Mortgage Brokers and Mortgage Rate Spreads: Their Pricing Influence Depends on Neighborhood Type

By

M. Cary Collins*
The University of Tennessee
Knoxville, TN 37996-0540

and

Keith D. Harvey
Boise State University
Boise, ID 83725-1630

PRELIMINARY DRAFT FOR PRESENTATION AT THE FMA MEETINGS 2007, ORLANDO, FLORIDA
NOT FOR QUOTATION OR REDISTRIBUTION

Abstract

Recent changes to The Home Mortgage Disclosure Act ("HMDA") make it possible for public interest groups, academics, and industry to review pricing on higher cost mortgage loans. Public interest groups believe mortgage brokers, as indirect or third-party lenders, are largely responsible for most of these reported high interest rate loans. Mortgage brokers have an incentive to solicit applications that are both creditworthy and where there is an information imbalance in favor of the mortgage broker. Empirical evidence on the price effects of mortgage broker participation in the loan process is limited, although recent findings indicate the APR on broker-originated loans is higher than that of lender-originated loan. We examine whether the existence, geographic density, and competition among mortgage brokers is highly correlated with the fraction of reported loans and the pricing of those credits regarded as high interest rate loans. Combining several lending & demographic sources, we create four clusters of substantially-similar U.S. counties & territories, representing more than 75 percent of the U.S. population and more than 99.8 percent of all 1-to-4 family refinancings of 1st and 2nd liens during 2005. Clustering these geographies based on credit, demographic, and structural differences proves key to evaluating the connection between reportable rate spreads & the presence of mortgage brokers in a given geography. In rural geographies, the higher the number of mortgage brokers, the higher the average reported rate spread, revealing a persistent economic rent for the broker. In urban geographies, higher numbers of mortgage brokers indicate greater competition among mortgage brokers to the benefit of borrowers.

---

* *Contact author*: M. Cary Collins, Ph.D., Department of Finance, 421 Stokely Management Center, College of Business Administration, The University of Tennessee, Knoxville, TN 37996-0540, US 1.865.974.1715, E-mail: Cary.Collins@Tennessee.Edu .

Mortgage Brokers and Mortgage Rate Spreads: Their Pricing Influence Depends on Neighborhood Type

Recent changes to The Home Mortgage Disclosure Act ("HMDA") make it possible for public interest groups, academics, and industry to review pricing on higher cost mortgage loans. For example, lenders are now required to report the interest rate spread over the comparable U.S. Treasury security for all first liens where the spread over the comparable Treasury is greater than 300 basis points. Lenders must also report the interest rate spread over the comparable Treasury for second liens where the difference between the two rates is greater than 500 basis points.[1] While these data represent one end of the pricing spectrum, they are also the most controversial. That is, these high interest rate loans are most often described by public interest groups as potentially predatory and increasingly associated with third-party or indirect mortgage lending. Public interest groups believe mortgage brokers, as indirect or third-party lenders, are largely responsible for most of these reported high interest rate loans.

Mortgage brokers act as independent information agents, collecting mortgage applications from individuals and determining the most suitable lender given the individual's creditworthiness. Mortgage brokers typically align with one or more financial services firms to provide mortgage applications to the lender. The broker is compensated in two ways for his/her services. First, the mortgage broker is paid a fee for providing creditworthy mortgage applications to the lender. This fee is typically flat or scaled to the size of the loan. Second, the mortgage broker is paid a "reserve" or spread based on the lending contract interest rate when the

[1] See 12 CFR Part 203, for the implementation of The Home Mortgage Disclosure Act (Regulation C) reporting, and changes to Section 203.5 for the implementation of these new rate spread reporting guidelines (Regulation C; Docket No. R–1219).

interest rate is higher than the buy rate from the lender. In this case, the mortgage broker earns a return when he/she generates a lending contract that is both creditworthy and at a rate higher than the required yield from the lender. The fee is based on projected interest income on the loan, meaning the profit involves both the rate and the size of the loan.

All other things equal, mortgage brokers have an incentive to solicit applications that are both creditworthy and where there is an information imbalance in favor of the mortgage broker. That is, mortgage brokers benefit from situations where the borrower does not understand his/her creditworthiness as well as the mortgage broker. Incomplete information on the part of the borrower enables the mortgage broker to earn a higher interest rate on the credit, meaning a higher fee.

Mortgage brokers also benefit when there are few alternatives to the use of indirect, or third party, lending. In geographies where other low-cost mortgage credit outlets are few, or where transportation/travel costs are significant, mortgage brokers may more easily earn fees for their services. With the advent of mortgage lending via the internet, a mortgage broker's advantage is most pronounced in low-to-moderate income geographies and where there are fewer walk-up lending alternatives.

We examine whether the existence, geographic density, and competition among mortgage brokers is highly correlated with the fraction of reported loans regarded as high interest rate loans, *ceteris paribus*. Holding other factors constant and accepting that mortgage brokers benefit more when there are fewer alternatives or where the costs of the alternatives are higher, we expect: (a) more high rate loans in areas with mortgage brokers – where alternatives are few or when transaction costs higher; and (b) a decrease in the rate spreads when there is a high density of mortgage brokers – as competition for scarce, high-quality borrowers increases.

II. The Relevant Literature

Mortgage brokers have come under persistent criticism from the press and consumer interest groups for engaging in predatory lending practices that take advantage of vulnerable borrowers, including the low income, elderly and minority segments of the market.[2] Reported abuses include: 1) steering borrowers to high cost loans for which the broker is compensated by the lender in the form of a Yield Spread Premium (YSP)[3]; 2) increasing loan volume by encouraging uneconomic refinancing of existing mortgages, or including balloon payments in newly originated loans that require near-term refinancing; 3) receiving compensation from lenders for convincing borrowers to accept large prepayment penalties that trap borrowers in high cost loans after their credit improves; 4) making inadequate disclosure of the true cost of a loan, including the broker's compensation; and 5) originating loans the borrower cannot afford, since the broker does not bear the cost of default.

Despite these criticisms, recent Mortgage Bankers Association Mortgage Origination Surveys estimate that approximately 50 percent of all mortgage loans and 71 percent of subprime mortgage loans in the second half of 2005 were originated through mortgage brokers, suggesting

[2] For some recent examples, see Tedeschi, Bob, "Tighter Reins for New York Brokers," *New York Times*, Oct 22, 2006; Samuels, Adrienne P., "Scrutiny Urged on Loan Officers," *Boston Globe*, Aug 28, 2006; Peterson, Jonathan, "Bonuses to Loan Brokers Scrutinized; Lenders often reward those who arrange high-interest mortgages. Many borrowers don't know they qualify for a lower rate, officials say," *Los Angeles Times*, Mar 27, 2006.

[3] The legitimate role of a YSP is to allow a cash deficient borrower with a high loan-to-value ratio to finance settlement costs over the life of a mortgage in the form of a higher interest rate on the loan. Brokers have been criticized for: 1) charging YSPs that result in an interest rate that is higher than the lender is willing to accept, resulting in kickbacks to the broker; 2) not disclosing that the YSP is built into the rate, and that the broker may receive a portion of the premium as compensation from the lender; 3) not offering borrowers the option of paying settlement cost in cash; and 4) including YSPs in loans for borrowers who could easily afford to pay settlement costs at initiation. See Hearing on "Predatory Mortgage Lending Practices: Abusive Uses of Yield Spread Premia," before the Senate Committee on Banking, Housing, and Urban Affairs, 107th Cong., 2d Sess. 3 (2002), available at banking.senate.gov/02_01hrg/010802/jackson.htm.

that brokers play a valuable role as an intermediary in the mortgage origination process.[4] Since brokers routinely deal with several lenders and have expertise in selecting loan products, they may reduce borrower search costs and make it more likely that the borrower will obtain a product that best serves their needs than if they attempted to shop for a loan on their own. As marketing specialists, brokers may also allow lenders to realize economies in the search process by matching them with customers at a cost that is lower than the lender could achieve on its own.

Empirical evidence on the price effects of mortgage broker participation in the loan process is limited, however, El Anshasy, Elliehausen and Shimazaki (2005) provides a direct comparison of the APR on broker-originated loans with that of lender-originated loans and finds that the former are less costly. The study employs data from the American Financial Services Association's (AFSA) subprime mortgage database. The data are drawn from ten subprime mortgage subsidiaries of AFSA-member companies, and include all closed-end first and second mortgages originated or purchased between the third quarter of 1995 and the first quarter of 2002. The resulting sample includes more than 1 million loans. An ordinary least squares regression model is estimated where the dependent variable is the APR on the loan and explanatory variables include a broker-originated dummy variable along with variables to control for various economic characteristics that influence the rate lenders charge on loans.

The results indicate the APR on broker-originated first mortgages is 1.13 percentage points lower than that on lender-originated loans. For second mortgages this difference is 1.98 percent. The study examines separately markets that might be especially prone to abuses. These include geographic areas that are at least 75 percent African-American or 75 percent Latino, low

[4] "Residential Mortgage Origination Channels," MBA Research Data Notes, Mortgage Bankers Association, Washington, D.C., 2006, available at www.mortgagebankers.org/files/Bulletin/InternalResource/44664_September2006-ResidentialMortgageOriginationChannels.pdf.

income geographies, and geographies with broker or lender licensing requirements that might inhibit competition. The findings suggest brokers provide significant cost advantages in all of these vulnerable segments.

McCoy (2005) criticizes the El Anshasy et al. results by pointing out that 32 percent of the borrowers in the second mortgage sample have credit scores that would qualify them for prime loans, suggesting evidence of steering these borrowers to higher cost subprime loans. In addition, some of the regression coefficients have unexpected signs, suggesting potential for collinearity or omitted variables bias. The loan-to-value ratio (LTV) has a negative correlation with APR in the 70 to 90 percent LTV range, contrary to the expectation that higher LTV exposes the lender to greater risk. El Anshasy et al. explain that the borrowers in this LTV range are predominantly high income applicants with high credit scores. High credit risk borrowers are not approved for high LTV loans, so a high LTV signifies low credit risk, a result consistent with Calomiris (1998). McCoy counters that data from Standard and Poor's show the opposite effect, with subprime borrowers exhibiting LTV ratios that are on average 10 points higher than those for prime borrowers.

Addressing the coefficients for borrower income in the El Anshasy et al. results, McCoy notes that as income falls from $100,000 to $35,000, so do APRs, inconsistent with the expectation that high income borrowers pay lower rates. El Anshasy et al. suggest that the income effect may reflect the other variables that are related to income, since high income borrowers in the subprime market have other characteristics that exclude them from the prime market, such as the inability to document income or unstable income. To address lender selection bias, El Anshasy et al. employ an instrumental variables approach. McCoy points out

that the results from this approach show that the probability of obtaining a loan from a broker is lower across all ethnic and racial groups, including white, black, Asian and Hispanic applicants, and suggests possible misspecification as a cause. Finally, the results show a negative relationship between loan term and APR, counter to the historical relationship of higher rates for longer term loans.

Collins (2005) argues that the lenders contributing to the AFSA data are among the most proactive lending institutions and are likely to exercise greater care when selecting brokers as compared with other more typical lenders. Thus, the El Anshasy et al. data may not be representative of the broker universe. He recommends the need for additional research on the pricing topic using a more comprehensive data set.

Backley, Niblack, Pahl, Risbey and Vockrodt (2006) analyze the impact of broker licensing policy in Minnesota and Wisconsin. Both states adopted licensing in 1988, but Wisconsin licenses both firms and individuals, while Minnesota licenses only firms. The data show a lower density of brokers in Wisconsin, with .32 brokers per 1,000 owner-occupied housing units compared to .52 brokers in Minnesota, suggesting that licensing is a barrier to entry. Regarding access to credit, Wisconsin has similar numbers of mortgage applications, including subprime applications, suggesting that lower broker density does not reduce overall credit availability. Minorities experience significantly higher denial rates in Wisconsin, however, possibly indicating that the minority population is adversely affected by licensing's impact on broker density.

Backley et al. also examine foreclosure rates in the two states. Both states experienced reductions in foreclosures after adopting licensing. This result is expected since licensing should

discourage the practice of pursuing loans where the broker knows the odds of default are high, but the broker does not bear the cost. Examining the difference between Wisconsin's higher foreclosure rates and Minnesota's rates, the authors find a reduction in the gap during the period after both states adopted licensing, suggesting that licensing provides benefits by reducing the incidence of unaffordable loans.

Lacour-Little and Chun (1999) argue that prepayment rates will be higher on third-party originated loans than lender-originated loans since brokers have incentives to encourage prior customers to refinance, thereby generating additional fees on each transaction. The authors use two sets of data to test their hypothesis. The first is from an unnamed national mortgage loan servicing firm, and includes 16,974 fixed-rate loans originated during 1992 and tracked for prepayment through 1997. To generalize the results to the broader market, the authors use data from a second provider, Mortgage Information Corporation (now LoanPerformance), which compiles aggregate delinquency, default and prepayment data on a national scale.

The loan level data are employed in a logistic regression to predict prepayment, while controlling for loan characteristics as well as the interest rate on the loan and borrower income. A dummy variable representing third-party origination is used to test the hypothesis that third-party loans exhibit faster prepayment speeds. The results show that for each of four loan products evaluated, third-party loans are about three times as sensitive to repayment incentives compared with lender originated loans. Using the national aggregate data, the authors compare mean prepayment speeds stratified by third-party originator versus lender originator. For the period 1994 to 1998 the results are consistent with the loan level findings across all products,

while the differences are smaller. Prepayment speeds prior to 1993, however, are not greater for third-party loans.

III. Data and Empirical Methods

The data for this study come from several sources, among them The Home Mortgage Disclosure Act (HMDA) data for the calendar year 2005, U.S. Census Bureau data for the decennial census of 2000 – with income updates for 2005, the Federal Depository Insurance Corporation's (FDIC's) Summary of Deposit data for June 2005, and the National Association of Mortgage Brokers (NAMB) database for yearend 2005.

Size of the studied sample. The HMDA data represent the substantial majority of mortgage lending performed in the U.S., resisting the criticisms of the American Financial Services Association (AFSA) data insomuch as all mortgage lenders and all mortgage loans are represented in the HMDA data while only a few large lenders are represented in the AFSA data for any particular time period. There are more than 32 million applications represented in the 2005 HMDA data, with almost 6.3 million originations resulting from refinancing an existing loan on 1-to-4 family residences. As Panel A of Table 1 reveals, there were more than 5.5 million first-lien originations resulting from refinancing an existing loan on 1-to-4 family residences in 2005, according to the HMDA data. In addition, there were more than 718,710 originations of second liens on these types of properties in 2005. Of this population, more than 99.80 percent are represented in our study.

Of the studied first lien originations, there are more than 1.4 million observations with reportable rate spreads, representing 25.74 percent of this type, as shown in Panel B of Table 1.

Of the second lien originations shown in Panel B of Table 1, there 214,990 observations with reportable rate spreads, representing almost 30 percent of the second lien originations. For first liens, the average reported rate spread was 4.84 %, while the second-lien average reported rate spread was 6.95 %. The implications here are that the average first lien is 484 basis points higher than the comparable U.S. Treasury security, while the average second lien is 695 basis points higher than the comparable U.S. Treasury.

Correlations among the target variable with the control and explanatory factors. Table 2 contains a correlation matrix of the target variable, Number of Mortgage Brokers, with the dependent, control, and other explanatory variables. All dependent, control, and explanatory variables have statistically significant relationships with the mortgage broker target variable, except for the proportion of Non Originated Offers expressed as a percent. For example, there is also a strong & negative correlation between the number of mortgage brokers in a geography and the likelihood of a rate spread being reported for that geography. This correlation and sign indicate that as the number of mortgage brokers increase, the frequency of reportable rate spreads tends to decline. There is also a strong & positive correlation between the number of mortgage brokers in a geography and the average reported rate spread for that geography. This correlation and sign indicate that as the number of mortgage brokers increase, the average rate spread reported tends to increase. These simple associations are somewhat contradictory on the surface. That is, the first relationship suggests that higher mortgage broker counts in a given geography lower the likelihood of the rate spread being reported. The second of these relationships suggests higher mortgage broker counts increase the size of the average reported rate spread.

Other positive relationships with the number of mortgage brokers are found for Loan Approval Rate, Population, Median Family Income, Median Value of the Housing Stock, and the proportion of the Housing Stock as Rental Units. In short, the number of brokers is higher in geographies where there are higher quality originations among more persons and where the lifestyle & housing are more affluent. The positive correlation with rental housing proportions may indicate geographies with high housing turnover.

Other negative relationships with the number of mortgage brokers are found for Median Age of the Population, Housing Units per 1,000 persons, Median Age of the Housing Stock, percent of the Housing Stock that is vacant, the proportion of persons in the geography who are below the poverty line for the geography, the civilian unemployment rate, and both the number of FDIC lenders & branches in the geography. In sum, these negative relationships indicate that mortgage broker counts are lower in geographies with older populations where the housing stock is plentiful – but older & largely vacant, and where the income levels & unemployment levels would suggest a lower likelihood of successfully brokering a mortgage lending transaction. The relationships further indicate that the count of mortgage brokers declines with an increase in traditional lending alternatives, as indicated by the correlations with the FDIC lender & branch counts for the geography.

Creating Geographic Clusters. Given the high correlations among explanatory factors for a given geography, we employ geographic clusters. Our methods are similar to those used by Amel and Rhoades (1988) and Halpern et al. (1999), each with its underpinnings from Hartigan (1975) and McQueen (1967). Cluster analysis allows us to use several demographic factors to

determine "neighborhood" rather than arbitrarily using one characteristic at a time or forcing the interpretation in a principal components analyses to groups, or baskets, of factors.[5]

To form the geographic clusters, we combine the 2005 HMDA data – which contain the spreads on high interest rate loans, with data from the 2005 FDIC Summary of Deposits; the U.S. Census information population and housing; the 2005 U.S. Department of Housing & Urban Development estimates of family income; and the National Association of Mortgage Brokers geo coded address information on licensed mortgage brokers. In so doing, we create the following variables to form our clusters: Percent of Loan Approvals, Percent of Non Originated Offers, Population, Median Age of the Population, Housing Units, Median Age & Value of the Housing Stock, Percent of Housing Stock as Rental or Vacant, Persons Below the Poverty Level for the Geography, Civilian Unemployment, Number of FDIC-insured lenders per 1,000 persons, and Number of branches of FDIC-insured lenders per 1,000 persons. These variables are defined as follows:

Percent of Loan Approvals. From the 2005 HMDA data, we determine the proportion of studied loans in our sample for which the credit was either originated or offered and compare that to the total of all completed applications. The result is an approval rate. The higher the approval rate, the higher the average quality of credits in the geography. Percent of Loan Approvals acts as a proxy for geographic credit quality.

Percent of Non Originated Offers. From the 2005 HMDA data, we determine the proportion of studied loans in our sample for which the credit was offered but not accepted and compare that to the total of all approvals. The result is an "shopping rate" for credit. That is, some geographies, or neighborhoods, have a higher likelihood of shopping for credit than

[5] The authors recognize the need to test several clustering approaches, ensuring the results are robust to the technique used to form neighborhoods. The next draft of the manuscript will note those tests.

seeking to make application at only one place. The higher the shopping rate, the higher the price sensitivity of the geography. Percent of Non Originated Offers acts as a proxy for geographic credit awareness.

Demographic & Structure County Control Variables. From the 2000 U.S. Census Bureau's decennial census, we create the following control variables:[6]

(a) Population;

(b) Median Age of the Population;

(c) Housing Units;

(d) Median Age & Value of the Housing Stock;

(e) Percent of Housing Stock as Rental and separately, as Vacant;

(f) Percent of Persons Below the Poverty Level; and

(g) Civilian Unemployment.

Competitive or Alternative Lending Outlet Controls. From the June 2005 FDIC Summary of Deposits data, we extract the geo coded location of all FDIC-insured lenders and all FDIC-insured branches. From those data, we create county-level controls for the number of FDIC-insured lenders per 1,000 persons; and separately, for the number of branches of FDIC-insured lenders per 1,000 persons. These two variables attempt to equalize the number of traditional lending outlets available across counties in a given cluster.

Panel A of Table 3 shows the largest four clusters, each having more than 50 counties in the cluster. Representative statistics for some of these variables, as well as the number of

[6] As an alternative to the county as a defined geography, which may represent several "neighborhoods" with blended traits, the authors have created Zip Code Tabulation Areas (ZCTAs) using zip code centroiding to determine the U.S. Census Bureau census tracts within a 3-mile radious of the zip code centroid. The next iteration of the manuscript will employ these geographies, better representing neighborhoods rather than the larger areas of counties.

counties, are also shown in Panel A of Table 3. Panel B of Table 3 reveals representative counties for each cluster.

The largest cluster, Cluster A, contains 2,393 counties – roughly ½ of all counties in the U.S. and its territories. Cluster A possesses the smallest average population per county, with the oldest average population. Cluster A has the lowest family income and the highest fraction of housing units per 1,000 persons. Cluster A has the lowest median home values, the lowest percentage of rental housing, and the highest percentage of vacant housing. Cluster A also has the highest proportion of county population below the poverty line and the worst loan offer rate, while possessing the highest number of average FDIC depository branches per 1,000 persons. Representative counties include Claiborne County, Mississippi; Roosevelt County, Montana; Eureka County, Nevada; and Aurora County, Oregon. In sum, the counties in this cluster are quite rural, sparsely populated by lower-income & older persons, and with high vacancy rates.

The smallest cluster, Cluster D, represented the diametric opposite of Cluster A, containing 69 counties with the highest average population per county, with the youngest average population. Cluster D has the highest family income and the most constrained housing markets, as demonstrated by the lowest fraction of housing units per 1,000 persons. Cluster D also has the highest median home values, the highest percentage of rental housing, and the lowest percentage of vacant housing. Cluster D finally has the best loan offer rate, while possessing the lowest number of average FDIC depository branches per 1,000 persons. The only ranking on which Cluster D is not diametrically opposite Cluster A is the percentage of persons below the poverty line, where Cluster D ranks 3rd to Cluster A's ranking of 4th. Representative counties include San Francisco County, California; New Haven County, Connecticut; The

District of Columbia; Honolulu County, Hawaii; and Westchester County, New York. In sum, the counties in this cluster are quite urban, heavily populated by higher-income & younger persons, and with low vacancy rates.

Mean Difference Tests based on the Number of Mortgage Brokers. For each cluster, we separate the included counties into three groups, those counties that: (a) have no mortgage brokers; (b) have less than the median number of mortgage brokers; and (c) have equal to or more than the median number of mortgage brokers. We perform mean difference tests across these three groups for each cluster on the following two variables: (a) the frequency of reported rate spreads relative to all originations; and (b) the average reported rate spread. These tests are separately performed on 1st lien originations and then for 2nd lien originations, but in each case restricted to refinancings of conventional 1-to-4 family dwellings. In short, no real estate transaction occurs insomuch as there is no sale or purchase, limiting the economic complications of whether the transaction creates other types of value for the homeowner.

Testable hypotheses: Comparative Advantage in Market Information to the Benefit of the Borrower. If mortgage brokers act as agents on behalf of borrowers, exercising better judgment in the choice lending outlets – both in gaining approval and in gaining approval at more favorable rates, then the presence of mortgage brokers have a dampening effect on both the likelihood of a reported rate spread and on the average reported rate spread. All other things equal and restricting the hypotheses to originated credits, we expect the higher the mortgage broker count in a given geography, the lower the likelihood of a reported rate spread – these represent the higher cost loans. We would also expect to find that the average rate spread declines with increases in the number of mortgage brokers in a given geography, as competitive

forces among full-information agents overwhelms the bargaining power of a small number of mortgage brokers with full-information. As such, our testable hypotheses are:

H_{o1}: % Reported Rate Spreads$_{(\text{Fewer Mortgage Brokers})}$ > % Reported Rate Spreads$_{(\text{More Mortgage Brokers})}$

H_{o2}: Average Reported Rate Spread$_{(\text{Fewer Mortgage Brokers})}$ > Average Reported Rate Spread$_{(\text{More Mortgage Brokers})}$

That is, geographies with fewer mortgage brokers should exhibit a higher frequency of reported rate spreads and a higher average reported rate spread than those geographies with more mortgage brokers.

Comparative Advantage in Market Information to the Benefit of the Broker. On the other hand, if mortgage brokers are able to privately use their better information on borrower creditworthiness and mortgage markets, extracting an economic rent from the borrower in securing the credit for him/her, then the presence of mortgage brokers has a magnifying effect on both the likelihood of a reported rate spread and on the average reported rate spread. In this scenario, we expect the higher the mortgage broker count in a given geography, the higher the likelihood of a reported rate spread and the higher the average reported rate spread. In this situation, we may still see declines in the average reported rate spreads as competition among brokers compete away the economic rent:

H_{o1}: % Reported Rate Spreads$_{(\text{Fewer Mortgage Brokers})}$ < % Reported Rate Spreads$_{(\text{More Mortgage Brokers})}$

H_{o2}: Average Reported Rate Spread$_{(\text{Fewer Mortgage Brokers})}$ < Average Reported Rate Spread$_{(\text{More Mortgage Brokers})}$

That is, geographies with fewer mortgage brokers should exhibit a lower frequency of reported rate spreads and a lower average reported rate spread than those geographies with more mortgage

brokers. As the number of mortgage brokers increase, however, competition may drive out the economic rent, hence our interest in evaluating geographies of three types: (1) no mortgage brokers in the geography; (2) few mortgage brokers in the geography; and (3) high numbers of mortgage brokers in the geography.

IV. Empirical Findings

Using mean difference tests, as shown in Panel A of Table 4 for the frequency of reported rate spreads and in Panel B of Table 4 for average reported rate spreads, we look across three groups of counties for each cluster: (a) no mortgage brokers registered within the county; (b) less than the cluster median count of mortgage brokers registered within the county; and (c) more than or equal to the cluster median count of mortgage brokers registered within the county.

Results regarding the frequency of reported rate spreads. In Cluster A – rural, sparsely populated by lower-income & older persons, and with high vacancy rates, we find those counties without any mortgage brokers have a very high 35.21 percent likelihood of a reported rate spread on first-lien refinancings, while those counties with less than the median number of mortgage brokers report a 32.17 percent likelihood of a reported rate spread. The difference here is significant at the 95 percent level, with a t statistic of 2.24. The difference in Cluster A between those counties with less than the median number of mortgage brokers and those with a greater number of mortgage brokers is 32.17 percent versus 26.73 percent, respectively. This difference is again significant at the 95 percent level, with a t statistic of 2.19. In sum, the greater the

number of mortgage bankers in these counties, the lower the likelihood of a reported rate spread. For second-lien refinancings we find no similar result.

In Cluster D – urban, heavily populated by higher-income & younger persons, and with low vacancy rates, there are no counties without mortgage brokers so we may only compare those counties with less than the median number of mortgage brokers and those with a greater number of mortgage brokers. The likelihood of a reported rate spread is then 29.59 percent versus 23.16 percent, respectively. This difference is significant at the 99 percent level, with a t statistic of 2.969. That is, the greater the number of mortgage bankers in these counties, the lower the likelihood of a reported rate spread. For second-lien refinancings we find no similar result.

The findings here for the frequency of reported rate spreads are consistent with both the mortgage broker as a better-informed market agent, as well as with the notion that higher numbers of mortgage brokers compete away any economic rent.

Results regarding the average reported rate spreads. Using Table 4, Panel B, and Cluster A – rural, sparsely populated by lower-income & older persons, and with high vacancy rates, we find those counties without any mortgage brokers have average reported rate spreads of 4.36 percent on first-lien refinancings. In contrast, those counties with less than the median number of mortgage brokers possess an average reported rate spread of 4.92 percent. The difference here is significant at the 99 percent level, with a t statistic of -7.00. The average reported rate spread difference in Cluster A between those counties with less than the median number of mortgage brokers and those with a greater number of mortgage brokers is 4.92 percent versus 5.01 percent, respectively. This difference is insignificant at reasonable levels, with a t

statistic of -1.00. In sum, the difference lies between those counties with mortgage brokers when compared to those counties without mortgage brokers.

For second-lien refinancings in Cluster A counties we find similar results insomuch as the counties without mortgage brokers have an average second-lien reported rate spread of 3.60 percent, while those counties with less than the median number of mortgage brokers have an average second-lien reported rate spread of 4.69 percent – a different of 109 basis points. This result is statistically significant at the 99 percent level, with a t statistic of -3.11. The average reported rate spread difference in Cluster A between those counties with less than the median number of mortgage brokers and those with a greater number of mortgage brokers for second liens is 4.69 percent versus 5.54 percent, respectively – a difference of 85 basis points. This difference is significant at the 95 percent level, with a t statistic of -2.03.

In sum, the differences lie along two lines for counties represented in Cluster A. First those counties with fewer mortgage brokers still have higher average reported rate spreads than those counties without mortgage brokers. Second, those counties with fewer mortgage brokers have lower average reported rate spreads than those counties with more mortgage brokers.

Unlike the findings based on the frequency of reported rate spreads, these pricing results lend favor to the notion mortgage brokers possess and employ a comparative advantage in assessing borrower creditworthiness and knowing the appropriate lending outlets, earning an economic rent in the process, but only in serving sparsely populated counties with lower-income & older persons, and with high vacancy rates. The results further suggest that this credit-and-market knowledge is not readily competed away, since the higher the number of mortgage brokers, the higher the average reported rate spread.

In Cluster D – urban, heavily populated by higher-income & younger persons, and with low vacancy rates, we find those counties with less than the median number of mortgage brokers possess an average reported rate spread of 4.91 percent. The average reported rate spread difference in Cluster D between those counties with less than the median number of mortgage brokers and those with a greater number of mortgage brokers is then 4.91 percent versus 4.80 percent, respectively. This difference is significant at the 95 percent level, with a t statistic of 2.63. In short, as the number of mortgage brokers increase in counties represented in Cluster D, the average reported rate spread declines.

For second-lien refinancings in Cluster D we find insignificant differences insomuch as the average reported rate spread difference in Cluster D between those counties with less than the median number of mortgage brokers and those with a greater number of mortgage brokers for second liens is 6.89 percent versus 6.88 percent, respectively – a difference of 1 basis point.

In sum, for urban counties, heavily populated by higher-income & younger persons, and with low vacancy rates, higher mortgage broker counts do not dampen the magnitude of the average reported rate spread.

These results lend favor to the notion mortgage brokers possess and employ a comparative advantage in assessing borrower creditworthiness and knowing the appropriate lending outlets, earning an economic rent for their comparative advantage but only in rural counties, sparsely populated by lower-income & older persons, and with higher vacancy rates. In urban counties, heavily populated by higher-income & younger persons, and with low vacancy rates, the comparative advantage of the mortgage broker is less evident. The overall pricing results further suggest that this credit-and-market knowledge is not readily competed

away in rural counties and in fact, grows as the number of mortgage brokers grow. The higher the number of mortgage brokers in these rural counties, the higher the average reported rate spread. A higher number of mortgage brokers in urban counties yields no difference in the average reported rate spread, however.

V. Summary and Conclusions

To test the notion that mortgage brokers possess a credit-and-market comparative advantage over borrowers, we designed tests that would control for credit, demographic, and structural differences. We did so by combining the 2005 HMDA data with data from the National Mortgage Brokers Association regarding membership for calendar year 2005, the U.S. Census Bureau's decennial census of 2000, the U.S. Department of Housing & Urban Development's annual update to median family incomes for 2005, and the FDIC's Summary of Deposits databases for June 2005. In so doing, we were able to create four clusters of substantially similar U.S. counties & territories. These clusters represent more than 75 percent of the U.S. population and more than 99.8 percent of all 1-to-4 family refinancings of 1st and 2nd liens during 2005.

To test for differences within clusters, we observed three groups of counties in each cluster: (a) those counties without any NAMB-registered mortgage brokers; (b) those with fewer than the median number of mortgage brokers for the cluster; and (c) those with equal to or more than the median number of mortgage brokers for the cluster. Our mean difference tests suggest

that the behavior of the mortgage broker in use of the comparative advantage he/she has on credit and lending markets depends on the type of credit market he/she faces.

In largely rural counties, sparsely populated by lower-income & older persons, and with higher vacancy rates, we find that the higher the number of mortgage brokers the lower the likelihood that a rate spread will reach the reporting threshold of HMDA. We also find, however, that the higher the number of mortgage brokers, the higher the average reported rate spread, indicating the mortgage brokers' comparative advantage is never fully competed away in these geographies.

In largely urban counties, heavily populated by higher-income & younger persons, and with lower vacancy rates, we also find the higher the number of mortgage brokers the lower the likelihood that a rate spread will reach the reporting threshold of HMDA. A result similar to the rural cluster results. On the other hand, we find no change in the average reported rate spread whether there are a high number or low number of mortgage brokers in these largely urban counties, indicating the mortgage brokers' comparative advantage is fully competed away in these geographies.

Separating these geographies into clusters based on credit, demographic, and structural differences proves key to determining the roles served by the mortgage broker in a given geography. A greater presence by mortgage brokers in a rural geography plays a distinctly different role in pricing than a greater presence of mortgage brokers in an urban geography.

VI. **References**

Amel, Dean F., and Stephen A. Rhoades, 1988, Strategic Groups in Banking, *The Review of Economics and Statistics*, 70:4 (1988), pp. 685-689.

Backley, Christopher W., Jeffrey M. Niblack, Cynthia J. Pahl, Terrence C. Risbey and Jeff Vockrodt, License to deal: Regulation in the mortgage broker industry, *Community Dividend*, Federal Reserve Bank of Minneapolis, 3 (2006).

Calomiris, Charles, High Loan-to-Value Mortgages: Problem or Cure?, *Journal of Lending and Credit Risk Management*, (September 1998), pp. 39-43.

Collins, Michael J., discussant at *Promises & Pitfalls: As Consumer Finance Options Multiply, Who Is Being Served and at What Cost*, a Federal Reserve System Community Affairs Research Conference, July, 2005.

El Anshasy, Amany, Gregory Elliehausen, and Yoshiaki Shimazaki, The Pricing of Subprime Mortgages by Mortgage Brokers and Lenders, *Georgetown University Working Paper*, July 2005.

Halpern, Paul, Robert Kieschnick, and Wendy Rotenberg, On the Heterogeneity of Leveraged Going Private Transactions, *The Review of Financial Studies*, 12:2 (1999), pp. 281-309.

Hartigan, J.A., *Clustering Algorithms*, (1975), New York: John Wiley & Sons, Inc.

Hartigan, J.A., Statistical Theory in Clustering, *Journal of Classification*, 2 (1985), pp. 63 - 76.

LaCour-Little, Michael, and Gregory H. Chun, Third-Party Originators and Mortgage Prepayment Risk: An Agency Problem?, *Journal of Real Estate Research*, 17:1 (1999), pp. 55-70.

McCoy, Patricia A., Banking on Bad Credit: New Research on the Subprime Home Mortgage Market, discussant at *Promises & Pitfalls: As Consumer Finance Options Multiply, Who Is Being Served and at What Cost*, a Federal Reserve System Community Affairs Research Conference, July, 2005.

MacQueen, J.B., Some Methods for Classification and Analysis of Multivariate Observations, *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, 1 (1967), pp. 281 - 297.

Satterthwaite, F.W., An Approximate Distribution of Estimates of Variance Components, *Biometrics Bulletin*, 2 (1946), pp. 110 - 114.

Table 1. Descriptive Statistics. Panel A shows refinancing counts of 2005 HMDA 1-to-4 family originations, both overall and used in this study, by lien status. Panel B shows refinancing counts of 2005 HMDA 1-to-4 family originations used in the study, decomposed in the study's proportion of HMDA-reportable rate spreads and the average HMDA-reportable rate spread, again by lien status.

Panel A. **2005 HMDA Composition: Overall v. Studied**

Lien Status	Overall	Studied	Survival (%)
1st Lien	5,520,410	5,512,830	99.86 %
2nd Lien	718,710	717,510	99.83 %

Panel B. **2005 HMDA Studied: Reported Rate Spreads & Average Rate Spread Data**

Lien Status	Study Total	Reported Rate Spreads	(%)	Average Rate Spread (%)
1st Lien	5,512,830	1,418,970	25.74 %	4.84 %
2nd Lien	717,510	214,990	29.96 %	6.95 %

Table 2. Correlation Matrix. County-based correlations of the mortgage broker count relative to demographic and study factors including: (a) the study variables, Percent of Reported Rate Spreads and Average Reported Rate Spread; (b) proxies for credit quality & credit shopping behaviors, Percent of Loan Approvals and Percent of Non Originated Offers; (c) demographic and housing variables, Population, Median Age of the Population, Housing Units, Median Age & Value of the Housing Stock, Percent of Housing Stock as Rental or Vacant, Persons Below the Poverty Level for the Geography, and Civilian Unemployment; and (d) proxies for access to credit alternatives, Number of FDIC-insured lenders per 1,000 persons and Number of branches of FDIC-insured lenders per 1,000 persons. P-values appear in parentheses beneath the Pearson correlation coefficients for these variables with the Number of Mortgage Brokers within the county.

Variable	Correlation		Variable	Correlation	
Reported Spreads (%)	-0.0573	***	Median Age of the Housing	-0.0486	***
	(0.001)		Stock (yrs)	(0.006)	
Average Reported	0.0328	*	Median Value of the	0.2737	***
Spread (%)	(0.065)		Housing Stock	(<.001)	
Loan Approval Rate (%)	0.1162	***	Housing Stock as	0.2110	***
	(<.001)		Rental (%)	(<.001)	
Non originated Offers (%)	-0.0286		Housing Stock as	-0.1084	***
	(0.113)		Vacant (%)	(<.001)	
Population	0.7338	***	Persons Below Poverty	-0.0778	***
	(<.001)		Level (%)	(<.001)	
Median Age of	-0.0356	**	Civilian Unemployment	-0.0365	**
Population (yrs)	(0.045)		(%)	(0.040)	
Median Family Income	0.2384	***	FDIC-insured lenders	-0.1612	***
	(<.001)		1,000 persons	(<.001)	
Housing Units per 1,000	-0.0666	***	FDIC-insured branches per	-0.1036	***
persons	(<.001)		1,000 persons	(<.001)	

*** Indicates significance at the .01 level or better.
** Indicates significance at the .05 level or better.
* Indicates significance at the .10 level or better.

Table 3. Cluster Descriptions. The clusters were formed at the county level based on similarities in: Percent of Loan Approvals, Percent of Non Originated Offers, Population, Median Age of the Population, Housing Units, Median Age & Value of the Housing Stock, Percent of Housing Stock as Rental or Vacant, Persons Below the Poverty Level for the Geography, Civilian Unemployment, Number of FDIC-insured lenders per 1,000 persons, and Number of branches of FDIC-insured lenders per 1,000 persons. Shown below are the four largest clusters. Representative statistics for some of these variables, as well as the number of counties, are shown below. Descending cluster rankings are shown in parentheses next to each representative statistic.

Panel A: **Clusters & Representative Statistics**

Cluster	Counties	Average Population	Median Age (yrs)	Family Income	Housing Units per 1,000 persons	Median Value: Housing Stock
A	2,393	23,026 (4^{th})	37.19 (4^{th})	$45,110 ($4^{th}$)	461.78 (4^{th})	$69,388 ($4^{th}$)
B	512	112,422 (3^{rd})	35.68 (3^{rd})	$59,794 ($3^{rd}$)	430.92 (3^{rd})	$117,616 ($3^{rd}$)
C	148	344,278 (2^{nd})	35.31 (2^{nd})	$66,739 ($2^{nd}$)	409.07 (2^{nd})	$134,063 ($2^{nd}$)
D	69	728,173 (1^{st})	34.77 (1^{st})	$71,627 ($1^{st}$)	403.03 (1^{st})	$162,540 ($1^{st}$)

Cluster	% Rental Housing	% Vacant Housing	% Below Poverty	Loan Offer Rate	FDIC Branches per 1,000 persons
A	20.76 % (4^{th})	15.49 % (4^{th})	16.27 % (4^{th})	56.36 % (4^{th})	0.50 (4^{th})
B	25.69 % (3^{rd})	10.35 % (3^{rd})	11.44 % (2^{nd})	64.87 % (3^{rd})	0.37 (3^{rd})
C	29.42 % (2^{nd})	7.47 % (2^{nd})	10.75 % (1^{st})	65.82 % (2^{nd})	0.32 (2^{nd})
D	35.22 % (1^{st})	6.00 % (1^{st})	11.48 % (3^{rd})	67.02 % (1^{st})	0.30 (1^{st})

Panel B: **Representative Counties for Each Cluster**

Cluster A. Lee County, AR; Atkinson County, GA; Hardin County, IL; Ellsworth County, KS, Claiborne County, MS; Chariton County, MO; Roosevelt County, MT; Deuel County, NE; Eureka County, NV; McIntosh County, ND; Alfalfa County, OK; Wheeler County, OR; Aurora County, SD; and Castro County, TX.

Cluster B: Sumter County, SC; Wichita County, TX; Vigo County, IN; Jackson County, MI; Lafourche Parish, LA; Macon County, IL; Harnett County, NC; Montgomery County, TN; Chautauqua County, NY; Kennebec County, ME; Manitowoc County, WI; and Cambria County, PA.

Cluster C: Anchorage Municipality, AK; Santa Barbara County, CA; El Paso County, CO; New London County, CT; Anne Arundel County, MD; Barnstable County, MA; Atlantic County, NJ; Rockland County, NY; Henrico County, VA; Spokane County, WA.

Cluster D: San Francisco County, CA; Denver County, CO; New Haven County, CT; The District of Columbia; Honolulu County, HI; Prince George's County, MD; Norfolk County, MA; Westchester County, NY; Delaware County, PA; Providence County, RI; Shelby County, TN; and Salt Lake County, UT.

Table 4. Mean Difference Tests Based on Mortgage Broker Concentration. These mean difference tests are performed across three groups for each of the county clusters depending on mortgage broker concentration: (1) Counties without mortgage brokers; (2) Counties with less than the median number of mortgage brokers; and (3) Counties with more than the median number of mortgage brokers. To arrive at the median mortgage broker count for a given cluster, only counties with mortgage brokers are ranked so there are a proportionately high number of counties represented in the first of these three groups. Panel A contains the results for the frequency of reported rate spreads, while Panel B contains the results for average reported rate spreads. Results are displayed both for first and second lien refinancing loans. T statistics are shown at right.

Panel A: **Frequency of Reported Rate Spreads (%)**

		No Mortgage Brokers		**Mortgage Broker Count < Median (LOW)**		**Mortgage Broker Count ≥ Median (HI)**		**Mean Difference Tests**	
Cluster	**Lien**	**(%)**	**N**	**(%)**	**N**	**(%)**	**N**	**0 v. LOW**	**LOW v. HI**
A	1st Lien	35.21 %	2,088	32.17 %	222	26.73 %	83	2.24**	2.19**,a
	2nd Lien	26.22 %	1,557	29.44 %	209	31.42 %	80	-1.11	-0.53 a
B	1st Lien	27.24 %	140	26.62 %	208	24.10 %	164	0.42	3.12***,a
	2nd Lien	29.05 %	140	30.58 %	207	30.07 %	164	-0.60	0.34a
C	1st Lien	29.63 %	8	28.46 %	68	24.22 %	72	0.20	3.06***,a
	2nd Lien	28.29 %	8	27.17 %	68	30.44 %	72	0.17	-1.69*
D	1st Lien	NA	--	29.59 %	34	23.16 %	35	NA	2.96***
	2nd Lien	NA	--	29.25 %	34	28.92 %	35	NA	0.13

[a] Variances are unequal. T statistics employ Satterthwaite's (1946) correction to degrees of freedom.
*** Indicates significance at the .01 level or better.
** Indicates significance at the .05 level or better.
* Indicates significance at the .10 level or better.

Table 4. Mean Difference Tests Based on Mortgage Broker Concentration (continued).

Panel B: **Average Reported Rate Spread (%)**

Cluster	Lien	No Mortgage Brokers (%)	N	Mortgage Broker Count < Median (LOW) (%)	N	Mortgage Broker Count ≥ Median (HI) (%)	N	Mean Difference Tests 0 v. LOW	LOW v. HI
A	1st Lien	4.36 %	2,088	4.92 %	222	5.01 %	83	-7.00***	-1.00[a]
	2nd Lien	3.60 %	1,557	4.69 %	209	5.54 %	80	-3.11***	-2.03**,[a]
B	1st Lien	4.92 %	140	4.85 %	208	4.79 %	164	1.40	2.18**,[a]
	2nd Lien	6.59 %	140	6.82 %	207	7.02 %	164	-0.72	-1.33[a]
C	1st Lien	4.83 %	8	4.89 %	68	4.76 %	72	-0.67	2.81***
	2nd Lien	7.26 %	8	7.01 %	68	6.99 %	72	0.76	0.20[a]
D	1st Lien	NA	--	4.91 %	34	4.80 %	35	NA	2.63**,[a]
	2nd Lien	NA	--	6.89 %	34	6.88 %	35	NA	0.17

[a] Variances are unequal. T statistics employ Satterthwaite's (1946) correction to degrees of freedom.
*** Indicates significance at the .01 level or better.
** Indicates significance at the .05 level or better.
* Indicates significance at the .10 level or better.