2004 V32 4: pp. 541–569
REAL ESTATE
ECONOMICS

The Effect of Conforming Loan Status on Mortgage Yield Spreads: A Loan Level Analysis

Brent W. Ambrose,* Michael LaCour-Little** and Anthony B. Sanders***

The magnitude of the effect of government-sponsored enterprise purchases on primary mortgage market rates has been a difficult research question with differing data and competing methodologies producing varying results. Here we present a new approach using loan level data and controlling for credit risk differentials between conforming and nonconforming loans. Our method also addresses econometric problems of endogeneity and sample selection bias. We find that conforming loans have yield spreads about 5.5% lower compared to other loans on a risk-adjusted basis. This is lower than previous estimates appearing in the literature.

Lenders in the primary mortgage market originate a range of contracts, many of which are then sold in the secondary market, either to the government-sponsored enterprises (Fannie Mae and Freddie Mac, hereafter the GSEs[1]) or pooled as collateral for private label mortgage-backed securities. Other loans are retained in portfolio or sold on a whole-loan basis. Loans vary in coupon, size, term, collateral and, of course, credit quality. Based on business strategy and risk preferences, together with information obtained during the underwriting process, and considering rules and price signals from the secondary market, lenders make a hold-versus-sell decision. What factors determine that decision, and what is the overall effect on rates in the primary market?

These questions are among many that are crucial to evaluating the role of the GSEs, whose special status in the economy generates both costs and benefits (Sanders 2002). GSE purchases provide liquidity to primary lenders and perhaps stability to the overall mortgage market (González-Rivera 2001, Naranjo and

*Gatton College of Business and Economics, University of Kentucky, Lexington, KY 40506-0053 or ambrose@uky.edu.
**Wells Fargo Home Mortgage and Washington University, St. Louis, Clayton, MO 63105 or michael.lacour-little@mortgage.wellsfargo.com.
***Finance Department, The Ohio State University, Columbus, OH 43210 or tsanders@postbox.acs.ohio-state.edu.

[1] There are other GSEs operating in the housing finance arena, including the Federal Home Loan Banks.

Toevs 2002). In addition, it is widely accepted that GSE activity reduces rates on conforming loans by expanding available funds to lenders (Phillips 1996). But the extent to which the GSEs reduce mortgage rates is controversial.[2] For example, Ambrose, Buttimer and Thibodeau (2001) show the impact of house price volatility on the jumbo/non-jumbo loan rate differential. Their simulations indicate that as much as 20% of the loan rate differential may be due to house price volatility.

On the other hand, GSEs benefit from an implied federal guaranty of their liabilities and are exempt from certain taxes and requirements that other financial intermediaries bear. A number of papers have examined the funding advantage of the GSE, most recently Ambrose and Warga (2002) and Nothaft, Pearce and Sevanovic (2002). The latter identifies seven prior studies, which provide a range of estimates of the funding advantage of between 23 and 72 basis points, depending on data used, comparison instrument and methodology, and the authors provide their own estimate of 27–30 basis points. Ambrose and Warga provide a broader comparison across risk classes and estimate an average funding advantage of 25–29 basis points over "AA"-rated banking sector bonds, 43–47 basis points over "A"-rated banking bonds and 76–80 basis points over "BBB"-rated banking sector bonds.

Our effort here again focuses on the benefit side of the ledger, that is, the rate reduction associated with conforming loan status. In contrast to research that relies on macro-level yield data, we return to the loan level approach as in Hendershott and Shilling (1989). But our data are more recent and include borrower credit score, a key piece of information that allows us to estimate the reduction in mortgage yield spreads attributable to conforming loan status on a risk-adjusted basis. In addition, we can more precisely separate conforming from nonconforming loans, a task that cannot be accomplished with the data used by McKenzie (2002), Ambrose and Buttimer (2004) and Passmore, Sparks and Ingpen (2002).[3] Finally, our approach corrects for two distinct econometric problems: endogeneity and sample selection bias. Endogeneity occurs because the loan-to-value ratio is jointly determined with note rate. Sample selection bias may occur because some loans that could be sold to the GSEs may not be.

---

[2] A long line of research has addressed this question, including Hendershott and Shilling (1989), Ambrose and Warga (1995), Cotterman and Pierce (1996), Kolari, Fraser and Anari (1998), Roll (2000), Todd (2001), González-Rivera (2001), Ambrose, Buttimer and Thibodeau (2001), Ambrose and Buttimer (2004), McKenzie (2002), Naranjo and Toevs (2002) and Passmore, Sparks and Ingpen (2002).

[3] The MIRS data, which is a sample of actual loan prices, only allows separation of loans into jumbo versus not-jumbo based on the conforming loan size limit, thereby missing most loans of conforming loan size which are ineligible for GSE purchase due to credit factors.

If the GSEs purchase lower risk loans, then a simple comparison of yields will be confounded by any credit risk differential.

To address this question we develop an *ex ante* model of yield spreads controlling for credit risk at the loan level. The model incorporates a variety of characteristics including credit score, borrower age and income and loan-to-value ratio (LTV). We also incorporate actual outcomes, that is, whether the loan was, in fact, sold into the secondary mortgage market or retained in portfolio by the originating lender, as well as bond market environmental factors.

The plan for the remainder of the paper is as follows. In the next section, we briefly review relevant literature, drawing analogies to the corporate bond market. In the second section, we sketch out the theory of mortgage valuation and develop our model of mortgage yield spreads. In the third section, we describe the data. The fourth section presents the basic regression model that is comparable to the existing literature. We then refine the model to control for credit risk and address econometric issues. The final section offers conclusions.

Literature Review

A number of studies have examined the conforming-nonconforming rate differential. In most cases rates on jumbo loans are taken as the relevant nonconforming loan category, although the reality is that some conforming loan size loans are, in fact, nonconforming, due to credit or documentation issues (*e.g.*, subprime or low-doc loans). Most previous studies rely on the Federal Housing Finance Board's monthly mortgage interest rate survey (MIRS), which only allows separation into jumbo versus non-jumbo categories and lacks important credit risk measures. McKenzie (2002) provides an excellent discussion of the issues associated with estimating the loan rate differential using the MIRS data. McKenzie also provides a summary of previous empirical estimates. Depending on the period examined and methodology employed, the jumbo/non-jumbo mortgage rate differential has ranged between a high of 60 basis points (Cotterman and Pearce 1996) to a low of 8 basis points (Naranjo and Toevs 2002).

In one of the first studies using MIRS data, Hendershott and Shilling (1989) found that conforming loans had interest rates 24 to 39 basis points lower than nonconforming loans after controlling for loan characteristics. They regressed effective mortgage interest rate against a set of variables to control for jumbo loan status, loan size, loan-to-value ratio, new versus existing home status, as well as dummy variables to capture temporal and regional variations. Subsequent studies (*e.g.*, ICF Inc. 1990, Cotterman and Pearce 1996, Ambrose, Buttimer and Thibodeau 2001, U.S. Congressional Budget Office 2001, McKenzie 2002, Naranjo and Toevs 2002, Passmore, Sparks and Ingpen

2002) have followed a similar methodology with minor variations in data screens designed to isolate the conforming/nonconforming effect as well as geographic differences (*e.g.*, McKenzie 2002, Ambrose and Buttimer 2004). Todd (2001) follows the Hendershott and Shilling methodology adding in the effect of origination costs using Freddie Mac and Federal Reserve aggregate data. Naranjo and Toevs (2002) extend the Hendershott and Shilling methodology to incorporate cointegration to correct for nonstationarity in rates.[4]

Studies of the determinants of yield spreads in the corporate bond literature have focused on credit spreads (see Altman and Saunders 1998 for a review). For mortgages, credit risk has traditionally been viewed as a function of borrower equity or loan-to-value ratio. Early research includes von Furstenberg (1969), von Furstenberg and Green (1974), Campbell and Dietrich (1983) and Cunningham and Capone (1990). Quercia and Stegman (1992) and Vandell (1993) provide surveys focusing on default, while Kau, Keenan and Kim (1994) develop theoretical default given stochastic collateral values. Among recent methodological innovations, Deng, Quigley and Van Order (2000) present a competing risk model of mortgage termination, both default and prepayment, accounting for unobserved borrower heterogeneity. Mortgage default research has also recently begun to incorporate borrower credit score (*e.g.*, Avery *et al.* 1996). In a related line of research outside of the mortgage literature, Angbazo, Mei and Saunders (1998) examine yield spreads for highly leveraged corporate loans.

Linking the mortgage literature to the broader finance literature, yield spreads on the firm's debt reflect underlying financial risk that depends on firm capital structure, just as default risk in mortgages is related to the borrower's equity. Likewise, we may think of borrower credit score as the analog to firm credit rating and individual borrower demographic characteristics as the analog to firm specific idiosyncratic risk factors.

Theoretical Predications and Model

The general approach to mortgage pricing is now well established. Mortgages are contingent claims contracts in which the mortgage value (V^M) depends critically on two stochastic processes, the market interest rate, $r(t)$, and the house value, $H(t)$. For instance, we may specify the interest rate process with a CIR diffusion process:

$$d(r) = \gamma(\Theta - r)\,dt + \sigma_r\sqrt{r}\,dz_r, \tag{1}$$

[4] Kolari, Fraser and Anari (1998) also utilize cointegration methods to investigate the impact of securitization on mortgage yield spreads using data from MIRS. However, their analysis does not examine the differential between conforming and nonconforming loans.

where Θ is the steady-state mean rate, γ is the speed of adjustment factor and σ_r is the volatility of interest rates.[5] Diffusion in collateral value affects mortgage value, too, so we may specify the evolution of house values, $H(t)$, by

$$\frac{dH}{H} = (\alpha - s)\,dt + \sigma_H\,dz_H, \tag{2}$$

where α is the instantaneous total return to housing, s is the service flow and σ_H is the volatility of housing returns. In (1) and (2), dz_r and dz_H are standard Wiener processes and the correlation between the movements of the two state variables (dz_H and $dz_{r)}$) is ρ.

Kau and Keenan (1995) show that under appropriate assumptions, the value of the mortgage (V^M) will satisfy the following partial differential equation (PDE):

$$\begin{aligned} &\frac{1}{2}H^2\sigma_H^2\frac{\partial^2 V^M}{\partial H^2} + \rho H\sqrt{r}\sigma_H\sigma_r\frac{\partial^2 V^M}{\partial H\partial r} + \frac{1}{2}r\sigma_r^2\frac{\partial^2 V^M}{\partial r^2} \\ &\quad + \gamma(\theta - r)\frac{\partial V^M}{\partial r} + (r - s)H\frac{\partial V^M}{\partial H} + \frac{\partial V^M}{\partial t} - rV^M = 0. \end{aligned} \tag{3}$$

Specifying the boundary conditions allows the valuation of the mortgage when the economic variables take on extreme values. With these boundary conditions, (3) can be solved to find the value of the mortgage contract.

We denote the present value of the remaining mortgage payments as $A(r(t), t)$. Since house prices and interest rates interact in determining the value of the right to terminate the mortgage, $J(r(t), H(t), t)$, the mortgage value is given as

$$V^M(r(t), H(t), t) = A(r(t), t) - J(r(t), H(t), t). \tag{4}$$

The right to terminate the mortgage is composed of the right to prepay the mortgage and the right to default.[6] Standard comparative statics show that $\partial V^M/\partial\sigma_H < 0$, $\partial V^M/\partial\sigma_r < 0$ and $\partial V^M/\partial H > 0$.[7] Assuming that V^M represents the mortgage value at origination (net of any discount points), then the yield

[5] Equation (1) is the standard Cox, Ingersoll and Ross (1985) interest rate model.

[6] Kau, Keenan and Kim (1993, 1994) show that the option pricing technique can be utilized to determine the probability of default.

[7] See Kau, Keenan, Muller and Epperson (1992, 1993) for a complete discussion of the fixed-rate mortgage contract comparative statics.

(y^M) is simply the internal rate of return that equates the expected mortgage payments (principal and interest) over the expected holding period to V^M.[8]

Following Merton (1974), we define the mortgage credit risk premium as the difference between the yield and the risk-free rate ($y^M - r$), where the Treasury bond rate serves as a proxy for the risk-free rate. Although Merton examined the relationship between the risk premium on discount debt issued by a firm, the comparative statics from the mortgage pricing model (4) and Merton's analysis of risky debt imply that the yield spread ($y^M - r$) is a function of the volatility of the underlying state variables (house values and interest rates) as well as the loan-to-value ratio at origination ($V^M(r(0), H(0), 0)/H(0)$). Thus, consistent with Merton, we have that

$$\partial S/\partial \sigma_H > 0, \quad \partial S/\partial \sigma_r > 0, \quad \text{and} \quad \partial S/\partial (V^M(r(0), H(0), 0)/H(0)) > 0$$

where $S = (y^M - r)$.

Many models of yield spreads have been developed in the corporate bond market following Merton (1974).[9] For example, Bakshi, Madan and Zhang (2000) test for the presence of firm-specific distress factors in discount rate models for corporate bonds. Results confirm that firm-specific factors (such as leverage and book-to-market ratios) as well as market interest rates explain differences in corporate bond yields. We follow this example and propose the following generalized model of the yield spread (S):

$$\begin{aligned} \ln S_i &= \alpha_0 + \alpha_1 \sigma_{H_i} + \alpha_2 \sigma_{r_i} + \alpha_3 \left(V_i^m / H_i \right) + \alpha_4 CONFORM_i \\ &\quad + \alpha_5 X_i + \sum_{j=1}^{11} \delta_j QTR_{ji} + \varepsilon_i \end{aligned} \tag{5}$$

where $CONFORM_i$ is a dummy variable denoting mortgages that meet the GSE conforming guidelines, X_i represents a vector of borrower-specific and market factors that may impact yield spreads, YR_i is a dummy variable denoting the year of origination and MON_i is a dummy variable denoting the month of origination.[10] Including the dummy variable *CONFORM* in (5) sets up our test

[8] In the empirical analysis, we assume a 10-year expected holding period rather than the full 30-year mortgage term.

[9] For example, see Duffie (1998, 1999), Collin-Dufresne, Goldstein and Martin (2001), Duffie and Singleton (1999), Bakshi, Madan and Zhang (2000) and Ericsson and Renault (2000).

[10] We include the yearly dummy variable to capture any variation in market credit conditions not specifically identified. We include the monthly dummy variable to capture any seasonal effects associated with mortgage origination.

of whether origination spreads are lower for loans that are eligible for GSE purchase compared to those which are not.

Data

Data used for this research consist of origination records on 26,179 conventional fixed-rate mortgages made between January 1995 and December 1997 by a national lender, the lender's correspondent lenders and mortgage brokers. Both conforming and nonconforming loans are included, although super-jumbos (loans with initial balances in excess of $650,000) are not. Table 1 reports the distribution of the loans by origination year and geography. We also compare the distribution of loans to the MIRS and Home Mortgage Disclosure Act (HMDA) databases for 1997 to ensure that our data are reasonably representative. Cross-sectionally, our sample is somewhat overrepresented by loans in New York/New Jersey and underrepresented in the Northwest and Southwest regions. However, the distribution of mortgage originations across years does follow the same pattern as that observed in the broader market.

We have relatively complete micro-level data for each loan in the sample. In addition to geography, loan amount and note rate, we observe points paid at the time of loan origination so we can correctly compute yields over a given horizon. Major credit quality indicators, loan-to-value ratio (*LTV*) and borrower credit score at time of origination (*FICO*), are also available, as is whether the loan had private mortgage insurance. Borrower demographic characteristics available include age (*BRWAGE*) and income (*INCOME*). Table 2 reports the descriptive statistics (mean, median and standard deviation) for the sample by origination year. Across all origination years, the mean *LTV* is 75.6%. About 69% have *LTV*s below 80%, 15% have *LTV*s between 80% and 90% and 16% have *LTV*s greater than 90% (see Table 2).

Across time, credit quality appears to have increased, with declining LTVs and increasing credit scores. The average FICO score increased from 715 in 1995 to 722 in 1997 and averaged 720 over the full sample. Average borrower age at origination (our proxy for wealth) increased from 40 years in 1995 to 42 years in 1997 and average borrower income at origination increased from $62,510 in 1995 to $87,960 in 1997. Consistent with housing market appreciation and a more affluent borrower pool, loan amounts increased over time from $107,700 in 1995 to $141,500 in 1997.

Our data contain both conforming and nonconforming loans; however, while we observe outcomes (*i.e.*, whether sold to Fannie Mae or Freddie Mac, sold into a private label MBS or retained in portfolio), we cannot precisely determine loan status at origination. About 71% of the loans were actually sold to the agencies

Table 1 ■ Distribution of mortgage loans by origination year and location.

	Sample				MIRS				HMDA
	1995 % Total Year	1996 % Total Year	1997 % Total Year	All Years % of Total	1995 % Total Year	1996 % Total Year	1997 % Total Year	All Years % of Total	1997 % Total Year
Region 1: New England									
CT	1.7%	2.3%	2.7%	2.4%	2.0%	1.9%	1.2%	1.6%	1.18%
ME	0.3%	0.2%	0.2%	0.2%	0.3%	0.3%	0.3%	0.3%	0.35%
MA	1.2%	1.5%	2.0%	1.7%	1.8%	2.1%	2.2%	2.1%	2.38%
NH	0.4%	0.2%	0.3%	0.3%	0.3%	0.3%	0.3%	0.3%	0.40%
RI	0.2%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.33%
VT	0.1%	0.1%	0.1%	0.1%	0.2%	0.2%	0.1%	0.2%	0.18%
Total	4.0%	4.4%	5.4%	4.8%	4.7%	4.9%	4.3%	4.6%	4.82%
Region 2: New York/New Jersey									
NJ	6.5%	6.4%	5.5%	6.0%	3.3%	3.7%	3.9%	3.7%	4.05%
NY	23.7%	33.2%	22.2%	25.1%	2.9%	2.5%	2.0%	2.4%	2.71%
Total	30.2%	39.5%	27.7%	31.0%	6.2%	6.2%	6.0%	6.1%	6.76%
Region 3: Mid-Atlantic									
DE	0.2%	0.4%	0.2%	0.2%	0.5%	0.5%	0.5%	0.5%	0.31%
DC	0.7%	0.7%	0.8%	0.7%	0.1%	0.1%	0.2%	0.2%	0.19%
MD	3.2%	2.8%	4.0%	3.5%	2.6%	2.6%	2.4%	2.5%	1.94%
PA	2.9%	2.2%	2.0%	2.2%	5.9%	5.2%	3.8%	4.8%	4.15%
VA	1.9%	2.7%	3.2%	2.7%	3.2%	3.1%	3.4%	3.2%	2.40%
WV	0.1%	0.1%	0.1%	0.1%	0.2%	0.2%	0.2%	0.2%	0.43%
Total	8.9%	8.8%	10.3%	9.6%	12.5%	11.7%	10.5%	11.4%	9.42%
Region 4: Southeast									
AL	0.4%	1.2%	1.0%	0.9%	0.6%	0.6%	0.5%	0.5%	1.37%
FL	6.3%	7.4%	5.6%	6.2%	7.8%	7.4%	5.7%	6.9%	6.40%
GA	1.2%	1.1%	1.4%	1.3%	1.5%	1.6%	1.9%	1.7%	3.06%

KY	0.3%	0.2%	0.2%	0.2%	0.9%	0.8%	0.8%	0.8%	1.37%
MS	0.1%	0.1%	0.1%	0.1%	0.5%	0.5%	0.4%	0.4%	0.85%
NC	1.4%	1.7%	2.1%	1.9%	3.0%	3.0%	3.1%	3.0%	3.20%
SC	0.5%	0.4%	0.5%	0.4%	1.6%	1.9%	1.6%	1.7%	1.60%
TN	0.8%	0.9%	0.5%	0.7%	1.3%	1.0%	1.2%	1.2%	1.88%
Total	11.1%	13.0%	11.3%	11.7%	17.1%	16.8%	15.2%	16.2%	19.72%
Region 5: Midwest									
IL	11.0%	9.5%	8.7%	9.5%	6.9%	5.8%	4.5%	5.6%	5.25%
IN	0.5%	0.3%	0.4%	0.4%	1.9%	1.9%	1.8%	1.9%	2.59%
MI	1.2%	1.0%	1.0%	1.0%	3.3%	3.2%	2.4%	2.9%	5.12%
MN	0.6%	0.7%	0.7%	0.7%	0.7%	0.8%	2.9%	1.6%	1.98%
OH	1.9%	0.8%	2.1%	1.8%	4.8%	4.6%	4.2%	4.5%	4.96%
WI	0.2%	0.1%	0.1%	0.1%	3.7%	3.6%	3.5%	3.6%	2.34%
Total	15.4%	12.4%	13.0%	13.5%	21.4%	19.9%	19.3%	20.1%	22.24%
Region 6: Southwest									
AR	0.7%	0.4%	0.4%	0.5%	0.2%	0.2%	0.3%	0.2%	0.69%
LA	0.3%	0.2%	0.3%	0.3%	0.4%	0.4%	0.4%	0.4%	1.29%
NM	0.3%	0.4%	0.6%	0.5%	0.3%	0.3%	0.5%	0.4%	0.66%
OK	0.4%	0.4%	0.4%	0.4%	0.6%	0.6%	0.6%	0.6%	0.87%
TX	2.7%	1.6%	2.2%	2.2%	6.2%	6.1%	5.1%	5.7%	4.70%
Total	4.5%	3.1%	3.9%	3.8%	7.7%	7.7%	6.9%	7.4%	8.22%
Region 7: Great Plains									
IA	0.5%	0.3%	0.5%	0.5%	0.6%	0.7%	1.3%	0.9%	0.86%
KS	0.5%	0.4%	0.8%	0.6%	1.1%	1.7%	1.3%	1.3%	0.82%
MO	4.2%	2.0%	2.1%	2.6%	2.5%	2.4%	1.6%	2.1%	2.10%
NE	0.1%	0.0%	0.0%	0.0%	0.4%	0.5%	0.6%	0.5%	0.49%
Total	5.3%	2.7%	3.4%	3.7%	4.6%	5.2%	4.8%	4.9%	4.26%
Region 8: Rocky Mountain									
CO	0.8%	0.9%	1.4%	1.1%	1.7%	1.7%	2.8%	2.2%	2.62%
MT	0.0%	0.0%	0.0%	0.0%	0.1%	0.2%	0.4%	0.3%	0.27%
ND	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.2%	0.1%	0.13%

Table 1 ■ continued.

	Sample				MIRS				HMDA
	1995 % Total Year	1996 % Total Year	1997 % Total Year	All Years % of Total	1995 % Total Year	1996 % Total Year	1997 % Total Year	All Years % of Total	1997 % Total Year
SD	0.2%	0.2%	0.1%	0.1%	0.1%	0.0%	0.3%	0.2%	0.20%
UT	0.6%	0.3%	0.4%	0.5%	0.2%	0.4%	1.0%	0.6%	1.26%
WY	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.2%	0.1%	0.16%
Total	1.7%	1.5%	2.0%	1.8%	2.3%	2.4%	4.9%	3.4%	4.64%
Region 9: Pacific									
AZ	0.0%	0.0%	0.0%	0.0%	2.2%	2.0%	2.7%	2.3%	2.29%
CA	14.5%	12.0%	18.4%	16.0%	13.2%	13.7%	14.1%	13.7%	11.47%
HI	0.0%	0.0%	0.0%	0.0%	0.5%	0.5%	0.5%	0.5%	0.31%
NV	2.2%	1.1%	1.3%	1.5%	1.5%	1.4%	1.4%	1.4%	0.89%
Total	16.7%	13.2%	19.8%	17.5%	17.3%	17.6%	18.7%	18.0%	14.96%
Region 10: Northwest									
AK	1.0%	0.7%	1.4%	1.1%	0.0%	0.0%	0.1%	0.1%	0.12%
ID	0.0%	0.0%	0.1%	0.1%	0.3%	0.4%	0.6%	0.4%	0.48%
OR	0.5%	0.2%	0.5%	0.4%	2.1%	2.4%	2.8%	2.5%	1.70%
WA	0.7%	0.4%	1.2%	0.9%	3.7%	4.7%	6.0%	4.9%	2.66%
Total	2.2%	1.4%	3.2%	2.6%	6.1%	7.5%	9.5%	7.9%	4.95%
Total #	6,600	6,030	13,549	26,179	128,782	133,458	183,052	445,292	7,796,111

This table reports the frequency distribution by year and location for the sample of 26,179 conventional fixed-rate mortgage loans originated between January 1995 and December 1997. Both conforming and nonconforming loans are included, although super-jumbos (loans with initial balances in excess of $650,000) are not. For comparison, we also report the distribution of loans in the Mortgage Interest Rate Survey (MIRS) compiled by the Federal Housing Finance Board and the distribution of all loans reported in the 1997 Home Mortgage Disclosure Act (HMDA) database.

Table 2 ■ Descriptive statistics (standard deviations reported in parentheses).

Variable	Label	Full Sample ($N = 14285$)		1995 ($N = 4695$)		1996 ($N = 3265$)		1997 ($N = 6325$)		Difference in Means F-test
		Median	Mean	Median	Mean	Median	Mean	Median	Mean	
YIELD	Effective yield (YTM)	7.875	7.913	8.015	8.123	8.147	8.108	7.653	7.725	2369.7***
			(0.489)		(0.546)		(0.457)		(0.388)	
YIELD_10	Effective yield (10 yr hold)	7.875	7.937	8.030	8.150	8.186	8.135	7.678	7.745	2543.1***
			(0.484)		(0.540)		(0.449)		(0.380)	
YLDSPD	*YIELD_10*–10-yr	1.604	1.625	1.820	1.815	1.610	1.611	1.530	1.538	1478.8***
	Treasury rate	(0.355)		(0.397)		(0.357)		(0.290)		
LTV	Average LTV	0.797	0.756	0.800	0.783	0.796	0.752	0.783	0.744	135.7***
			(0.163)		(0.171)		(0.162)		(0.158)	
LTV80	LTV $\leq$ 80%	1.000	0.685	1.000	0.561	1.000	0.696	1.000	0.740	348.1***
			(0.465)		(0.496)		(0.460)		(0.438)	
LTV8090	LTV $>$ 80% and LTV $\leq$ 90%	0.000	0.152	0.000	0.155	0.000	0.164	0.000	0.145	6.2**
			(0.359)		(0.496)		(0.460)		(0.438)	
LTV9095	LTV $>$ 90% and LTV $\leq$ 95%	0.000	0.125	0.000	0.230	0.000	0.098	0.000	0.087	471.2***
			(0.331)		(0.421)		(0.297)		(0.281)	
LTV95	LTV $>$ 95%	0.000	0.038	0.000	0.054	0.000	0.042	0.000	0.028	47.2***
			(0.191)		(0.227)		(0.201)		(0.166)	

Table 2 ■ continued.

Variable	Label	Full Sample ($N = 14285$)		1995 ($N = 4695$)		1996 ($N = 3265$)		1997 ($N = 6325$)		Difference in Means
		Median	Mean	Median	Mean	Median	Mean	Median	Mean	F-test
FICO	Fair Isaac credit score	730.000	720.359 (56.726)	725.000	715.833 (59.268)	731.000	721.815 (55.496)	733.000	721.915 (55.883)	33.7***
BRWAGE	Borrower age	39.000	41.501 (11.307)	38.000	40.258 (11.317)	40.000	41.698 (11.337)	40.000	42.018 (11.244)	54.3***
INCOME	Borrower income	62.148	79.592 (96.773)	51.660	62.510 (51.498)	62.496	79.478 (87.206)	69.828	87.963 (114.922)	160.8***
ORIBAL	Loan amount	128,000	153,934 (99,023)	107,700	121,871 (73,496)	128,000	152,679 (97,332)	141,500	170,111 (106,473)	570.5***
AGENCY	1 = originated for sale to agency	1.000	0.709 (0.454)	1.000	0.693 (0.461)	1.000	0.727 (0.445)	1.000	0.709 (0.454)	8.3***
PORTFOLIO	1 = originated for retained portfolio	0.000	0.030 (0.170)	0.000	0.061 (0.239)	0.000	0.028 (0.164)	0.000	0.016 (0.124)	147.5***
PVTLABEL	1 = originated as private label mbs	0.000	0.261 (0.439)	0.000	0.245 (0.430)	0.000	0.245 (0.430)	0.000	0.275 (0.447)	14.6***
CONFORM	1 = conforming loan	1.000	0.755 (0.430)	1.000	0.760 (0.427)	1.000	0.778 (0.415)	1.000	0.742 (0.437)	15.5***

This table reports the descriptive statistics for the sample of 26,179 conventional fixed-rate mortgage loans originated between January 1995 and December 1997. Both conforming and nonconforming loans are included, although super-jumbos (loans with initial balances in excess of $650,000) are not. Loans were originated throughout the United States. ***significant at the 1% level, **significant at the 5% level, *significant at the 10% level.

(Fannie Mae and Freddie Mac), 26% were sold as private label MBS and 3% were retained in portfolio. It seems reasonable to assume that loans actually sold to the GSEs were conforming at origination and that jumbo loans were not. The question is whether the small number of conforming-size loans that were not sold to the GSEs could have been. We use some admittedly arbitrary criteria to make this determination. Since mortgages to borrowers with credit scores lower than 620 are classified by regulatory agencies as "subprime" and not usually purchased by the GSEs without significant credit enhancement, we use this value as one credit quality cutoff. We then classify a mortgage as conforming if (1) the loan was sold to one of the GSEs, (2) the borrower FICO score at origination is above 620, the loan amount was below the conforming loan limit in place at time of origination and the LTV is less than 80% or (3) the borrower FICO score at origination is above 620, the loan amount was below the conforming loan limit in place at time of origination and the loan has private mortgage insurance if the LTV is greater than 80%.[11] Based on this classification scheme, we see that approximately 80% of the mortgages were conforming loans at origination and were thus eligible for sale to the agencies. Hence, by our calculations, about 90% of conforming production was ultimately sold to the GSEs.[12]

Table 2 reports the mean (and median) yield and yield spreads for the sample. We report both the effective yield-to-maturity (*YIELD*) as well as the effective yield assuming a 10-year expected holding period (*YIELD_10*). The effective yield controls for any points paid by the borrower at origination. In the subsequent analysis, we use *YIELD_10* since most loans are repaid prior to maturity. The yield spread (*YLDSPD*) is calculated as the difference between the effective yield assuming a 10-year holding period and the then current 10-year Treasury. The data show a general downward drift in the effective yield and yield spreads mirroring overall rate movements from 1995 to 1997.

We begin our analysis of the difference between conforming and nonconforming yield spreads with the simple nonparametric comparison shown simply in Table 3. We show the mean and median effective yield and yield spread for conforming and nonconforming loans assuming a 10-year expected holding period. Consistent with previous empirical studies, conforming loans in our

[11] We also estimated the credit spread models using a credit score of 660 to denote conforming status. Overall, results remained consistent and do not appear to be sensitive to our choice of credit score cutoff.

[12] Unfortunately, our classification scheme does not control for nonborrower credit quality underwriting criteria and thus may lead to some loans that are nonconforming due to property characteristics being incorrectly categorized as conforming. While the number of these cases is probably small, the effect will be to bias down the conforming/nonconforming differential.

Table 3 ■ Mean effective yield and yield spreads by origination year, segmented by conforming/nonconforming and jumbo/non-jumbo status (standard deviations reported in parentheses).

	1995			1996			1997			F-test	
	N	*YLDSPD*	*YIELD_10*	*N*	*YLDSPD*	*YIELD_10*	*N*	*YLDSPD*	*YIELD_10*	*YLDSPD*	*YIELD_10*
Conforming	5,016	1.789	8.066	4,694	1.596	8.124	10,060	1.533	7.722	974.0***	1715.1***
		(0.383)	(0.510)		(0.363)	(0.463)		(0.290)	(0.392)		
Nonconforming	1,584	1.897	8.417	1,336	1.667	8.175	3,489	1.550	7.811	549.3***	1217.6***
		(0.426)	(0.544)		(0.327)	(0.397)		(0.292)	(0.333)		
T-test		8.93***	22.7***		6.88***	4.04***		2.91***	12.99***		
Jumbo	530	1.844	8.236	1,233	1.665	8.164	3,264	1.544	7.801	239.2***	642.8***
		(0.384)	(0.509)		(0.302)	(0.378)		(0.269)	(0.303)		
Non-jumbo	6,070	1.813	8.143	4,797	1.598	8.128	10,285	1.536	7.728	1237.4***	2012.6***
		(0.398)	(0.542)		(0.368)	(0.466)		(0.297)	(0.400)		
T-test		−1.80*	−4.02***		−6.72***	−2.89***		−1.56	−11.10***		

***significant at the 1% level, **significant at the 5% level, *significant at the 10% level.

sample had lower effective yields and yield spreads than nonconforming loans across all years. Table 3 also shows the effective yields and yield spreads for loans above and below the GSE conforming limits in place at the time of origination (*JUMBO* and *NON-JUMBO*). We see that mortgages with loan amounts above the conforming loan limit had slightly higher yield spreads than non-jumbo loans. *T*-tests within years and *F*-tests across years show statistically significant differences for both *YLDSPD* and *YIELD_10*.

Impact of Conforming Loan Status

Base Model

Since our data set differs from the MIRS data used in prior studies, we begin with a simple model following those used by previous researchers (*e.g.*, Ambrose and Buttimer 2004, Cotterman and Pearce 1996, Hendershott and Shilling 1989, McKenzie 2002). Specifically, we estimate the following model of mortgage origination spreads via ordinary least squares (OLS) based on the simplified jumbo versus non-jumbo classification:

$$\ln S_i = \beta_0 + \beta_1 \ln BAL_i + \sum_{k=1}^{3} \beta_{1+k} \ln LTV_{ik} + \beta_5 JUMBO_i + \sum_{k=1}^{3} \beta_{5+k} REG_{ik} + \sum_{k=1}^{11} \beta_{8+k} QTR_{ik} + \varepsilon_i \tag{6}$$

where BAL_i is the loan amount, LTV_i is the mortgage loan-to-value ratio, $JUMBO_i$ is a dummy variable denoting jumbo loan size, REG_i is a set of dummy variables controlling for location (South, Midwest and West, with North the reference category) and QTR_i represents a series of dummy variables controlling for the year and quarter of origination (the fourth quarter of 1997 is the reference). We incorporate the log of *BAL* and *LTV* into the econometric specification to control for nonlinearities in loan size and LTV.

Table 4 reports the estimated parameter coefficients for (6). Consistent with previous studies, the *JUMBO* parameter coefficient is positive and significant, indicating that jumbo mortgages have significantly higher yield spreads than loans below the conforming loan limit. The estimate of 0.142 implies that jumbos have origination spreads about 15% higher than mortgages below the conforming loan limit.[13] Multiplying this percentage difference by the mean

[13] The point estimate is calculated as $e^{\beta} - 1$.

Table 4 ■ OLS regression estimates for the base yield spread model and the adjusted model controlling for "conforming" loan status and borrower credit quality.

		Model 1		Model 2	
Variable	Label	Coeff.	Std. Err.	Coeff.	Std. Err.
INTERCEPT	Intercept	1.828***	0.036	3.468***	0.114
CONFORM	1 = conforming loan			−0.032***	0.006
JUMBO	1 = jumbo loan	0.142***	0.005	0.113***	0.007
Log(*LTV*)	Log of loan-to-value ratio	0.036***	0.005	0.019***	0.005
Log(*FICO*)	Fair Isaac credit score			−0.274***	0.015
CREDITSPD	AAA-BBB bond spread			0.684***	0.080
Log(*ORIBAL*)	Loan amount	−0.118***	0.003	−0.114***	0.003
HPI_STDEV	House price index volatility			0.001*	0.001
SIG_GS1	Interest rate volatility			−0.225***	0.031
YLDCURVE	Yield curve slope			−0.471***	0.017
q_2	Nonjudicial, no deficiency			−0.024***	0.005
q_3	Judicial and deficiency			0.011***	0.003
q_4	Judicial, no deficiency			−0.032*	0.020
QTR95_1	1st quarter 1995	0.088***	0.010	0.444***	0.032
QTR95_2	2nd quarter 1995	0.171***	0.006	0.402***	0.024
QTR95_3	3rd quarter 1995	−0.005	0.005	0.159***	0.016
QTR95_4	4th quarter 1995	0.183***	0.005	0.209***	0.016
QTR96_1	1st quarter 1996	−0.123***	0.007	0.123***	0.019
QTR96_2	2nd quarter 1996	−0.074***	0.006	0.196***	0.017
QTR96_3	3rd quarter 1996	0.007	0.006	0.194***	0.015
QTR96_4	4th quarter 1996	0.097***	0.006	0.232***	0.013
QTR97_1	1st quarter 1997	−0.076***	0.008	0.104***	0.012
QTR97_2	2nd quarter 1997	−0.063***	0.005	0.099***	0.009
QTR97_3	3rd quarter 1997	−0.068***	0.004	0.040***	0.006
SOUTH	1 = Located in South Region	−0.033***	0.004	−0.024***	0.004
MIDWEST	1 = Located in Midwest Region	0.007*	0.004	0.003	0.004
WEST	1 = Located in West Region	0.001	0.003	0.023***	0.005
Adjusted R^2		0.208		0.250	

The dependent variable is the log of the origination yield spread, where the origination yield spread is calculated assuming a 10-year holding period. ***significant at the 1% level, **significant at the 5% level, *significant at the 10% level.

sample yield spread of 162.5 basis points implies that jumbo loans have yield spreads 24.8 basis points higher than non-jumbo loans. This result is completely consistent with other studies. For example, Hendershott and Shilling (1989) found that jumbo loan rates were 24 to 39 basis points higher than conforming loans, and, over the period from 1989 to 1993, Cotterman and Pearce (1996) estimated loan rate differentials between 24 and 60 basis points. Our result is also consistent with McKenzie (2002), who estimates the jumbo/conforming loan rate differential between 11 and 31 basis points over the period from 1986 to 2000, a longer period.

Base Model: Impact of Loan Risk Characteristics

The drawback to the base model of jumbo/non-jumbo spreads is that it does not control for loan risk characteristics and the fact that some conforming size loans may not actually be conforming. As a first step in addressing this problem, Table 4 also reports an extended specification that includes borrower risk characteristics and other factors. We estimate this model over the full sample (1995 to 1997 originations) and include two indicator variables, *CONFORM* and *JUMBO,* to decompose the spread differential. This specification also includes borrower credit score (*FICO*) and loan balance. To capture the time-varying market influences we also include the yield difference between the "AAA" bond index and the "Baa" bond index (*CREDSPD*) as well as a measure of yield curve slope (*YLDCURVE*), proxied by the difference between the 10-year and 1-year Treasury bond rates. Finally, following Ambrose and Pennington-Cross (2000), we include a set of dummy variables that categorize states based on judicial versus nonjudicial foreclosure laws and deficiency versus nondeficiency judgment states to control for differences in legal environment. q_1 indicates states that have nonjudicial foreclosure available and allow deficiency judgments;[14] q_2 indicates states that have nonjudicial foreclosure available and do not allow deficiency judgments;[15] q_3 indicates states that require judicial foreclosure and allow deficiency judgments[16] and finally q_4 indicates states that require judicial foreclosure and do not allow deficiency judgments.[17]

Based on the theoretical model of yield spreads outlined above, we expect yield spreads to be related to house price and interest rate volatility and loan-to-value ratio. We construct a proxy for house price volatility by calculating the standard deviation in the OFHEO state-level house price index over the eight quarters prior to the quarter of origination. We use the standard deviation in the 1-year Treasury bond rate over the 15 months prior to origination as a proxy for rate volatility.

[14] q_1 indicate AL, AR, DC, GA, HI, MO, IA, MA, MD, MI, MS, RI, NE, NH, NM, NV, NY, TN, UT, VA, WV, WY and CO.

[15] q_2 indicate AK, AZ, CA, ID, OK, ME, MN, MT, NC, OR, SD, TX and WA.

[16] q_3 indicate CT, DE, FL, IL, IN, KS, KY, NJ, OH, PA, SC and VT.

[17] q_4 indicate LA, ND and WI.

Results indicate that it is possible to decompose the 15% jumbo–non-jumbo differential into two distinct parts: an exogenous loan size effect and an endogenous credit quality effect reflecting GSE underwriting. The positive and significant coefficient for *JUMBO* measures the size effect: Loans exceeding the GSE conforming loan limit have yield spreads 12.0% above loans with balances below the limit. The negative coefficient for *CONFORM* indicates that loans meeting the conforming guidelines have yield spreads 3.2% lower than nonconforming loans. Again, multiplying these estimates by the sample mean suggests that conforming loans have yield spreads that are 5.1 basis points lower than nonconforming, non-jumbo loans, while jumbo loans have yield spreads that are 19.4 basis points above the nonconforming, non-jumbo segment and 24.5 basis points above conforming loans.

We also note that adding borrower risk characteristics and incorporating market risk factors increases the model's adjusted R^2 from 20.8% to 25%. We now turn to additional econometric issues and further refinement of the model to address them.

Econometric Issues

Endongeneity

To reduce default risk, lenders impose LTV restrictions. However, since LTV is such an important component of underwriting, it can be argued that LTV, mortgage amount and contract rate are jointly determined. Moreover, the endogenous relationship between the level of mortgage debt and the amount of housing consumption is well known. For example, Ling and McGill (1998) develop a simultaneous equation model where mortgage demand is a function of borrower income, nonhousing wealth, desired housing consumption and demographic characteristics, and housing consumption is a function of the level of mortgage debt as well as economic and demographic factors. To control for this endogenous relationship, we specify *LTV* and house value as a simultaneous equations system and utilize the predicted loan-to-value ($\hat{LTV}$) in the subsequent yield spread regression. We estimate the following system via two-stage least squares regression:

$$\begin{aligned}\log(house_i) = {} & \beta_0 + \beta_1 LTV_i + \beta_2 BRWAGE_i + \beta_3 BRWAGE_i^2 + \beta_4 r_{mkt} \\ & + \beta_5 INCOME_i + \beta_6 INCOME_i^2 + \beta_7 YR95 + \beta_8 YR96 \\ & + \sum_{k=2}^{10} \beta_{7+k} REG_{i,k} + \varepsilon_i \end{aligned} \tag{7}$$

$$LTV_i = \alpha_0 + \alpha_1 \log(house_i) + \alpha_2 FICO_i + \alpha_3 r_{mkt} + \alpha_4 INCOME_i + \alpha_5 INCOME_i^2 + \alpha_6 q_2 + \alpha_7 q_3 + \alpha_8 q_4 + \alpha_9 YR95 + \alpha_{10} YR96 + \sum_{k=2}^{10} \alpha_{9+k} REG_{i,k} + \varepsilon_i \quad (8)$$

where $INCOME_i$ is borrower i's income at origination, $BRWAGE_i$ is borrower i's age at origination, r_{mkt} is the current mortgage rate at origination as proxied by the Freddie Mac 30-year fixed-rate mortgage rate, $FICO_i$ is borrower i's credit score at origination,[18] *YR95* and *YR96* are dummy variables denoting year of origination (1997 is the control year), and REG_k ($k = 1 \ldots 4$) are a series of dummy variables denoting the region where the loan is located (Region 1, North, is the control region). Use of borrower age and income together provides a rough proxy for total wealth, or permanent income, and we include the square of these variables to control for any nonlinear effects in wealth or income. We also include the previously described set of dummy variables that classify states based legal environment ($q_1 \ldots q_4$).

Table 5 reports the results of estimating Equations (7) and (8) via two-stage least squares. Consistent with Ling and McGill (1998), we find that loan-to-value ratios are lower for higher value homes and for borrowers with higher credit ratings. We also find that borrower income is nonlinearly related to LTV, with individuals having greater income capable of supporting higher LTV ratios. We also find that LTV ratios are higher in states that do not allow lenders to pursue deficiency judgments (q_2 and q_4). Based on the model coefficients reported in Table 5, we calculate the predicted *LTV* ratio for each borrower, which is used in the enhanced yield spread model.

Table 6 reports our enhanced yield spread model using the predicted LTV values. The coefficient on *CONFORM* remains significantly negative, implying that conforming loans do have lower yield spreads than nonconforming loans. However, correcting for the endogenous relationship for LTV causes the magnitude of the conforming loan differential to shrink. The estimated coefficient in Table 6 implies that conforming loans have yield spreads that are 3.0% lower than nonconforming loans. Again, this implies that conforming loans have yield spreads that are approximately 24.4 basis points smaller than jumbo loans.

[18] In the empirical estimation, we use the log of *FICO*.

Table 5 ■ Estimation of the endogenous relationship between LTV and house values.

Parameter	Variable	Estimate	Approx. Std. Err.	t-stat.	Approx. $Pr > \|t\|$
		LTV equation			
α_1	*Intercept*	1.699	0.036	47.740	<.0001
α_2	Log*(house value)*	−0.071	0.002	−32.850	<.0001
α_3	Log*(FICO)*	−4.3E-04	1.7E-05	−25.850	<.0001
α_4	r_{mkt}	0.029	0.003	9.620	<.0001
α_5	*INCOME*	1.7E-04	1.7E-05	9.870	<.0001
α_6	$INCOME^2$	−1.8E-08	2.4E-09	−7.510	<.0001
α_7	q_2	0.011	0.004	2.890	0.004
α_8	q_3	−0.007	0.002	−2.800	0.005
α_9	q_4	0.029	0.015	2.000	0.045
α_{10}	*YR95*	0.009	0.002	3.800	0.000
α_{11}	*YR96*	−0.009	0.003	−3.510	0.000
α_{12}	*South*	−3.8E-04	0.003	−0.120	0.903
α_{13}	*Midwest*	−0.009	0.003	−3.220	0.001
α_{14}	*West*	0.004	0.004	0.930	0.352
		House equation			
β_1	*Intercept*	12.332	0.081	152.870	<.0001
β_2	*LTV*	−1.164	0.027	−43.370	<.0001
β_3	*BRWAGE*	0.032	0.002	19.800	<.0001
β_4	$BRWAGE^2$	−3.9E-04	1.7E-05	−22.870	<.0001
β_5	r_{mkt}	−0.033	0.009	−3.590	0.000
β_6	*INCOME*	0.004	4.3E-05	97.660	<.0001
β_7	$INCOME^2$	−4.4E-07	6.9E-09	−63.800	<.0001
β_8	*YR95*	−0.179	0.008	−23.760	<.0001
β_9	*YR96*	−0.042	0.008	−5.090	<.0001
β_{10}	*South*	−0.145	0.009	−16.930	<.0001
β_{11}	*Midwest*	−0.170	0.008	−20.620	<.0001
β_{12}	*West*	0.214	0.008	28.200	<.0001

The table reports the Simultaneous Two Stage Least Squares (2SLS) regression estimates of the following system of equations:

$$\begin{aligned} \log(house_i) &= \beta_0 + \beta_1 LTV_i + \beta_2 BRWAGE_i + \beta_3 BRWAGE_i^2 + \beta_4 r_{mkt} \\ &\quad + \beta_5 INCOME_i + \beta_6 INCOME_i^2 + \beta_7 YR95 + \beta_8 YR95 + \varepsilon_i \\ LTV_i &= \alpha_0 + \alpha_1 \log(house_i) + \alpha_2 FICO_i + \alpha_3 r_{mkt} + \alpha_4 INCOME_i + \alpha_5 INCOME_i^2 \\ &\quad + \alpha_6 q2 + \alpha_7 q3 + \alpha_8 q4 + \alpha_9 YR95 + \alpha_1 YR96 + \varepsilon_i \end{aligned}$$

The dependent variables are the house price at origination and the loan-to-value ratio at origination. The independent variables are as follows: $INCOME_i$ is borrower income at origination; $BRWAGE_i$ is borrower age at origination; r_{mkt} is the current mortgage rate at origination as proxied by the Freddie Mac 30-year fixed-rate mortgage rate; $FICO_i$ is borrower credit score at origination; *YR95* and *YR96* are dummy variables denoting year of origination (1997 is the control year); q_1 indicates states that have nonjudicial foreclosure available and allow lenders to obtain deficiency judgments; q_2 indicates states that have nonjudicial foreclosure available but do not allow deficiency judgments; q_3 indicates states that require judicial foreclosure and allow deficiency judgments; and q_4 indicates states that require judicial foreclosure and do not allow deficiency judgments.
Adjusted R^2—LTV equation: 14.2%.
Adjusted R^2—House equation: 45.0%.

Table 6 ■ Adjusted model controlling for conforming, jumbo status and borrower credit.

		Model 3			
Variable	Label	Coeff.	Std. Err.	*T*-stat.	*P*-value
INTERCEPT	Intercept	2.217	0.194	11.437	0.000
CONFORM	1 = conforming loan	−0.030	0.006	−5.411	0.000
JUMBO	1 = jumbo loan	0.114	0.007	16.459	0.000
Log(*LTV*)	Log of loan-to-value ratio	0.373	0.046	8.120	0.000
Log(*FICO*)	Fair Isaac credit score	−0.119	0.025	−4.719	0.000
CREDITSPD	AAA-BBB bond spread	0.694	0.080	8.727	0.000
Log(*ORIBAL*)	Loan amount	−0.085	0.004	−19.560	0.000
HPI_STDEV	House price index volatility	0.001	0.001	1.670	0.095
SIG_GS1	Interest rate volatility	−0.235	0.031	−7.615	0.000
YLDCURVE	Yield curve slope	−0.489	0.018	−27.880	0.000
Q_2	Nonjudicial, no deficiency	−0.029	0.005	−5.744	0.000
Q_3	Judicial and deficiency	0.015	0.003	4.474	0.000
Q_4	Judicial, no deficiency	−0.047	0.020	−2.408	0.016
QTR95_1	1st quarter 1995	0.436	0.032	13.442	0.000
QTR95_2	2nd quarter 1995	0.398	0.024	16.814	0.000
QTR95_3	3rd quarter 1995	0.155	0.016	9.881	0.000
QTR95_4	4th quarter 1995	0.205	0.016	13.189	0.000
QTR96_1	1st quarter 1996	0.136	0.019	7.123	0.000
QTR96_2	2nd quarter 1996	0.201	0.017	11.645	0.000
QTR96_3	3rd quarter 1996	0.195	0.015	12.668	0.000
QTR96_4	4th quarter 1996	0.237	0.013	17.840	0.000
QTR97_1	1st quarter 1997	0.103	0.012	8.550	0.000
QTR97_2	2nd quarter 1997	0.096	0.009	10.571	0.000
QTR97_3	3rd quarter 1997	0.041	0.006	6.554	0.000
SOUTH	1 = Located in South Region	−0.025	0.004	−6.206	0.000
MIDWEST	1 = Located in Midwest Region	0.007	0.004	1.684	0.092
WEST	1 = Located in West Region	0.022	0.005	4.298	0.000
Adjusted R^2		0.251			

Estimated with predicted LTVs.

Sample Selection Bias

As discussed previously, raw data suggests that conforming loan rates are consistently lower than nonconforming rates. But such a pattern could emerge for at least two distinct reasons. First, the differential could reflect a liquidity premium, since it is less costly to sell (or swap) loans to the GSEs, compared to the creating complex private label mortgage-backed securities. Second, the differential could be attributable to credit risk, that is, if the GSEs purchase only loans with lower credit risk, then any comparison to rates for higher risk loans would be confounded by this difference.

To control for this potential sample selection bias, we estimate the "treatment effects" model (see Green 1997). This procedure involves estimating the following conforming loan selection model:

$$CONFORM_i = \gamma Z_i + \xi_i, \tag{9}$$

where Z_i is a vector of characteristics that determines whether the loan is a conforming loan and ξ_i is an error term. Since the GSEs have congressionally mandated missions to increase the supply of mortgage credit for low- and moderate-income borrowers, we assume that Z_i includes borrower income (*INCOME*) and age (*BRWAGE*, proxying for borrower wealth) along with dummy variables for year to reflect the yearly adjustments in the conforming loan limit. We estimate (9) as a probit model with the following form:

$$\Pr(CONFORM_i = 1) = \frac{\phi(-\gamma Z_i)}{1 - \Phi(-\gamma Z_i)} \tag{10}$$

and

$$\Pr(CONFORM_i = 0) = [1 - \Pr(CONFORM_i = 1)], \tag{11}$$

where ϕ is the standard normal probability density function (pdf) and Φ is the standard normal cumulative distribution function (cdf). From the probit model coefficients, γ, we compute the inverse Mills ratio (λ_i) as

$$\lambda_i = \phi(\hat{\gamma} Z_i)/\Phi(\hat{\gamma} Z_i). \tag{12}$$

Flannery and Houston (1999) discuss that if ε_i and ξ_i are jointly normally distributed, then

$$E(\varepsilon_i \mid CONFORM_i) = \rho\sigma_\xi E(\xi_i \mid CONFORM_i), \tag{13}$$

where ρ is the correlation between ε_i and ξ_i, and σ_ξ is the standard deviation of ξ_i. Then in the second step, we reestimate Equation (5) via least squares with λ_i included as an explanatory variable:[19]

$$\ln S_i = \alpha_0 + \alpha_1\sigma_{H_i} + \alpha_2\sigma_{r_i} + \alpha_3\left(V_i^M/H_i\right) + \alpha_4 CONFORM_i + \beta X_i + \sum_{j=1}^{11} \delta_j QTR_{ji} + \varphi\lambda_i + \varepsilon_i \tag{14}$$

[19] The two-step estimation procedure follows Heckman (1979).

Table 7 ■ Sample selection correction.

Panel A: Probit model[a]				
Parameter	Estimate	Error	Chi Square	*Pr* > ChiSq
INTERCEPT	2.451	0.080	946.740	<.0001
BRWAGE	0.005	0.001	32.370	<.0001
INCOME	−0.008	0.000	3,308.890	<.0001
Log(*LTV*)	−1.727	0.071	591.960	<.0001
YR95	−0.069	0.023	9.190	0.002
YR96	0.088	0.023	14.380	0.000

The inverse Mills ratio coefficient (φ) is a measure of ($\rho\sigma_\xi$) in (14); hence, if the parameter estimate of λ is zero, sample selection bias is not present. However, Willis and Rosen (1979) show that including λ corrects for selectivity bias in the sample observations. Furthermore, if the λ coefficient is statistically significant, then this suggests that the mortgage status as conforming impacts the mortgage spread at origination. Finally, following Flannery and Houston (1999), a positive (negative) φ implies that $\rho > (<) 0$, which suggests that the origination spread on conforming loans is higher (lower) than nonconforming loans.

Table 7, Panel A presents the results of the first-stage probit model (Equation (9)). Results show that the probability of a loan being conforming is positively related to borrower age and negatively related to borrower income. This result seems intuitive since older households would have accumulated greater funds for down payment and higher income households would be more likely to purchase higher priced housing requiring jumbo loans. In addition, the probability of being conforming declines as the loan-to-value ratio increases, consistent with underwriting of credit risk.

Table 7, Panel B presents the second-stage results with asymptotically corrected standard errors for the yield spread model that includes the inverse-Mills ratio to control for sample selection bias. The significantly positive coefficient for λ indicates that the loan spread is indeed affected by conforming loan status, from which we may conclude that a simple OLS model of *SPREAD* without including λ would suffer from omitted variables bias.[20] In addition, since λ is positive, this implies that conforming loan origination spreads are actually larger than would be estimated under the simple OLS regression. In other words, the base OLS coefficients presented in Tables 4 and 6 are biased.

[20] See Flannery and Houston (1999) for a discussion of the interpretation of the inverse-Mills ratio in the context of the impact of bank examinations on market value.

Table 7 ■ continued.

Variable	Label	Parameter	Std. Error	T-stat.	P-value
INTERCEPT	Intercept	3.112	0.229	13.584	0.000
CONFORM	1 = conforming loan	−0.096	0.011	−9.143	0.000
JUMBO	1 = jumbo Loan	0.107	0.007	15.443	0.000
Log(*FICO*)	Fair Isaac credit score	−0.215	0.029	−7.555	0.000
Log(*LTV*)	Log of loan-to-value ratio	0.682	0.079	8.591	0.000
CREDITSPD	AAA-BBB bond spread	−0.109	0.006	−19.995	0.000
Log(*ORIBAL*)	Loan amount	0.134	0.057	2.366	0.018
HPI_STDEV	House price index volatility	0.001	0.001	1.581	0.114
SIG_GS1	Interest rate volatility	−0.230	0.031	−7.454	0.000
YLDCURVE	Yield curve slope	−0.478	0.018	−27.133	0.000
q_2	Nonjudicial, no deficiency	−0.026	0.005	−5.180	0.000
q_3	Judicial and deficiency	0.013	0.003	3.785	0.000
q_4	Judicial, no deficiency	−0.039	0.020	−1.971	0.049
QTR95_1	1st quarter 1995	0.441	0.032	13.623	0.000
QTR95_2	2nd quarter 1995	0.400	0.024	16.941	0.000
QTR95_3	3rd quarter 1995	0.158	0.016	10.055	0.000
QTR95_4	4th quarter 1995	0.207	0.016	13.364	0.000
QTR96_1	1st quarter 1996	0.130	0.019	6.776	0.000
QTR96_2	2nd quarter 1996	0.201	0.017	11.619	0.000
QTR96_3	3rd quarter 1996	0.197	0.015	12.814	0.000
QTR96_4	4th quarter 1996	0.237	0.013	17.840	0.000
QTR97_1	1st quarter 1997	0.103	0.012	8.563	0.000
QTR97_2	2nd quarter 1997	0.098	0.009	10.760	0.000
QTR97_3	3rd quarter 1997	0.040	0.006	6.459	0.000
SOUTH	1 = Located in South Region	−0.025	0.004	−6.212	0.000
MIDWEST	1 = Located in Midwest Region	0.004	0.004	1.046	0.295
WEST	1 = Located in West Region	0.025	0.005	4.711	0.000
λ	Inverse Mill's ratio	0.039	0.005	7.381	0.000

Panel B: Second stage OLS regression with consistent asymptotic standard errors[b]

[a]This panel reports the maximum-likelihood parameter estimates for the first-stage probit model of whether a loan is a conforming loan. The dependent variable is a dummy variable equal to 1 if the loan is conforming and 0 otherwise. The independent variables are as follows: *INCOME* is borrower income at origination; *BRWAGE* is borrower age at origination; *LTV* is the actual loan-to-value ratio at origination; and *YR95* and *YR96* are dummy variables denoting year of origination (1997 is the control year).
Log-likelihood: −12258.5.

[b]This panel reports the OLS regression estimates of the following equation:

$$\ln S_i = \alpha_0 + \alpha_1\sigma_{H_i} + \alpha_2\sigma_{r_i} + \alpha_3\left(V_i^M/H_i\right) + \alpha_1 CONFORM + \beta X_i + \sum_{j=1}^{11}\delta_i QTR + \varphi\lambda_i + \varepsilon_i$$

The dependent variable is the log of the yield spread calculated as the difference between the effective loan yield assuming a 10-year holding period and the 10-year Treasury rate. The inverse Mills ratio (λ) is defined as $\lambda_i = \phi(\hat{\gamma}Z_i)/\Phi(\hat{\gamma}Z_i)'$, where γ are the parameter coefficients of the first-stage probit model reported in Panel A.

Other results are consistent with our expectations about the relationship between yield spreads and loan and borrower characteristics. Yield spreads are positively related to LTV and negatively related to borrower credit score. Furthermore, yield spreads are positively related to rate volatility.

Turning to the general macro-economic variables, we find that mortgage yield spreads are negatively associated with changes in the general default risk premium and negatively related to changes in the yield curve. The negative coefficient for the bond market yield spread (*CREDSPD*) indicates that a 1 basis point increase in the corporate bond market yield spread results in a 1.2 basis point decrease in mortgage yield spreads (assuming a conforming loan with an LTV ratio of 80%). Likewise, a 1 basis point increase in the yield curve translates into approximately a 0.7 basis point decline in the mortgage yield spread. The legal environment dummy variables indicate that loans originated in states that do not allow deficiency judgments have lower origination spreads. Given that these states have more borrower-friendly laws, it is rather surprising that the negative coefficient suggests that loans originated in these states actually have lower yield spreads than loans originated in other more lender-friendly states.

Finally, the parameter estimate for the conforming loan dummy variable continues to be negative and statistically significant while the estimate for the jumbo size dummy is positive and significant. The *JUMBO* coefficient indicates that mortgages with loan amounts above the GSE conforming loan limit have yield spreads that are on average approximately 11.3% higher than mortgages with loan amounts below the conforming loan limit. This implies that mortgages above the conforming loan limit have yield spreads that are 18.4 basis points above loans that are below the conforming loan limit (evaluated at the average yield spread of 162.5 basis points). Following Greene (1997), the impact of conforming loan status is given by $\alpha_4 + \varphi\lambda$. Thus, we estimate the impact of conforming loan status as -0.0554, indicating that conforming loan yield spreads are 5.5% lower than nonconforming loans.[21] This implies that conforming loans have yield spreads that are 9 basis points lower than nonconforming loans (evaluated at the average yield spread for the sample). Thus, combining the jumbo and conforming loan effects suggests that conforming loans have yield spreads that are 27.7 basis points lower than jumbo loans (evaluated at the average yield spread of 162.5 basis points for the sample).[22] Comparing this estimate with the 24.5 basis point conforming-jumbo differential calculated using the coefficients in Tables 4 and 6, we see that adjusting for sample selection bias increases the magnitude of the estimate, though not dramatically.

[21] $-0.0554 = \exp(-0.096 + 0.039) - 1$.

[22] The total differential is found by subtracting the jumbo effect from the conforming effect: 27.7 bp = 18.4 bp − (−9 bp).

Our decomposition of the yield spread differential into the loan size effect and an underwriting effect has implications for the debate surrounding the magnitude of the interest savings the GSEs provide to consumers. In a simulation of the effect of house price volatility on the yield spread differential, Ambrose, Buttimer and Thibodeau (2001) found that 20% of the jumbo/non-jumbo yield differential could be explained by differences in underlying collateral property price dynamics.[23] Thus, applying this simulation result to our analysis implies that 3.68 basis points of the jumbo/non-jumbo effect may result from differences in the volatility of typical properties that collateralize loans above and below the conforming loan limit.[24] To summarize, of the 27.7 basis point differential, 32% (9 basis points) results from the GSE conforming loan underwriting guidelines, 13% (3.68 basis points) results from differences in property price volatility and 53% (14.72 basis points) results from the conforming loan limit barrier. This implies that the volatility-adjusted yield spread is approximately 24 basis points.

A reasonable interpretation of the conforming loan underwriting differential is that this is the source for the pass-through of the GSE benefits associated with their charters. As a result of the GSEs' special relationship with the Federal government, Ambrose and Warga (2002) indicate that the GSEs enjoy a debt funding advantage between 25 and 29 basis points over comparable "AA" banking sector firms and between 43 and 47 basis points over comparable "A" banking sector firms.[25] Combining these funding advantage estimates with our result of the conforming loan differential implies that the GSEs retain between 5 and 50% of their debt cost advantage.[26] In other words, our analysis suggests that the GSEs pass through between 50 and 95% of their debt funding advantage to borrowers in the conforming loan market in the form of lower interest rates.

Conclusions

This paper focuses on the rate reduction associated with conforming loan status. Our analysis goes beyond the traditional jumbo/non-jumbo classification and

[23] The Ambrose, Buttimer and Thibodeau (2001) simulation was based on properties located in Dallas, Texas, and thus may not reflect the national market.

[24] 3.68 bp $= 0.2 \times 18.4$ bp.

[25] The Ambrose and Warga (2002) analysis covers the period between 1995 and 1999—roughly consistent with the mortgage origination data employed here.

[26] At the low end, the 5% is calculated assuming the GSE funding advantage is 25 basis points and the volatility adjusted yield spread differential is 24 bp ([25 bp − 24 bp]/25 bp). At the high end, the 50% is calculated assuming the GSE funding advantage is 47 basis points ([47 bp − 24 bp]/47 bp).

includes borrower credit score, a key piece of information that allows us to estimate the reduction in mortgage yield spreads attributable to conforming loan status on a risk-adjusted basis.

Our results confirm that mortgage yield spreads are positively related to loan-to-value ratio and negatively related to borrower credit score. This is consistent with the finance literature, which relates yield spreads to firm capital structure and credit ratings. We also find that conforming loans have lower yield spreads, after controlling for borrower and loan level risk characteristics and the broader bond market environment, though our point estimates are smaller than those found in previous studies. Correcting for endogeneity and sample selection bias shrinks the magnitude of the conforming loan yield spread advantage still further.

Finally, we utilize our decomposition of the mortgage yield differential into the component parts to identify the magnitude of the benefit the GSEs provide to the market via the pass-through of their debt funding advantage. Depending on the choice of benchmark comparison, we estimate that the GSEs pass through between 50 and 95% of the debt funding advantage they enjoy over comparable financial institutions as a result of the implicit guarantee arising from their Congressional charters.

We thank Jim Follain, Bob Van Order and the seminar participants at the Homer Hoyt Institute, Rice University, Cambridge University and the University of Wisconsin at Madison for their helpful comments and suggestions on earlier versions of this paper. Any errors are our own responsibilities.

References

Altman, E.I. and A. Saunders. 1998. Credit Risk Measurement: Developments Over the Last 20 Years. *Journal of Banking and Finance* 21: 1721–1742.

Ambrose, B.W. and R.J. Buttimer, Jr. 2004. GSE Impact on Rural Mortgage Markets. *Regional Science and Urban Economics*. Forthcoming.

Ambrose, B.W., R.J. Buttimer, Jr. and T. Thibodeau. 2001. A New Spin on the Jumbo/Conforming Loan Rate Differential. *Journal of Real Estate Finance and Economics* 23(3): 309–335.

Ambrose, B.W. and A. Pennington-Cross. 2000. Local Economic Risk Factors and the Primary and Secondary Mortgage Markets. *Regional Science and Urban Economics* 30(6): 683–701.

Ambrose, B. and A. Warga. 1995. Pricing Effects in Fannie Mae Agency Bonds. *Journal of Real Estate Finance and Economics* 11: 235–249.

———. 2002. Measuring Potential GSE Funding Advantages. *Journal of Real Estate Finance and Economics* 25(2/3): 129–151.

Angbazo, L., J. Mei and A. Saunders. 1998. Yield Spreads in the Market for Highly Leveraged Transaction Loans. *Journal of Banking and Finance* 22: 1249–1282.

Avery, R., R. Bostic, P. Calem and G. Canner. 1996. Credit Risk, Credit Scoring, and the Performance of Home Mortgages. *Federal Reserve Bulletin* 82: 621–648.

Bakshi, G., D. Madan and F. Zhang. 2000. What Drives Default Risk? Lessons from Empirically Evaluating Credit Risk Models. Mimeo. College Park: University of Maryland.

Campbell, T. and J.K. Dietrich. 1983. The Determinants of Default on Conventional Residential Mortgages. *Journal of Finance* 48(5): 1569–1581.

Collin-Dufresne, P., R.S. Goldstein and J.S. Martin. 2001. The Determinants of Credit Spread Changes. *Journal of Finance* 56(6): 2177–2208.

Cotterman, R.F. and J.E. Pearce. 1996. The Effects of the Federal National Mortgage Association and the Federal Home Loan Mortgage Corporation on Conventional Fixed-Rate Mortgage Yields. In *Studies on Privatizing Fannie Mae and Freddie Mac*. Washington, DC: U.S. Department of Housing and Urban Development.

Cox, J.C., J.E. Ingersoll, Jr. and S.A. Ross. 1985. An Intertemporal General Equilibrium Model of Asset Prices. *Econometrica* 53(2): 363–384.

Cunningham, D. and C. Capone. 1990. The Relative Termination Experience of Adjustable to Fixed-Rate Mortgages. *Journal of Finance* 45(5): 1687–1703.

Deng, Y., J.M. Quigley and R. Van Order. 2000. Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options. *Econometrica* 68(2): 275–307.

Duffie, G. 1998. The Relation Between Treasury Yields and Corporate Bond Yield Spreads. *Journal of Finance* 53: 2225–2242.

———. 1999. Estimating the Price of Default Risk. *Review of Financial Studies* 12(1): 197–226.

Duffie, D. and K. Singleton. 1999. Modeling Term Structures of Default Risky Bonds. *Review of Financial Studies* 12(4): 687–720.

Ericsson, J. and O. Renault. 2000. Liquidity and Credit Risk. Mimeo. Montreal: McGill University.

Flannery, M.J. and J.F. Houston. 1999. The Value of a Government Monitor for US Banking Firms. *Journal of Money, Credit and Banking* 31(1): 14–34.

González-Rivera, G. 2001. Linkages Between Secondary and Primary Markets for Mortgages. *Journal of Fixed Income* 11(1): 29–36.

Greene, W.H. 1997. *Econometric Analysis*. 3rd ed. Upper Saddle River, NJ: Prentice Hall.

Heckman, J.J. 1979. Sample Selection Bias as a Specification Error. *Econometrica* 47: 153–161.

Hendershott, P.H. and J.D. Shilling. 1989. The Impact of Agencies on Conventional Fixed-Rate Mortgage Yields. *Journal of Real Estate Finance and Economics* 2: 101–115.

ICF, Incorporated. 1990. 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. Washington, DC.

Kau, J.B. and D.C. Keenan. 1995. An Overview of the Option-Theoretic Pricing of Mortgages. *Journal of Housing Research* 6(2): 217–244.

Kau, J.B., D.C. Keenan and T. Kim. 1993. Transaction Costs, Suboptimal Termination, and Default Probabilities. *Journal of the American Real Estate and Urban Economics Association* 21(3): 247–264.

———. 1994. Default Probabilities for Mortgages. *Journal of Urban Economics* 35: 278–296.

Kau, J.B., D.C. Keenan, W.J. Muller, III and J.F. Epperson. 1992. A Generalized Valuation Model for Fixed-Rate Residential Mortgages. *Journal of Money, Credit, and Banking* 24: 279–299.

———. 1993. Option Theory and Floating-Rate Securities with a Comparison of Adjustable- and Fixed-Rate Mortgages. *Journal of Business* 66(4): 595–618.

Kolari, J.W., D.R. Fraser and A. Anari. 1998. The Effects of Securitization on Mortgage Market Yields: A Cointegration Analysis. *Real Estate Economics* 26(4): 677–693.

Ling, D.C. and G.A. McGill. 1998. Evidence on the Demand for Mortgage Debt by Owner-Occupants. *Journal of Urban Economics* 44: 391–414.

McKenzie, J.A. 2002. A Reconsideration of the Jumbo/Non-Jumbo Mortgage Rate Differential. *Journal of Real Estate Finance and Economics* 25(2/3): 197–214.

Merton, R.C. 1974. On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *Journal of Finance* 29(2): 449–470.

Naranjo, A. and A. Toevs. 2002. 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* 25(2/3): 173–196.

Nothaft, F.E., J.E. Pearce and S. Stevanovic. 2002. Debt Spreads Between GSEs and Other Corporations. *Journal of Real Estate Finance and Economics* 25(2/3): 151–172.

Passmore, W., R. Sparks and J. Ingpen. 2002. GSEs, Mortgage Rates, and the Long-Run Effects of Mortgage Securitization. *Journal of Real Estate Finance and Economics* 25(2/3): 215–243.

Phillips, S.M. 1996. The Place of Securitization in the Financial System: Implications for Banking and Financial Policy. Leon T. Kendall and Michael J. Fishman, editors. *A Primer on Securitization*. Cambridge: MIT Press.

Roll, R. 2000. Benefits to Homeowners from Mortgage Portfolios Retained by Fannie Mae and Freddie Mac. Working paper. University of California at Los Angeles.

Quercia, R.G. and M.A. Stegman. 1992. Residential Mortgage Default: A Review of the Literature. *Journal of Housing Research* 3(2): 341–376.

Sanders, A.B. 2002. Government Sponsored Agencies: Do the Benefits Outweigh the Costs? *Journal of Real Estate Finance and Economics* 25(2/3): 121–129.

Todd, S. 2001. The Effects of Securitization on Consumer Mortgage Costs. *Real Estate Economics* 29(1): 29–54.

U.S. Congressional Budget Office. 2001. Interest Rate Differentials Between Jumbo and Conforming Mortgages. Washington, DC: U.S. Congressional Budget Office.

Vandell, K. 1993. Handing Over the Keys: A Perspective on Mortgage Default Research. *AREUEA Journal* 21(3): 211–246.

Von Furstenberg, G. 1969. Default Risk on FHA-Insured Home Mortgages as a Function of the Term of Financing: A Quantitative Analysis. *Journal of Finance* 24: 459–477.

Von Furstenberg, G. and R.J. Green. 1974. Home Mortgage Delinquency: A Cohort Analysis. *Journal of Finance* 29: 1545–1548.

Willis, R. and S. Rosen. 1979. Education and Self-Selection. *Journal of Political Economy* 87(5, part 2): 507–536.