy variable for loans with principal below $100,000 to capture the potential scale effects in underwriting. Passmore *et al.* [forthcoming] argue that the proper regression specification omits any scale measure because:

> When one includes the loan size as an explanatory variable, the resulting estimate of the jumbo-conforming spread essentially double counts the decrease in mortgage rates arising because of a large loan size… (page 23).

We find this explanation puzzling. As stated earlier, if the loan origination process includes any fixed costs, competition will force larger loans to bear lower interest rates in a competitive market for originations. Omitting loan size then biases the coefficient of the jumbo dummy downward: since jumbo loans are relatively cheap dollars to lend, their loan rates will naturally be lower, *ceteris paribus*.

We evaluate the effect of loan size in the four regressions reported in Table 6. We start with our base specification from the third column of Table 3, then try three alternative *PRIN* treatments: *ln(PRIN)*, a dummy variable (*SMALL*) equal to unity for loans whose principal is below $100,000, and completely omitting *PRIN* from the explanatory variables.

First, compare column 1 with columns 3 and 4. Our base case using *INV_PRIN* to measure loan size (col. 1) estimates the spread to be 24 bps. Using the *SMALL* dummy variable alone (col. 3) produces a slightly worse fit and an estimated jumbo-conforming spread of 17 bps; and omitting *PRIN* altogether (col. 4) yields a still-worse fit and an estimated spread of 15 bps. Except for the intercept terms, the other estimated coefficients remain pretty stable across all four specifications. This means, however, that the specifications in columns 3 and 4 have a troubling implication: Adding the coefficient on the jumbo dummy, *J*, to that of *DEC_NC* (the dummy for December near conformers) yields a significantly negative value ($p < 0.0001$), which implies that late-December near conformers are 5 to 7 bps *cheaper* than the average conforming mortgage. This illogical implication leads us to believe that these two regressions are mis-specified.

However, choosing between the *INV_PRIN* specification in column 1 and the *ln(PRIN*) specification in column 2 is not easy. The fit is identical, but measuring loan scale by *ln(PRIN*) instead yields an estimated jumbo-conforming spread that is 5 bps higher. This turns out to be a general result: In most of our regression variants, the *ln(PRIN*) specification gives a larger estimated value. Why is that? When we compare the two specifications' estimated coefficients, the only two that differ meaningfully are the intercept and the coefficient on the loan size variable. Figure 4 suggests how this works by illustrating the implied scale economies under the two specifications. The inverse function imposes far faster *marginal* cost declines for small loans. Hence, replacing *INV_PRIN* by *ln(PRIN)* decreases the intercepts of the (unreported) versions of the Table 4 regressions applying to the smaller conforming loans. The intercepts for large loans in the Table 4 regressions do not differ much, because the slopes of the estimated *INV_PRIN* and *ln(PRIN)* functions are very similar for larger jumbo loans.

Although it seems clear that we want to include loan size among the explanatory variables, the choice between *INV_PRIN* and *ln(PRIN)* as measures of loan size is not so clear. McKenzie [2002] criticized the log transformation because it can imply negative origination and servicing costs for large loans. But we elided this problem by measuring *PRIN* in millions of dollars; thus even the largest mortgage in our sample implies a positive cost of origination and servicing. McKenzie (p. 203) argues that the inverse form is "more appropriate if servicing costs are invariant with respect to loan size because it is monotonic and asymptotic to zero as loan size gets very large." (p. 203). But can theory really help us decide which is more appropriate? We are doubtful. While it is easy to find cost structures for which *INV_PRIN* is exactly the correct specification, there is no guarantee that these structures apply in the actual data.

Suppose that origination costs are $c_0 + c_1$ *PRIN,* where c_0 and c_1 are non-negative constants. (Ignore the costs and revenues associated with servicing the mortgage loan for now.) The market will price loans to recoup these origination costs, but it is not obvious how this will occur. If closing fees just compensate for the lender's processing costs, MIRS data should report

$$RM = RC + \left(\frac{FEE}{PRIN}\right)/10 = RC + \left(\frac{c_0}{PRIN} + c_1\right)/10\,,$$

where *RC* is the annual mortgage contract rate. This implies that *INV_PRIN* is at least partially correct, and that the estimated coefficient on *INV_PRIN* equals (c_0 /10). However, borrowers could also "repay" their origination costs by agreeing to a higher coupon rate on their debt. The ensuing quasi-rents are like an interest-only strip, whose expected value at closing equals the lender's origination costs. Closing fees are now zero and it is no longer clear that *INV_PRIN* is the right functional form. The situation is further complicated if some of the originator's compensation is provided by profits on servicing the loan (or selling the servicing rights).

Assume that gross servicing costs are given by $s_0 + s_1\, PRIN$, where s_0 and s_1 are non-negative constants. Servicing revenues also have two components, the fee (usually 25 bps) and some value derived from float (s_2 per dollar of *PRIN*). Competition among mortgage originators should drive the benefits to zero, so in the absence of a contract rate premium, closing fees will be

$$FEE = (c_0 + c_1 PRIN) + s_0 + PV(s_1 PRIN) - PV[PRIN(.25\% + s_2)]$$

Therefore

$$RM = RC + [(c_0 + s_0)/PRIN + c_1 - PV((.25\% + s_2 - s_1))] /10$$

From these simple examples, it should be clear that one cannot identify "the" proper measure of scale effects without further information about the originator's production function and the determinants of a loan's duration. Unfortunately, our sensitivity analysis shows that the magnitude of the estimated jumbo-conforming spread does depend on how the regression incorporates loan-size effects. However, the estimated spread is sizable in both cases.

b. Alternative Data Trims

In our main results, we follow previous researchers in excluding near-conforming loans and deciding which "outliers" to omit by comparing each mortgage's reported effective rate to Freddie Mac's monthly average rate on new fixed-rate mortgages. Our base sample is derived by applying McKenzie's [2002] rule that omits all loans with effective rates more than 50 bps below, or more than 200 bps above, the Freddie Mac average for the origination month. However, as a robustness check, we re-ran all our regressions over three alternative samples:

- "Less Trimmed": Excludes near-conformers and loans whose *RM* was more than 100 bps below, or more than 200 bps above, the Freddie Mac average.
- "Not Trimmed": Includes all loans that were identified in MIRS as fixed rate contracts, except near-conforming loans.
- "Add Near-conforming": Adds near-conforming loans to the primary sample.

All three of our methods for estimating the jumbo-conforming spread (using the dummy *J*, using *DEC_NC*, and the more complex decomposition method) are robust to changing the sample in the three ways listed above. The most noticeable change when using the estimated coefficient of *J* to measure the spread was generated by the "Not Trimmed" alternative, which produced estimates that were on average 1 bps greater than those reported in Tables 2 and 3. When using the *DEC_NC* variable, the "Less Trimmed" and "Not Trimmed" alternatives reduced the estimated spread from the reported 24 bps to 23 bps. The decomposition method estimates were generally robust to both the "Less Trimmed" and "Not Trimmed" alternatives; the estimates were essentially unchanged for the "Less Trimmed" version, and the "Not Trimmed" sample yielded an average estimate about 3 bps greater than that of our base case. However, changing to the "Add Near-conforming" sample did affect the decomposition results somewhat, as expected. Remember, we have found that many near-conforming loans are priced like large conforming loans. So it was not surprising that including them lowered the average decomposition estimate of the *difference* between conforming and nonconforming loan rates from 25 bps to 21 bps.

c. Market Controls

Using our primary sample, we tried modifying the specification in three ways. First, Ambrose *et al.* [2004] concluded that the state's bankruptcy regime significantly affects effective mortgage yield spreads. So we constructed Ambrose *et al.*'s four state-law dummies to control for (a) judicial versus non-judicial foreclosure laws and (b) deficiency versus non-deficiency judgment states. All results, including the $\overline{R}^2$ statistics, were virtually unchanged when we included the state-law dummies.

Second, we allowed for the possibility that loan pricing depends more on financial market conditions during the lock-in period, which may be one or two months earlier than the closing. Similar results are obtained when we lag the market conditions variables (*CM_1*, *CM_10*, *VOL_10*, *SPREAD*, and *HPI_STD*) either one or two months. The estimated coefficients on *J* remained essentially unchanged, the *DEC_NC* estimated coefficients were about 3 bps greater in absolute value, and the estimates from the decomposition method were essentially unchanged. None of these alternatives produced a substantially higher $\overline{R}^2$.

Third, we measured the debt market's risk aversion with a different variable. Our *SPREAD* series is the difference between yields on U.S. industrial bonds (Bbb-Aaa) with ten-years to maturity, obtained from Bloomberg. We created an alternative measure (*SPREAD_H15*) using the Fed's H.15 Statistical Release data for indices of Moody's seasoned corporate bonds rated Aaa and Baa. Because the H.15 measures are likely calculated from bonds with various call provisions and maturities, it is not clear that *SPREAD_H15* reflects an appropriate ten-year maturity measure.[20] In addition, through December 6, 2001, the H.15 rates were averages of both utility and industrial bonds instead of only industrial bonds as it is now reported. When we replace the Bloomberg *SPREAD* variable with *SPREAD_H15*, that variable's estimated coefficient was slightly larger (an average of 0.532 versus 0.392 from Table 4), but the estimated coefficients of *J* from Tables 2 and 3 were essentially unchanged. Likewise, the coefficients estimated using *SPREAD_H15* yielded virtually identical measures for the decomposition method shown in Table 5. The *DEC_NC* estimates were about 1 bp greater in absolute value.

d. Splitting the sample

[20] The correlation coefficient for SPREAD and SPREAD_H15 during our sample period is only 0.62.

We split our sample period into two parts, for two reasons. First, it provides another check on our regression specification: If it fits well in all sub-periods, our confidence in its correctness increases. Second, it has been suggested that the jumbo-conforming spread has declined in recent years (Passmore *et al.* [forthcoming]). Often sample splits are arbitrary, but our sample period includes one regulatory innovation that might have reduced the jumbo-conforming spread by making private securitizations cheaper.

Private securitizations rely on "junior tranches" to provide credit protection to the senior (often Aaa-rated) tranches collateralized by a loan pool. Because the junior tranches absorb most of the pool's default risk, they are difficult for the loan originator to sell. Before November 2001, banks that retained junior tranches of their own pools were required to hold risk-based capital against the junior tranche plus all senior tranches, just as if the bank had instead retained the entire security. By contrast, if a bank purchased another originator's junior tranche, the capital requirement was based only on the dollar amount purchased. Thus, banks securitizing jumbo mortgages were forced either to sell their junior tranches and pay a "lemon's discount," or retain them and incur large capital charges. In November 2001, federal regulators implemented a ratings-based approach to capital requirements that effectively lowered the required capital for firms that issue private-label mortgage backed securities. This action should have reduced the jumbo-conforming spread by making it less expensive for banks to securitize jumbos.

To test this hypothesis, column (4) of Table 3 adds a dummy variable (*POST_REG*) defined as 1 if the observation describes a jumbo mortgage funded after November 2001 and zero otherwise. The coefficient of -0.072 (t = 22.03) indicates that jumbo loans were 7 bp less expensive following this regulatory change. Combined with the equation's estimated jumbo coefficient (0.27), we estimate that the jumbo-conforming rate differential was 27 bps through November 2001 and 20 bps thereafter. We also undertake the Blinder-Oaxaca analysis for the two sub-periods. The decomposition method estimates that the average jumbo-conforming rate differential decreased by about 5 bps after this regulatory innovation. The evidence is thus that the jumbo-conforming spread declined later in our sample period, perhaps due to a reduction in the cost of issuing private mortgage backed securities.

VII. Summary and Conclusions

The effect of GSE conforming status on the cost of a mortgage has been studied often. It is, for example, a first step in assessing the social benefits of the housing GSEs. Nearly all previous studies have employed the Federal Housing Finance Board's Monthly Interest Rate Survey (MIRS) data and a basic regression specification derived from Hendershott and Shilling [1989]. Previous estimates of the jumbo-conforming spread range from 8 to 60 basis points, and at least some authors claim that the spread has declined over time. In this paper, we use the MIRS data to assess the impact of alternative regression specification and data selection decisions on the estimated spread. We conclude that the jumbo-conforming spread is relatively large (roughly 25 bps) and has fallen somewhat after late 2001.

Although we base our analysis on the Hendershott-Shilling model, we augment that model to recognize that some differences in loan pricing reflect real economic effects beyond the loan's conforming status. Jumbo and conforming loan rates exhibit different sensitivities to LTV, the market's aversion to default risk, and short versus long-term Treasury rates. They also follow surprisingly different seasonal patterns. It is important to control for these differences when estimating the value of conforming status, and we do so in three distinct ways.

Our first method of estimating the jumbo-conforming spread involves a single equation which generalizes Hendershott and Shilling by permitting selected coefficients to differ between smaller and larger loans. These regressions imply a jumbo-conforming differential of 23-29 bps for our 1997-2003 sample period.

Our second method expands the single-equation format to take advantage of the fact that the conforming loan limit rose every year during our sample period. A jumbo loan according to *this year's* conforming loan limit may fall below *next year's* limit, so near-conforming jumbo loans closed late in the year should be much like conforming loans. Indeed, we find a significant effect of this sort, but only for loans closed in November and December. In particular, the estimation results indicate that a near-conforming loan originated in December carries an effective rate that is about 24 bps below the same-sized loan made the previous January. We also find that December near-conforming loans are priced the *same as* conforming loans closed the following month. These results provide particularly strong evidence of a positive jumbo-conforming spread, since the pricing shifts so dramatically in such a short period of time.

Our third method for estimating the value of conforming status takes account of *all* loan characteristics by estimating separate pricing regressions for conforming and jumbo loans. Using a technique originated for a quite different purpose by Blinder [1973], we compute the effect of conforming (jumbo) status by estimating what the rate on an otherwise-identical jumbo (conforming) loan would have been if it were priced according to the conforming (jumbo) loan equation. When we use the conforming loans equation to price jumbo loans, Table 5 indicates a jumbo-conforming spread in the range of 22 to 25 bps. Pricing conforming loans as if they were jumbo indicates a spread of 24 to 29 bps.

All of these estimates are pretty similar, but we also find that the treatment of loan scale effects can change the estimated jumbo-conforming spread by a non-trivial amount. When loan size is measured by the log of principal instead of its inverse, the spread estimates vary more widely (between 20 and 30 bps), although their mean is the same as in Table 5: 25 bps. We further find that near-conforming loans require special treatment. While it is tempting to compare the rates on jumbo loans just over the conforming limit to those on conforming loans just under, the smallest jumbos have unusual features that make this comparison treacherous. For example, the rates on at least some of the smaller jumbo loans must reflect unobserved credit problems, for why else would a customer take such a loan? And as just mentioned, late in the year the smallest jumbo loans will be priced as if they were conforming, which would give a false impression that conforming status has no value.

What do these results imply about the social value of Fannie Mae and Freddie Mac? Taken at face value, they indicate that private securitization arrangements are more costly than securitizations occurring through the GSEs' mortgage backed securities, at least during our sample period. However, these estimates derive from observed institutional arrangements, and Passmore *et al*. [forthcoming] assert that the GSEs' conforming loan securitizations raise the cost of selling private mortgage pools. By this reasoning, 25 bps would be an upper bound on the GSEs' effect on mortgage rates. Conversely, curtailing GSE securitizations might remove valuable liquidity from the market, driving up all mortgage rates.

As with any statistical evidence, we must acknowledge the possibility that our regressions are not perfectly specified, and hence fail to value the features of jumbo versus conforming loans properly. And a full assessment of the social value of the housing GSEs involves other considerations that we have not even discussed here. But that analysis is needed only if we start with evidence that the jumbo-conforming spread is of an economically meaningful size. Although we find that the spread is difficult to estimate precisely, it is extremely unlikely that it is small enough to ignore.

REFERENCES

Ambrose, Brent W., Richard Buttimer, and Thomas Thibodeau, "A New Spin on the Jumbo/Conforming Loan Rate Differential," *Journal of Real Estate Finance and Economics*, Volume 23, Issue 3, pages 309-335, November 2001.

Ambrose, Brent W., Michael LaCour-Little, and Anthony B. Sanders, "The Effect of Conforming Loan Status on Mortgage Yield Spreads: A Loan Level Analysis," *Real Estate Economics*, Volume 32, Issue 4, pages 541-569, Winter 2004.

Blinder, Alan S., "Wage Discrimination: Reduced Form and Structural Estimates," *The Journal of Human Resources*, Volume 8, Issue 4, pages 436-455, Fall 1973.

Blinder, Alan S., Mark J. Flannery, and James Kamihachi, "The Value of Housing Related Government Sponsored Enterprises: A Review of a Preliminary Draft Paper by Wayne Passmore," Fannie Mae Papers, Vol. III, Issue 2, May 2004.

Cotterman, Robert F., and James E. Pearce, "The Effect of the Federal National Mortgage Association and the Federal Home Loan Mortgage Corporation on Conventional Fixed-Rate Mortgage Yields," *Studies on Privatizing Fannie Mae and Freddie Mac, U.S. Department of Housing and Urban Development,* Office of Policy Development and Research, pages 97-168, 1996.

Hendershott, Patric H., and James D. Shilling, "The Impact of the Agencies on Conventional Fixed-Rate Mortgage Yields," *Journal of Real Estate Finance and Economics*, Volume 2, Issue 2, pages 101-115, 1989.

ICF Incorporated, "Effects of the Conforming Loan Limit on Mortgage Markets," Report prepared for the U.S. Department of Housing and Urban Development, Office of Policy Development and Research, 1990.

Fama, Eugene F., and James D. MacBeth, "Risk, Return, and Equilibrium: Empirical Tests," *Journal of Political Economy*, Volume 81, Issue 3, pages 607-637, May/June 1973.

Greene, William H., "Commentary on 'The GSE Implicit Subsidy and Value of Government Ambiguity,'" Fannie Mae, 2004.

Jaffee, Dwight, "The Interest Rate Risk of Fannie Mae and Freddie Mac," *Journal of Financial Services Research*, Volume 24, Issue 1, pages 5-30, 2003.

Jaffee, Dwight, "On Limiting the Retained Mortgage Portfolios of Fannie Mae and Freddie Mac," mimeo, University of California at Berkeley, 2005.

Kau, James B., and Donald C. Keenan, "An Overview of the Option-Theoretic Pricing of Mortgages," *Journal of Housing Research*, Volume 6, Issue 2, pages 217-244, 1995.

McKenzie, Joseph A., "A Reconsideration of the Jumbo/Non-Jumbo Mortgage Rate Differential," *Journal of Real Estate Finance and Economics*, Volume 25, Issue 2-3, pages 197-213, 2002.

Naranjo, Andy and Alden Toevs, “The Effects of Purchases of Mortgages and Securitization by Government Sponsored Enterprises on Mortgage Yield Spreads and Volatility,” *Journal of Real Estate Finance and Economics*, Volume 25, Issue 2-3, pages 173-195, 2002.

Oaxaca, Ronald, “Sex Discrimination in Wages,” in O. Ashenfelter and A. Rees (eds.), Discrimination in Labor Markets (Princeton, NJ: Princeton University Press), 1973.

Passmore, Wayne, “The GSE Implicit Subsidy and Value of Government Ambiguity,” Board of Governors of the Federal Reserve System, December 22, 2003.

Passmore, Wayne, Shane M. Sherlund, Gillian Burgess, “The Effect of Housing Government-Sponsored Enterprises on Mortgage Rates,” forthcoming in *Real Estate Economics*.

Passmore, Wayne, Roger Sparks, and Jamie Ingpen, “GSEs, Mortgage rates, and the Long-Run Effects of Mortgage Securitization,” *Journal of Real Estate Finance and Economics*, Volume 25, Issue 2-3, pages 215-242, 2002.

Pearce, James E., and J. C. Miller, III, “Freddie Mac and Fannie Mae: Their Funding Advantage and Benefits to Consumers,” working paper Welch Consulting, College Station, Texas, 2001.

Sanders, Anthony B., “Government Sponsored Agencies: Do the Benefits Outweigh the Costs?,” *Journal of Real Estate Finance and Economics*, Volume 25, Issue 2-3, pages 121-127, 2002.

U.S. Congressional Budget Office (CBO), “Interest Rate Differentials Between Jumbo and Conforming Mortgages,” 2001.

Figure 1: Mortgage principal distribution, by % of applicable conforming limit

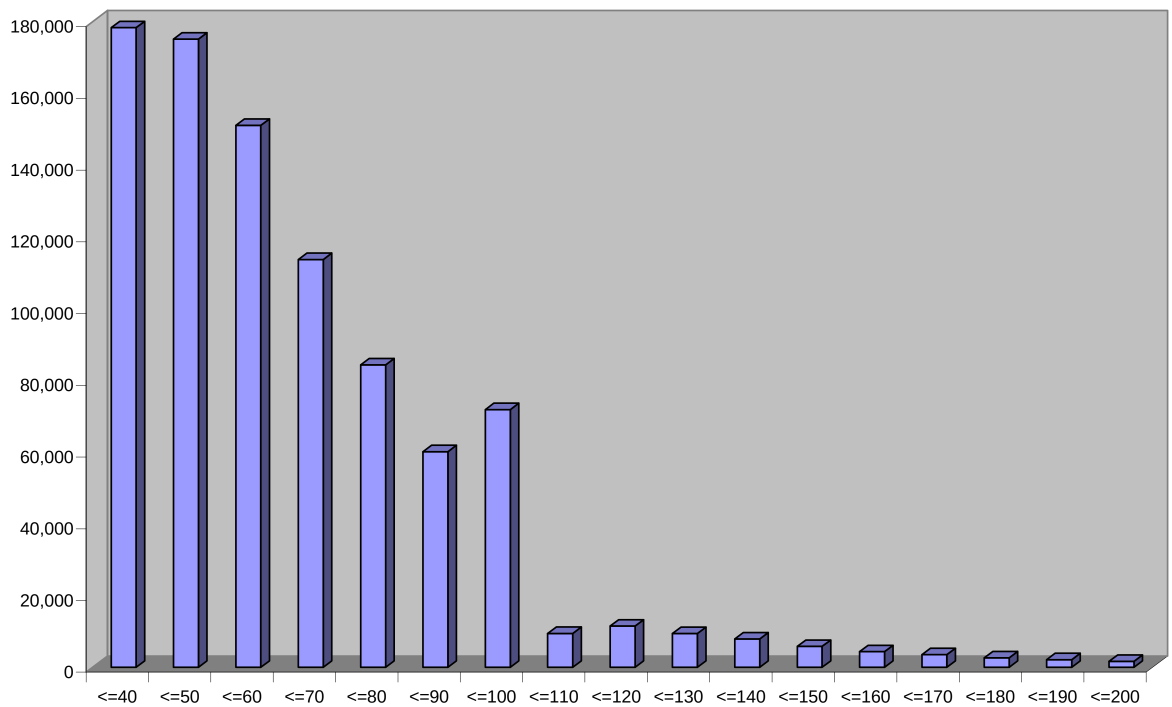

Figure 2: Mortgage principal distribution, by % of applicable conforming limit (different scale)

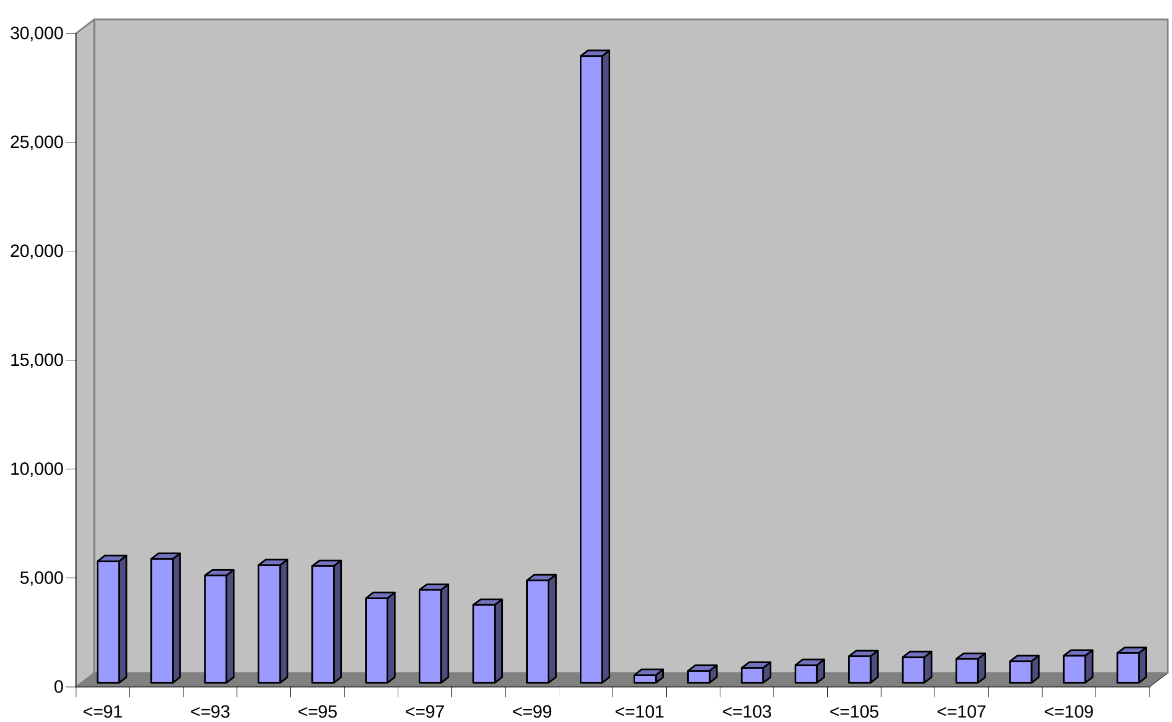

Figure 3: Mortgage principal distribution per month by type, 1998-2002

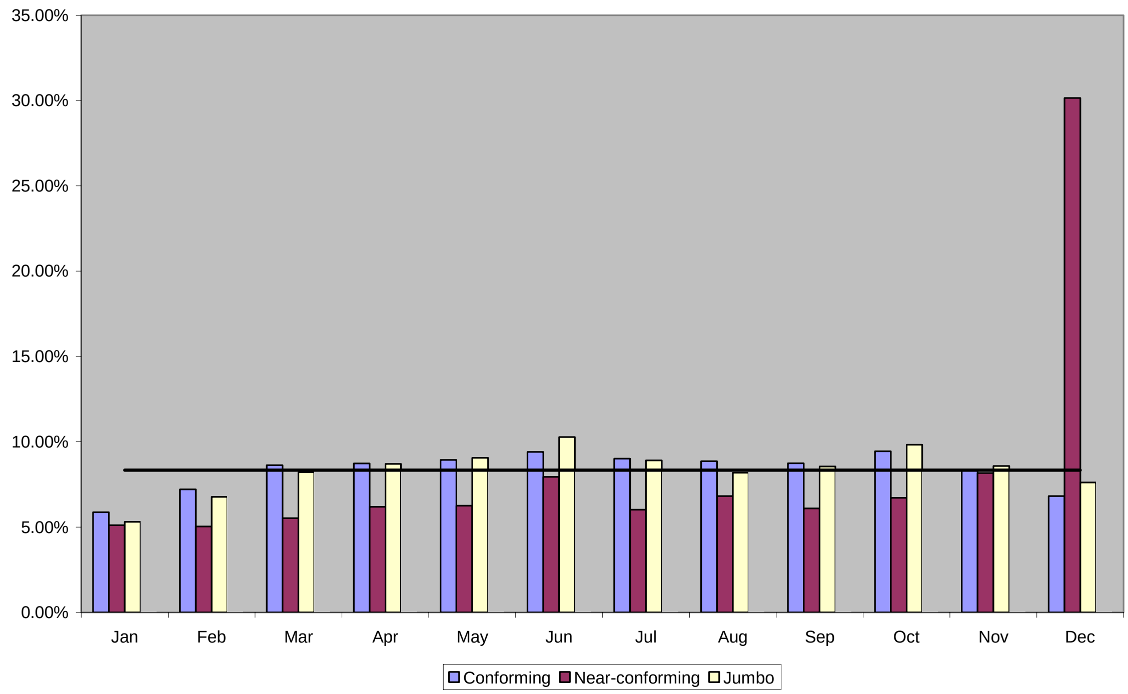

Figure 4: Alternative measures of loan size

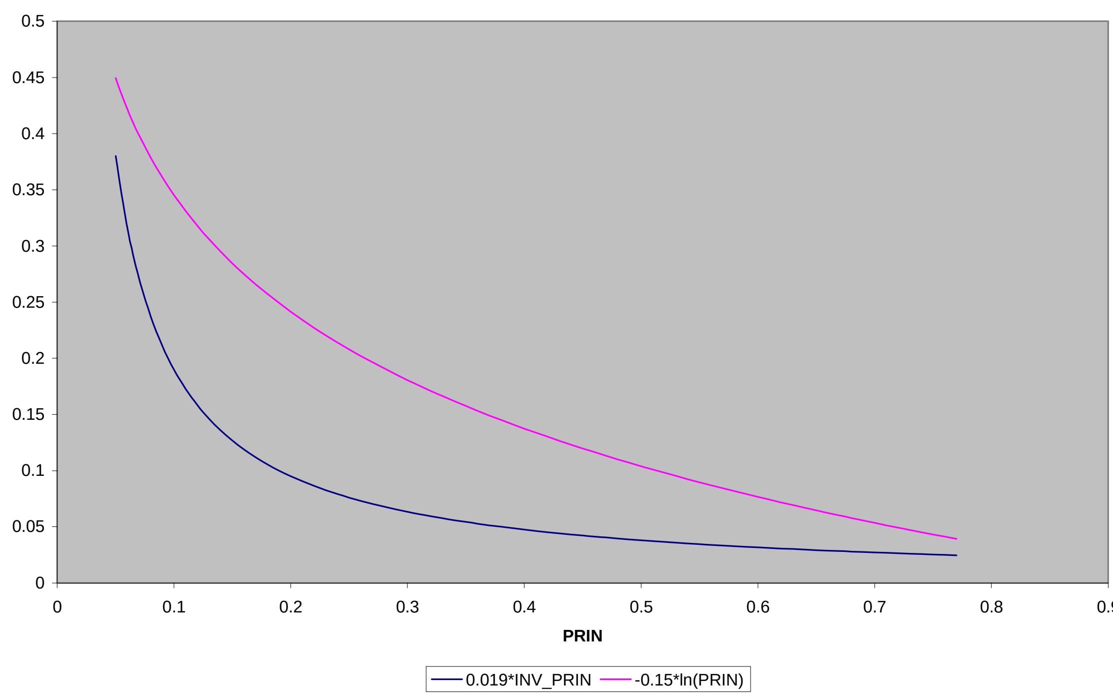

Table 1: Summary statistics

	Mean				Median			
	All	**Conf.**	**Jumbo**	**T-statistic**	**All**	**Conf.**	**Jumbo**	**Z-statistic**
Mortgage loans:								
Number	964,634	911,956	52,678		964,634	911,956	52,678	
Percent		94.5%	5.5%			94.5%	5.5%	
RM (%)	7.224	7.222	7.262	11.605	7.172	7.162	7.281	17.893
RM-CM_10 (%)	2.030	2.024	2.125	53.853	1.990	1.984	2.097	58.681
LTV (%)	79.413	79.575	76.617	-45.771	80.000	80.000	79.991	-78.360
LTV_0	0.283	0.280	0.342	30.506	0.000	0.000	0.000	30.492
LTV_1	0.339	0.330	0.489	75.254	0.000	0.000	0.000	75.034
LTV_2	0.141	0.143	0.114	-18.767	0.000	0.000	0.000	-18.764
LTV_3	0.237	0.247	0.056	-101.048	0.000	0.000	0.000	-100.517
PRIN	156,360	144,684	358,494	817.117	138,000	133,600	348,000	378.516
ln(PRIN)	-1.956	-2.009	-1.050	559.688	-1.981	-2.013	-1.056	378.516
INV_PRIN	7.755	8.034	2.924	-373.893	7.246	7.485	2.874	-378.516
SMALL	0.230	0.243	0.000	-130.187	0.000	0.429	0.000	-129.058
J	0.055	n/a	n/a	n/a	0.000	n/a	n/a	n/a
FREDDIE (%)	7.024	7.029	6.928	-33.290	7.010	7.010	6.990	-29.906
FEES	0.643	0.646	0.590	-26.212	1.000	1.000	1.000	-26.203
BANK	0.060	0.060	0.057	-2.530	0.000	0.000	0.000	-2.530
MTGCO	0.676	0.680	0.615	-30.896	1.000	1.000	1.000	-30.881
THRIFT	0.264	0.260	0.328	34.162	0.000	0.000	0.000	34.142
NEW	0.173	0.170	0.235	38.566	0.000	0.000	0.000	38.536
CA	0.164	0.151	0.394	147.795	0.000	0.000	0.000	146.150
MD	0.031	0.030	0.052	27.736	0.000	0.000	0.000	27.725
NJ	0.039	0.038	0.060	25.828	0.000	0.000	0.000	25.819
VA	0.035	0.034	0.052	21.982	0.000	0.000	0.000	21.977
OTHER_STATE	0.730	0.747	0.442	-155.162	1.000	1.000	0.000	-153.261
CM_SLOPE	1.112	1.110	0.683	9.253	0.683	1.154	0.691	9.287
CM_1 (%)	4.082	4.088	3.983	-13.932	4.690	4.690	4.690	-16.495
CM_10 (%)	5.194	5.198	5.137	-17.533	5.233	5.233	5.233	-17.397
SPREAD (%)	0.840	0.842	0.816	-26.582	0.881	0.890	0.866	-23.952
SPREAD_H15 (%)	0.887	0.888	0.881	-5.759	0.824	0.824	0.824	-9.600
VOL_10	0.174	0.174	0.174	-1.100	0.158	0.158	0.158	-5.745
HPI_STD	9.528	9.413	11.526	76.900	7.701	7.593	9.334	65.368

NOTES TO TABLE 1: Descriptive statistics for our primary sample. RM is the mortgage's effective rate of interest. RM-CM_10 is RM less the 10-year constant maturity U.S. Treasury bond. PRIN is the initial balance of the loan. ln(PRIN) is the natural log of (PRIN/$1,000,000). INV_PRIN is ($1,000,000/PRIN). SMALL is a dummy variable equal to unity if PRIN < $100,000. J is a dummy equal to unity if the loan exceeded the conforming limit at closing. LTV is the mortgage's loan-to-value ratio. LTV_0, LTV_1, LTV_2, and LTV_3 are dummies equal to one if LTV does not exceed 75%, is greater than 75% but does not exceed 80%, is greater than 80% but does not exceed 90%, or exceeds 90%, respectively. FEES is a dummy variable equal to unity if RM includes the amortization of fees and/or charges. NEW is a dummy variable equal to unity if the loan funded the purchase of new construction. BANK, MTGCO, and THRIFT are dummies equal to unity if the loan was closed by a commercial bank, mortgage company, or thrift institution, respectively. CA, MD, NJ, and VA are dummy variables indicating the state in which the property is located, and OTHER_STATE is a dummy equal to unity if CA, MD, NJ, and VA are all equal to zero. FREDDIE is the monthly average conventional mortgage rate reported by the FHLMC. CM_1 (CM_10) is the monthly average yield on the 1-year (10-year) constant maturity U.S. Treasury bond. CM_SLOPE is the monthly CM_10 minus CM_1. SPREAD is the monthly average yield spread (Baa-Aaa) for U.S. industrial bonds with 10-years to maturity. SPREAD_H15 is the monthly average yield spread (Baa-Aaa) for U.S. corporate bonds calculated from the Federal Reserve's H.15 Statistical Release. VOL_10 is the standard deviation of the 10-year Treasury over the preceding 60 trading days. HPI_STD is the standard deviation of the OFHEO state-level house price index over the preceding eight quarters. The t-statistics (z-statistics) test the hypothesis that the means (medians) are equal for conforming and jumbo mortgages.

Definitions of all variables:

Loan Terms:

RM:	The contract rate of interest plus fees and charges amortized over a 10-year period multiplied by 100.
RM-CM_10:	The loan's effective interest rate (RM) minus CM_10 (defined below).
PRIN:	The principal amount of the mortgage.
ln(PRIN):	The natural logarithm of the loan's principal amount, in million dollars.
INV_PRIN:	The principal amount divided into $1,000,000.
SMALL:	A dummy variable equal to unity when the loan's principal amount is less than $100,000.
J:	A dummy variable equal to unity when the loan's principal amount exceeds the conforming limit at closing.

Loan-to-value Dummies:

LTV:	The mortgage's principal amount divided by the property purchase price multiplied by 100.
LTV_0:	A dummy variable equal to unity when the loan-to-value ratio does not exceed 0.75. (omitted category)
LTV_1:	A dummy variable equal to unity when the loan-to-value ratio exceeds 0.75 but is less than or equal to 0.80.
LTV_2:	A dummy variable equal to unity when the loan-to-value ratio exceeds 0.80 but is less than or equal to 0.90.
LTV_3:	A dummy variable equal to unity when the loan-to-value ratio exceeds 0.90.
J_LTV_0:	A dummy variable equal to unity when LTV_0=1 and JUMBO=1. (omitted category)
J_LTV_1:	A dummy variable equal to unity when LTV_1=1 and JUMBO=1.
J_LTV_2:	A dummy variable equal to unity when LTV_2=1 and JUMBO=1.
J_LTV_3:	A dummy variable equal to unity when LTV_3=1 and JUMBO=1.

Origination Dummies:

FEES:	A dummy variable equal to unity when fees and charges were associated with the loan.
NEW:	A dummy variable equal to unity when the loan is associated with a newly built house.
BANK:	A dummy variable equal to unity when the loan is originated by a commercial bank (as opposed to a mortgage company or thrift institution). (omitted category)
MTGCO:	A dummy variable equal to unity when the loan is originated by a mortgage company (as opposed to a bank or thrift institution).
THRIFT:	A dummy variable equal to unity when the mortgage loan is originated by a thrift institution (as opposed to a bank or mortgage company).

State Dummies:

CA: A dummy variable equal to unity when the purchased property is located in California.

MD: A dummy variable equal to unity when the purchased property is located in Maryland.

NJ: A dummy variable equal to unity when the purchased property is located in New Jersey.

VA: A dummy variable equal to unity when the purchased property is located in Virginia.

OTHER_STATE: A dummy variable equal to unity when the purchased property is not located in California, Maryland, New Jersey, or Virginia. (omitted category)

<u>Year Dummies</u>:

D_1997...D_2003: Seven dummy variables, D_1997…D_2003, that equal unity when the loan is closed in 1997…2003, respectively. (D_2003 is omitted)

<u>Seasonal Dummies</u>:

JAN…DEC: Twelve dummy variables, JAN…DEC, equal to unity when the loan is closed in January...December, respectively. (JAN is omitted)

JUL_NC…DEC_NC: Six dummy variables, JUL_NC…DEC_NC, equal to unity when JUMBO=1, the loan is closed in JUL…DEC, and PRIN is less than or equal to the next year's conforming loan limit.

<u>Market Conditions</u>:

FREDDIE: The monthly average contract rate for closed 30-year conventional rate mortgages, compiled by the Federal Home Loan Mortgage Corporation.

CM_1: The monthly average of daily yields on the constant maturity Treasury bond with one year to maturity, multiplied by 100.

CM_10: The monthly average of daily yields on the constant maturity Treasury bond with ten years to maturity, multiplied by 100.

CM_SLOPE: The monthly difference of average daily yields on the constant maturity Treasury bonds with ten years and one year to maturity, multiplied by 100 (CM_10 – CM_1).

SPREAD: The monthly average yield spread (Baa-Aaa) for U.S. industrial bonds with ten years to maturity, multiplied by 100 (obtained from Bloomberg).

SPREAD_H15: The monthly average ((Baa-Aaa) for U.S. corporate bonds, multiplied by 100, calculated from the Federal Reserve's H.15 Statistical Release.

VOL_10: The standard deviation of the daily 10-year Treasury rate over the sixty trading days preceding the month-end at which loans are closed, multiplied by 100.

HPI_STD: The standard deviation of OFHEO's state-level house price index over the preceding eight quarters for the state in which the mortgage property is located.

POST_REG: A dummy variable equal to unity when JUMBO=1 and the loan was closed after November 2001.

Table 2: Single-equation models of mortgage rate determination

	(1)	(2)	(3)	(4)	(5)
INTERCEPT	6.511	5.666	2.063	2.927	2.627
J	0.380	0.217	0.201	0.224	0.226
LTV_1	0.063	0.028	0.024	0.017	0.016
LTV_2	0.244	0.121	0.096	0.103	0.102
LTV_3	0.215	0.139	0.116	0.119	0.119
INV_PRIN	0.056	0.021	0.016	0.019	0.019
FEES	0.298	0.072	0.029	0.048	0.048
MTGCO	0.003*	0.083	0.083	0.098	0.102
THRIFT	-0.285	-0.015	0.003**	0.019	0.026
NEW	0.123	0.037	0.029	0.030	0.029
CA	0.050	0.057	0.005	0.040	0.046
MD	-0.021	0.046	0.053	0.051	0.051
NJ	0.114	0.067	0.021	0.049	0.051
VA	-0.047	0.008	0.006	0.005***	0.004***
CM_1			0.101	0.085	0.115
CM_10			0.734	0.610	0.652
VOL_10			0.564	0.307	0.255
SPREAD			0.584	0.237	0.375
HPI_STD			0.008	0.003	0.003
D_1997		1.836		0.159	-0.105
D_1998		1.139		0.036	-0.194
D_1999		1.437		0.116	-0.122
D_2000		2.202		0.507	0.186
D_2001		1.118		0.182	-0.018
D_2002		0.704		0.182	0.045
FEB					0.014
MAR					-0.032
APR					0.015
MAY					-0.025
JUN					0.090
JUL					0.114
AUG					0.140
SEP					0.195
OCT					0.120
NOV					0.019
DEC					0.053
$\overline{R}^2$	16.33%	69.92%	78.15%	80.06%	80.69%
N	964,634	964,634	964,634	964,634	964,634

All coefficient estimates differ significantly from zero (two-tailed test) at the 1% confidence level, <u>except:</u>

*Not significantly different from zero

**Significant at the 10% level

***Significant at the 5% level

NOTES TO TABLE 2: Variations of our version of Hendershott and Shilling's [1989] first test. We report a variety of specifications of the form (3), to assess the impact of various specifications on the jumbo dummy coefficient. LTV_1, LTV_2, and LTV_3, equal unity when LTV exceeds 75% but not 80%, exceeds 80% but not 90%, and exceeds 90%, respectively, and are relative to LTV_0 (LTV≤75%). INV_PRIN is $1,000,000/PRIN. FEES indicates that positive fees and/or charges are associated with the loan. MTGCO and THRIFT estimates are relative to BANK. NEW is a dummy equal to unity when the loan funded new construction. State dummies are relative to OTHER_STATE. CM_1 (CM_10) is the monthly average of the 1-year (10-year) constant maturity U.S. Treasury. VOL_10 is the standard deviation of the daily yield on the 10-year Treasury over the preceding 60 trading days. SPREAD is the monthly industrial bond yield difference (Baa-Aaa). HPI_STD is the standard deviation of the OFHEO state-level house price index over the preceding eight quarters. Year dummies are relative to D_2003. Month dummies are relative to JAN. All estimates are statistically significant at the 1% level (two-tailed test), unless noted otherwise.

Table 3: Extensions of the basic regression specification

	(1)	(2)	(3)	(4)
INTERCEPT	2.627	2.626	2.627	2.627
J	0.226	0.240	0.241	0.269
LTV_1	0.016	0.019	0.019	0.019
LTV_2	0.102	0.103	0.103	0.104
LTV_3	0.119	0.118	0.119	0.119
INV_PRIN	0.019	0.019	0.019	0.019
CM_1	0.115	0.115	0.115	0.115
CM_10	0.652	0.652	0.652	0.653
VOL_10	0.255	0.255	0.254	0.256
SPREAD	0.375	0.375	0.375	0.374
HPI_STD	0.003	0.003	0.003	0.003
J_LTV_1		-0.032	-0.032	-0.040
J_LTV_2		-0.011***	-0.009**	-0.026
J_LTV_3		0.059	0.059	0.042
JUL_NC			-0.015*	-0.021*
AUG_NC			0.019*	0.013*
SEP_NC			0.057	0.053
OCT_NC			0.074	0.068
NOV_NC			-0.088	-0.097
DEC_NC			-0.238	-0.236
POST_REG				-0.072
$\overline{R}^2$	80.69%	80.70%	80.64%	80.65%
N	964,634	964,634	970,575	970,575

All coefficient estimates differ significantly from zero (two-tailed test) at the 1% confidence level, except:

*Not significantly different from zero

**Significant at the 10% level

***Significant at the 5% level

NOTES TO TABLE 3: Extensions of our base regression (Table 2, Column 5). Many coefficients are not reported here to save space, but are available upon request from the authors. LTV_1, LTV_2, and LTV_3, equal unity when LTV exceeds 75% but not 80%, exceeds 80% but not 90%, and exceeds 90%, respectively, and are relative to LTV_0 (LTV≤75%). INV_PRIN is $1,000,000/PRIN. CM_1 (CM_10) is the monthly average of the daily yields for the 1-year (10-year) constant maturity U.S. Treasury bond. VOL_10 is the standard deviation of the trailing 60-day yield on the 10-year constant maturity U.S. Treasury bond. SPREAD equals the monthly industrial bond yield difference (Baa-Aaa). HPI_STD is the standard deviation of the OFHEO's state-level house price index over the preceding eight quarters. The J_LTV_i dummies equal unity when the respective LTV_i dummies equal unity and JUMBO=1. The month_NC dummies equal unity when the loan was closed in the respective month and JUMBO=1, but PRIN does not exceed the next year's conforming loan limit. POST_REG is a dummy that equals unity when JUMBO=1 and the mortgage was closed after November 2001. All estimates are statistically significant at the 1% level (two-tailed test), unless noted otherwise.

Table 4: Pricing equations various loan size groups

	(1)	(2)	(3)	(4)	(5)	(6)
Loan Group:	25-100	25-80	80-100	NC-120	120-200	NC-200
INTERCEPT	2.611	2.606	2.617	3.242	3.263	3.270
LTV_1	0.018	0.019	0.009	-0.014***	0.002*	-0.002*
LTV_2	0.103	0.100	0.120	0.085	0.097	0.094
LTV_3	0.118	0.114	0.147	0.171	0.194	0.185
INV_PRIN	0.019	0.019	0.032	0.029*	0.022	0.019
FEES	0.048	0.052	0.025	0.041	0.035	0.037
MTGCO	0.105	0.112	0.048	0.026***	0.043	0.039
THRIFT	0.025	0.027	0.000*	-0.012*	0.044	0.028
NEW	0.028	0.025	0.044	0.054	0.042	0.045
CA	0.046	0.049	0.038	0.065	0.035	0.044
MD	0.051	0.039	0.087	0.078	0.045	0.055
NJ	0.052	0.050	0.062	0.028***	0.037	0.035
VA	0.006	0.001*	0.029	-0.027***	-0.010*	-0.015***
CM_1	0.111	0.113	0.104	0.213	0.217	0.216
CM_10	0.658	0.657	0.660	0.520	0.516	0.517
VOL_10	0.255	0.250	0.283	0.249	0.288	0.279
SPREAD_10	0.372	0.372	0.371	0.424	0.403	0.407
HPI_STD	0.003	0.003	0.002	0.002	0.001	0.001
D_1997	-0.103	-0.111	-0.078	-0.216	-0.209	-0.210
D_1998	-0.198	-0.210	-0.148	-0.261	-0.235	-0.242
D_1999	-0.120	-0.129	-0.080	-0.255	-0.236	-0.242
D_2000	0.187	0.179	0.223	0.073*	0.072***	0.071
D_2001	-0.024	-0.032	0.018***	0.047**	0.054	0.051
D_2002	0.042	0.040	0.057	0.088	0.089	0.087
FEB	0.015	0.017	0.004*	-0.017*	-0.007*	-0.010*
MAR	-0.029	-0.027	-0.041	-0.088	-0.097	-0.095
APR	0.018	0.020	0.003*	-0.035***	-0.060	-0.053
MAY	-0.021	-0.019	-0.038	-0.096	-0.097	-0.097
JUN	0.094	0.097	0.078	-0.011*	0.004*	-0.000*
JUL	0.118	0.121	0.101	0.028**	0.028	0.027
AUG	0.145	0.148	0.127	0.021*	0.045	0.037
SEP	0.198	0.202	0.180	0.110	0.111	0.111
OCT	0.122	0.124	0.104	0.084	0.074	0.078
NOV	0.020	0.024	-0.003*	-0.006*	-0.019**	-0.016**
DEC	0.056	0.059	0.033	0.008*	0.001*	0.003*
$\overline{R}^2$	80.79%	80.61%	81.86%	80.27%	80.07%	80.11%
N	911,956	779,940	132,016	15,042	37,636	52,678

All coefficient estimates differ significantly from zero (two-tailed test) at the 1% confidence level, except:

*Not statistically different from zero

**Significant at the 10% level

***Significant at the 5% level

NOTES TO TABLE 4: Results for regressions on six size subsets of our primary sample. LTV_1, LTV_2, and LTV_3, equal unity when LTV exceeds 75% but not 80%, exceeds 80% but not 90%, and exceeds 90%, respectively. INV_PRIN is $1,000,000/PRIN. FEES indicates that positive fees and/or charges are associated with the loan. MTGCO and THRIFT estimates are relative to BANK. NEW indicates that the loan funded the purchase of new construction. State dummies are relative to OTHER_STATE. CM_1 (CM_10) is the monthly average of the 1-year (10-year) constant maturity U.S. Treasury bond. VOL_10 is the standard deviation of the daily yield on the 10-year Treasury over the preceding 60 trading days. SPREAD is the monthly industrial bond yield difference (Baa-Aaa). HPI_STD is the standard deviation of the OFHEO's state-level house price index over the preceding eight quarters. Year dummies are relative to D_2003. Month dummies are relative to JAN. All estimates are statistically significant at the 1% level (two-tailed test), unless noted otherwise.

Table 5: Decomposition Method of Estimating the Jumbo-Conforming Spread

Cells report {[Mean actual RM] – [Mean fitted RM]}

		the comparison size group, whose coefficients are applied			
		25-80	**80-100**	**NC-120**	**120-200**
the base size group, from which loan properties are taken	**25-80**		-0.059	-0.291	-0.264
	80-100	-0.003		-0.250	-0.245
	NC-120	0.222	0.235		-0.004
	120-200	0.226	0.251	0.005	

NOTES TO TABLE 5: The lower-left triangle reports the mean values of $\boldsymbol{(b-a)w^{j}_{it}}$ from equation (6), where *j* is the base size group (whose coefficients are ***b*)** and *c* is the comparison group (whose coefficients are ***a***). The groups are those defined in Table 4: 25-80, 80-100, NC – 120, 120-200 percent of the conforming limit. For example, the bottom entry in the first column (0.226) indicates that the average loan between 120 and 200 percent of the conforming loan limit would have been priced 21.8 bps lower had it been smaller (between 25 and 80% of the limit) with the same other features. Conversely, the upper-right triangle reports the mean values of $\boldsymbol{(b-a)w^{c}_{it}}$ from equation (7). For example, the top right cell indicates that loans between 25 and 80 percent of the conforming limit would have been priced 26.4 bps higher had they been valued like loans from the 120-200 size group.

Table 6: Various treatments of scale economies

	(1)	(2)	(3)	(4)
INTERCEPT	2.627	2.479	2.744	2.767
J	0.241	0.290	0.174	0.153
LTV_1	0.019	0.017	0.015	0.008
LTV_2	0.103	0.101	0.101	0.095
LTV_3	0.119	0.116	0.119	0.118
INV_PRIN	0.019			
ln(PRIN)		-0.149		
SMALL			0.102	
CM_1	0.115	0.115	0.114	0.113
CM_10	0.652	0.652	0.654	0.657
VOL_10	0.254	0.254	0.256	0.259
SPREAD	0.375	0.374	0.378	0.377
HPI_STD	0.003	0.003	0.002	0.001
J_LTV_1	-0.032	-0.034	-0.029	-0.025
J_LTV_2	-0.009**	-0.018	-0.008*	-0.009**
J_LTV_3	0.059	0.045	0.062	0.057
JUL_N	-0.015*	-0.038***	-0.002*	-0.005*
AUG_N	0.019*	-0.004*	0.033***	0.031**
SEP_N	0.057	0.035***	0.071	0.071
OCT_N	0.074	0.052	0.089	0.089
NOV_N	-0.088	-0.110	-0.073	-0.072
DEC_N	-0.238	-0.262	-0.223	-0.223
$\bar{R}^2$	80.64%	80.63%	80.47%	80.18%
N	970,575	970,575	970,575	970,575

All coefficient estimates differ significantly from zero (two-tailed test) at the 1% confidence level, except:

*Not significantly different from zero

**Significant at the 10% level

***Significant at the 5% level

NOTES TO TABLE 6: Various treatments of scale economies, based on our Table 3, Column 3 regression. Many coefficients are not reported here to save space, but are available upon request from the authors. LTV_1, LTV_2, and LTV_3, equal unity when LTV exceeds 75% but not 80%, exceeds 80% but not 90%, and exceeds 90%, respectively, and are relative to LTV_0 (LTV≤75%). INV_PRIN is $1,000,000/PRIN. ln(PRIN) is the natural logarithm of PRIN/$1,000,000. The SMALL dummy equals unity when PRIN < $100,000. CM_1 (CM_10) is the monthly average of the daily yields for the 1-year (10-year) constant maturity U.S. Treasury bond. VOL_10 is the standard deviation of the trailing 60-day yield on the 10-year constant maturity U.S. Treasury bond. SPREAD equals the monthly industrial bond yield difference (Baa-Aaa). HPI_STD is the standard deviation of the OFHEO's state-level house price index over the preceding eight quarters. The J_LTV_i dummies equal unity when the respective LTV_i dummies equal unity and JUMBO=1. The month_NC dummies equal unity when the loan was closed in the respective month and JUMBO=1, but PRIN does not exceed the next year's conforming loan limit. POST_REG is a dummy that equals unity when JUMBO=1 and the mortgage was closed after November 2001. All estimates are statistically significant at the 1% level (two-tailed test), unless noted otherwise.