How Subprime Borrowers and Mortgage Brokers Shared the Pie

Antje Berndt, Burton Hollifield and Patrik Sandås

July 2014

How Subprime Borrowers and Mortgage Brokers Shared the Pie[*]

Antje Berndt[†]　　　　Burton Hollifield[‡]　　　　Patrik Sandas[§]

Abstract

We develop an equilibrium model for origination fees charged by mortgage brokers and show how the equilibrium fee distribution depends on borrowers' valuation for their loans and their information about fees. We use non-crossing quantile regressions and data from a large subprime lender to estimate conditional fee distributions. Given the fee distribution, we identify the distributions of borrower valuations and informedness. The level of informedness is higher for larger loans and in better educated neighborhoods. We quantify the fraction of the surplus from the mortgage that goes to the broker, and how it decreases as the borrower becomes more informed.

[†]Poole College of Management, North Carolina State University, Raleigh, NC, 27695, Email: aberndt@ncsu.edu.

[‡]Tepper School of Business, Carnegie Mellon University, Pittsburgh, PA, 15213. Email: burtonh@andrew.cmu.edu.

[§]McIntire School of Commerce, University of Virginia, Charlottesville, VA, 22904. Email: patriks@virginia.edu. Sandås is a research affiliate of SIFR.

Borrowers in the residential mortgage market take out a loan to purchase a new home or to refinance an existing mortgage. For loans originated through an independent mortgage broker, the broker charges the borrower a fee for the origination services provided. Different borrowers often pay different fees for the same loan. We develop a model that links the conditional variation in broker fees to the conditional variation in the value borrowers assign to their loan and the amount of information borrowers have about broker compensation. Our framework provides consumer advocates and regulators with a tool to analyze how mortgage origination charges may change in response to better educating borrowers about how to evaluate loan proposals, or to better educating borrowers about the origination process and the importance of being informed about fees.

Our goal is to quantify the average fraction of the overall surplus from the mortgage that the broker is able to capture. That fraction depends on the borrower's valuation for the loan, and on how informed the borrower is. We use our model to empirically identify the conditional distributions of borrower valuations and informedness from the conditional broker fee distribution. The fee distribution is estimated from loan-level mortgage origination and broker compensation data, using a non-crossing quantile regression technique that is new to the real estate and finance literature. Our data contain detailed records of almost 47,000 brokered loans funded by New Century Financial Corporation between 2004 and 2006. During those years, New Century was one for the top three subprime lenders in the U.S..

Focusing on subprime borrowers, we use a simple model of bargaining between borrower and broker where the broker learns the borrower's reservation value for the fees and has all the bargaining power. The broker sets her fees equal to the borrower's reservation value. The borrower's reservation value does not exceed the value he assigns to the loan, which is measured as the value of the benefits he expects to draw from owning the home in excess of the expected present value of the mortgage payments and the outside option of not obtaining the loan. If the outside option is to rent a similar home or to keep the existing mortgage for

1

some time before making the mortgage transaction, the borrower's valuation is his perceived cost of waiting to take out the loan plus the present value of the fees he expects to pay later.

Borrowers form their reservation value for the fees in one of three ways: (i) *Fully informed borrowers* learn the broker's reservation value, possibly by shopping from different brokers, and set their reservation value for the fees equal to the broker's reservation value; (ii) *Partially informed borrowers* consult with a friend about the fees that the friend paid in a similar recent mortgage transaction and set their reservation value equal to the minimum of the fees paid by the friend and their private valuation for the loan; (iii) *Minimally informed borrowers* neither shop from different brokers nor consult with a friend. They set their reservation value equal to their private valuation for the loan.[1]

We show how the equilibrium distribution of borrower reservation values and broker fees—which is one where the distribution of reservation values is the same as the distribution of fees—depends on the distributions of borrower valuations and informedness. We invert this link to express the borrower valuation distribution as a function of the equilibrium fee distribution and the borrower informedness distribution.

For a given fee distribution, an increase in the level of informedness in the borrower population is associated with a shift of the entire borrower valuation distribution towards larger values. This is an important feature of the model as it allows us to identify the level of informedness in the borrower population from knowledge of a single point on the borrower valuation distribution. We develop a framework for measuring the median borrower valuation and use the median valuation to identify the implied distribution of borrower informedness. This in turn allows us to compute the implied distribution of borrower valuations.

---

[1]The notion of borrowers consulting with friends goes back to Woodward and Hall (2009), who distinguish between borrowers who are fully informed and borrowers who consult with friends. In their paper, borrowers who consult with friends also gather information from the broker and set their reservation value equal to the minimum of the fees their friends paid plus some upward effect that the broker has on the reservation value. Woodward and Hall do not cap the borrower's reservation value at his valuation for the loan. Moreover, they do not consider a scenario where the borrower neither learns the broker's reservation value nor consults with friends.

We use non-crossing quantile regressions to obtain estimates for the broker fee distribution conditional on loan and borrower characteristics. Classical quantile regression analysis estimates each quantile function separately. The resulting regression functions may cross in finite samples. To ensure that the estimated fee distribution is well defined, we implement non-crossing quantile regressions, which estimate the quantile functions simultaneously and under an explicit non-crossing constraint.

For a hybrid refinance loan and a reference specification of loan and borrower characteristics, expected fees per loan are about $4.4K. Survey evidence in Lacko and Pappalardo (2007) and ICF Macro (2009) suggests that only a small fraction of borrowers are fully informed, which allows us to identify the broker's reservation value for the fees—that is, the fees paid by a fully informed borrower—as a low quantile of the estimated fee distribution. A fully informed borrower pays $0.9K per loan, a partially informed borrower pays an average of $3.6K, and a minimally informed borrower pays an average of $5.0K. More than half of the borrower population is minimally informed.

The overall surplus from the mortgage is equal to the amount by which the borrower's valuation exceeds the broker's reservation value for the fees. The fraction of the surplus that goes to the broker is zero for fully informed borrowers and one for minimally informed borrowers. For partially informed borrowers, we estimate it to be 72% on average. If a partially informed borrower were to consult with one extra friend, he would save an average of $0.6K per loan. The fraction of the overall surplus from the mortgage that such a borrower leaves on the table would be reduced by thirteen percentage points.

All else the same, expected fees are estimated to be higher for larger loans or when the percentage of neighbors with a B.A. degree is lower. The effect of loan size on fees is associated with a size effect on borrower valuations and a size effect on borrower informedness. For larger loans, borrowers' valuations tend to be higher and borrowers tend to be more informed. The effect of education on fees, on the other hand, mostly translates into an effect

on borrower informedness. When the percentage of neighbors with a college degree is lower, borrowers tend to be less informed about fees. The distribution of borrower valuations, however, remains nearly unchanged.

We develop a framework of how borrowers decide whether or not to consult with a friend. We show that large borrowers are more inclined to gather additional information about fees because they perceive the cost of waiting to take out the loan to be high relative to fees. Borrowers with more educated neighbors, on the other hand, have the same perception of the cost of waiting and the fees as those with fewer educated neighbors. They do, however, consult with a friend more frequently, perhaps because they benefit from the advice of educated neighbors to become more informed about fees.

Our findings are qualitatively similar across different loan purpose and product type combinations. For each stratum of the data, partially informed borrowers leave a large fraction of the overall surplus from the loan on the table. Our results are robust to extending the model by allowing valuation distributions to depend on the level of borrower informedness.

1. Linking Broker Fees, Borrower Informedness and Borrower Valuations

We develop a model of the mortgage origination process to understand how broker origination fees are determined and what they may reveal about the borrower's valuation for the loan. We focus on loans originated in the wholesale market, where independent mortgage brokers act as financial intermediaries matching borrowers with lenders. Brokers assist borrowers in the selection of the loan and in completing the loan application, and provide services to wholesale lenders by generating business and helping them complete the paperwork.

Consider a borrower who arrives at a broker requesting a mortgage. The broker evaluates the borrower's and property's characteristics. He also reviews wholesale rate sheets distributed by potential lenders. These rate sheets state the minimum rate—called the base rate—at which a given lender is willing to finance a loan, as a function of loan and borrower

characteristics. Rate sheets also inform the broker about the yield spread premium (YSP), if any, that the lender pays to the broker for originating the loan at a rate higher than the base rate. Based on this information, the broker provides the borrower with a financing option. A financing option is a specification of the terms of the loan, such as the product type, level of income documentation, loan amount and mortgage rate. With the financing option on the table, the borrower and broker bargain over the broker's fees.

The borrower has the outside option of not taking out the loan. For example, he could wait and take out a loan at a later time. We therefore assume that the borrower has a reservation value for the fees, $\overline{f}$. The reservation value is defined as the largest amount of fees that the borrower is willing to pay for the loan. It does not exceed the borrower's perceived benefits from taking out the loan, ν, so that $\overline{f} \leq \nu$. We refer to ν as the borrower's valuation for the loan. It is defined as the dollar value of the benefits the borrower expects to draw from purchasing the house or refinancing the existing mortgage in excess of the expected present value of the mortgage payments and the perceived dollar value of the outside option of not obtaining the loan. If the borrower learns that he can realize fees below ν, perhaps by consulting with recent mortgagors about the fees they paid or by shopping from other brokers, then the borrower's reservation value for the fees would be strictly lower than his valuation for the loan: $\overline{f} < \nu$. In any case, once the fee f is set the borrower's net benefit from interacting with the broker is $\overline{f} - f$.

The broker's reservation value for the fees, $\underline{f}$, is equal to

$$\underline{f} \;=\; c - y, \tag{1}$$

where y denotes the YSP paid by the lender and c denote the broker's cost of originating the loan. Broker costs are the costs the broker expects to incur between the time she strikes a deal with the borrower and the time the loan closes. They include the broker's time costs

of dealing with the borrower as well as any administrative costs incurred by the broker for intermediating the mortgage. To keep things simple, we abstract from any unobserved heterogeneity in broker costs. The broker's net benefit from originating the loan is $f - \underline{f}$. The borrower's and broker's joint gains from trade is the sum of their respective benefits, $(\overline{f} - f) + (f - \underline{f}) = \overline{f} - \underline{f}$.

We consider a simple model of bargaining between the borrower and broker where the broker learns the borrower's reservation value $\overline{f}$ and has all the bargaining power. The broker sets the fee equal to the borrower's reservation value,

$$ f \;\; = \;\; \overline{f}, \tag{2} $$

and, as long as $\underline{f} \leq \overline{f}$, a deal is made between the borrower and the broker. The broker then submits a funding request to one or more lenders. The lender reviews the application material and responds with a decision to fund the loan or not.[2] If the loan is funded, the broker receives the fees and YSP at the loan closing. The broker sets the terms of the loan so as to maximize the expected gains from trade, subject only to the constraint $\underline{f} \leq \overline{f}$.

We focus on funded loans. Equation (2) implies that the surplus associated with the interaction between borrower and broker goes entirely to the broker. Even though the borrower's net benefit from his interaction with the broker is zero, his surplus from obtaining the mortgage, $\nu - f = \nu - \overline{f}$, is positive as long as the borrower's valuation for the loan exceeds his reservation value for the fees. The overall surplus from the mortgage is defined as the sum of the borrower's surplus from obtaining the mortgage and the broker's surplus from originating the mortgage: $(\nu - \overline{f}) + (\overline{f} - \underline{f}) = \nu - \underline{f}$. It is equal to the dollar amount by which the borrower's valuation for the loan exceeds the broker's reservation value for the

---

[2]We suppose that a lender will fund the loan as long as the broker collects and transfers the requested application materials and secures a rate at or above the lender's base rate. Since the broker is paid only if the loan is made, she will only offer fundable proposals to the borrower and will ensure that the application materials are presented to the lender in a timely fashion.

fees. The fraction of the overall surplus from the mortgage that goes to the broker is

$$\gamma \; := \; \frac{\overline{f} - \underline{f}}{\nu - \underline{f}}, \tag{3}$$

and the fraction that goes to the borrower is $1 - \gamma$.

In what follows, we consider a given set of loan and borrower characteristics that includes YSP but excludes broker fees. Our goal is to quantify the distributions of $\overline{f}$, ν, and γ across all mortgage transactions with such characteristics. To do so, we propose a model for how borrowers form their reservation values $\overline{f}$, and solve it for the equilibrium distribution of $\overline{f}$.[3] In particular, we classify borrowers by how informed they are about attainable broker fees:

- **Fully informed borrower:** The borrower is an expert in the mortgage market in that he learns the broker's reservation value for the fees, $\underline{f}$, by shopping from different brokers. He sets his reservation value $\overline{f}$ equal to $\underline{f}$.

- **Partially informed borrower:** The borrower does not shop from different brokers. Instead, the borrower consults with a friend who is entirely honest about the broker fee f_1 he paid in a similar recent mortgage transaction. The borrower forms his reservation value as $\overline{f} = \min\{f_1, \nu\}$.

- **Minimally informed borrower:** The borrower neither shops from different brokers nor consults with a friend. He sets his reservation value $\overline{f}$ equal to ν.

We define π_0, π_1 and π_2 as the fractions of fully, partially and minimally informed borrowers, and refer to $\pi = (\pi_0, \pi_1, \pi_2)$ as the distribution of borrower informedness. We consider the case where only a small fraction of borrowers are experts, i.e., fully informed, and where

---

[3]Our model for the borrower reservation value $\overline{f}$ is an extension of the framework in Woodward and Hall (2009), which does not allow for minimally informed borrowers and does not ensure that the reservation value of borrowers is capped at ν.

most borrowers are non-experts, i.e., partially or minimally informed.[4] Non-expert borrowers are either unaware of the fact that shopping from different brokers is likely to result in a lower fee, or assign a high personal cost to fee shopping.

Non-expert borrowers lower their fees by consulting with a friend only if the friend's fees are less than the borrower's valuation for the loan: $f_1 < \nu$. We assume that f_1 is a random draw from the population of recent borrowers with the same characteristics, and that no change in mortgage-market conditions has occurred since they borrowed. We also assume that non-expert borrowers have to decide whether or not to consult with a friend prior to observing their private valuations ν—perhaps because the time it takes them to learn ν and to consult with a friend prevents them from performing these tasks sequentially—and that ν is independent of their decision and independent of f_1.[5]

We now solve for the equilibrium distribution of borrower reservation values. Let $F_{\bar{f}}$ denote the cumulative distribution function of borrower reservation values, F_f the cumulative distribution function of broker fees, and F_ν the cumulative distribution function of borrower valuations. Unlike $F_{\bar{f}}$ and F_f, F_ν is computed conditional on the borrower being a non-expert. Suppose that F_ν has support on $[\underline{f}, \bar{\nu}]$, for $\bar{\nu} > \underline{f}$. For any $a \in [\underline{f}, \bar{\nu}]$, we have

$$
\begin{aligned}
F_{\bar{f}}(a) &= \pi_0 + \pi_1 Prob(\min\{f_1, \nu\} \le a) + \pi_2 F_\nu(a) \\
&= \pi_0 + \pi_1 \{1 - (1 - F_f(a))(1 - F_\nu(a))\} + \pi_2 F_\nu(a).
\end{aligned}
\tag{4}
$$

The equilibrium distribution of borrower reservation values or, equivalently, the equilibrium distribution of broker fees is one where $F_{\bar{f}} = F_f$. Imposing the equilibrium condition,

---

[4]Two borrower surveys, Lacko and Pappalardo (2007) and ICF Macro (2009), find that many borrowers shop from just one broker and that only few borrowers shop from a large number of brokers. This motivates us to keep π_0 small. Further anecdotal evidence regarding the hurdles to comparison shopping is provided by Engel and McCoy (2011).

[5]A mechanism for how non-expert borrowers decide whether or not to consult with a friend is described in Section 5. A specification of such a mechanism, however, is not necessary when solving for the equilibrium distribution of $\bar{f}$. Situations in which non-expert borrowers' valuations depend on their level of informedness are analyzed in Section 7.

Equation (4) becomes

$$F_f(a)(1 - \kappa_1(a)) + \kappa_1(a) - \kappa_0(a) \;=\; 0, \tag{5}$$

where $a \in [\underline{f}, \overline{\nu}]$, $\kappa_1(a) = \pi_1(1 - F_\nu(a))$ and $\kappa_0(a) = 1 - \pi_2(1 - F_\nu(a))$. In Appendix A, we prove that there is a unique equilibrium to the model:

Proposition 1 *Let F_ν be some distribution function with support on $[\underline{f}, \overline{\nu}]$. For $\pi_0, \pi_1, \pi_2 > 0$, the equilibrium distribution of broker fees is well-defined by Equation (5), has support on $[\underline{f}, \overline{\nu}]$ and is unique.*

Proposition 2 states that independent of π, the equilibrium fee distribution lies to the left of the valuation distribution, meaning it is shifted towards smaller values.

Proposition 2 *For $\pi_0, \pi_1, \pi_2 > 0$, the equilibrium distribution of broker fees satisfies $F_f(a) \geq F_\nu(a)$ for all $a \in [\underline{f}, \overline{\nu}]$.*

Empirically, we observe loan-level mortgage origination and broker compensation data and can estimate the broker fee distribution conditional on a set of loan and borrower characteristics. But neither the conditional distribution of borrower informedness nor the conditional distribution of borrower valuations is directly available. In the next section, we show how to use the equilibrium condition (5) to learn about π and F_ν.

2. Identification of the Borrower Informedness and Valuation Distributions

For a given set of loan and borrower characteristics and the associated fee distribution F_f, we now show how to compute the implied distributions of borrower informedness and valuations. First, we invert Equation (5) and express F_ν as a function of F_f and π:

$$F_\nu(a) \;=\; 1 - \frac{1 - F_f(a)}{\pi_1(1 - F_f(a)) + \pi_2} \tag{6}$$

for $a \in [\underline{f}, \overline{\nu}]$, and $F_\nu(a) = 0$ for $a < \underline{f}$. Proposition 3 states the conditions under which F_ν defined in Equation (6) is a well-behaved distribution function, that is, the conditions under which we can invert the fee distribution for the valuation distribution given knowledge of π:

Proposition 3 *Let F_f be some distribution function with support on $[\underline{f}, \overline{\nu}]$. For $\pi_0, \pi_1, \pi_2 > 0$ and $F_f(\underline{f}) \geq \pi_0/(1 - \pi_1)$, F_ν given by Equation (6) is a well-defined distribution function with support on $[\underline{f}, \overline{\nu}]$.*

Second, we impose the technical condition $F_\nu(\underline{f}) = 0$, which implies

$$\pi_1 = \frac{1 - F_f(\underline{f}) - \pi_2}{1 - F_f(\underline{f})}. \tag{7}$$

By combining Equations (6) and (7), we can rewrite F_ν as a function of F_f and π_2:

Corollary 1 *Let F_f be some distribution function with support on $[\underline{f}, \overline{\nu}]$ and $F_f(\underline{f}) \in (0, 1)$. Let $\pi_2 \in \big(0, 1 - F_f(\underline{f})\big)$, and let π_1 be defined as in Equation (7). Then F_ν given by Equation (6) is a well-defined distribution function with support on $[\underline{f}, \overline{\nu}]$. For each $a \in [\underline{f}, \overline{\nu}]$,*

$$F_\nu(a) = \frac{(F_f(a) - F_f(\underline{f}))\, \pi_2}{(1 - F_f(\underline{f}))\,(1 - F_f(a)) + (F_f(a) - F_f(\underline{f}))\, \pi_2}. \tag{8}$$

Third, we observe that Equation (8) expresses $F_\nu(a)$ as an increasing function of the fraction of π_2, for each a. An increase in π_2 is therefore associated with a shift of the entire valuation distribution towards smaller values. This is an important feature of our model as it allows us to identify π from knowledge of a single point on the borrower valuation distribution F_ν. We compute a benchmark borrower valuation b_ν and specify $F_\nu(b_\nu)$. Then Equation (8) yields

$$\pi_2 = \frac{F_\nu(b_\nu)}{1 - F_\nu(b_\nu)} \frac{(1 - F_f(\underline{f}))\,(1 - F_f(b_\nu))}{(F_f(b_\nu) - F_f(\underline{f}))}. \tag{9}$$

10

Given π_2, π_1 and π_0 can be computed using Equation (7) and the identity $1 = \pi_0 + \pi_1 + \pi_2$. Once we have identified π, we can obtain the entire valuation distribution via Equation (6).

In the remainder of this section we propose an approach to quantify b_ν and $F_\nu(b_\nu)$. To this end, consider a non-expert borrower with benchmark valuation b_ν. Recall that the borrower's valuation for the loan is his perceived net benefit from taking out the loan. If the borrower expects to pay off the mortgage after T months, either because he expects to sell the home or refinance the loan at that point, b_ν is given as

$$b_\nu \;=\; H_{0 \to T} + V_T - P_{0 \to T} - o. \tag{10}$$

Here, $H_{0 \to T}$ is the value the borrower receives from occupying the house between now and time T, V_T is the value of the house at time T, and $P_{0 \to T}$ is the value of the payments associated with taking out a mortgage today—net of broker fees—plus the costs of unwinding the mortgage at time T. The value of the outside option of not obtaining the loan is denoted by o. All quantities are expressed in today's dollars and are as perceived by the borrower.

We assume that the borrower's outside option is to take out the same type of loan as the one being offered now m months later, where $m \ll T$. For purchase loans, the outside option is to rent the same or a similar home for m months before purchasing the same or a similar home. For refinance loans, it is to keep the existing mortgage for m more months before refinancing. Before taking out the loan m months later, the borrower consults with one friend and expects to pay the same broker fee as that friend.

Given the borrower's outside option, he expects to live in the same or a similar house between now and time T, and to own the home by time T, whether or not he takes out the loan now or later. This is important because it implies that the borrower's benefits from the outside option are the same as his benefits from the current loan: $H_{0 \to T} + V_T$.[6] The value

---

[6]For purchase loans, this assumption ignores any perceived soft benefits from living in the home for the first m months as owner rather than renter. One could account for any potential soft benefits that the borrower

11

of the outside option is therefore given by

$$o \;=\; H_{0 \to T} + V_T - \widetilde{P}_{0 \to m} - \widetilde{P}_{m \to T} - \widetilde{f}_m, \tag{11}$$

where $\widetilde{P}_{0 \to m}$ denotes the value of the payments associated with renting a similar home for the first m months or keeping the existing mortgage for m more months, $\widetilde{P}_{m \to T}$ denotes the value of the payments associated with taking out the loan at time m—net of broker fees—plus the cost of unwinding the mortgage at time T, and $\widetilde{f}_m$ denotes the broker fees paid when taking out the loan in m months. As before, all quantities are expressed in today's dollars and are as perceived by the borrower. We use "$\sim$" to distinguish payments associated with waiting m months before taking out a loan from those associated with taking out the loan today.

Substituting Equation (11) into Equation (10) yields

$$b_\nu \;=\; b_k + b_{\widetilde{f}},$$

where $b_k = \widetilde{P}_{0 \to m} + \widetilde{P}_{m \to T} - P_{0 \to T}$ is the borrower's pre-fee cost of waiting to take out the loan and $b_{\widetilde{f}} = \widetilde{f}_m$ is his estimate of discounted average future fees. We refer to b_k and $b_{\widetilde{f}}$ as the benchmark cost of waiting and benchmark future fees, respectively. In the next section, we introduce our data and show how to compute empirical measures for b_k and $b_{\widetilde{f}}$.

In our empirical applications, we interpret b_ν as the median borrower valuation, so that $F_\nu(b_\nu) = 0.5$ and $F_\nu(b_\nu)/(1 - F_\nu(b_\nu)) = 1$ in Equation (9).[7]

3. Data, Benchmarks and Conditioning Variables

We describe our loan-level mortgage market data, show how to compute empirical measures for b_k and $b_{\widetilde{f}}$, and describe the conditioning variables used in our empirical analysis.

---

derives from "homeowner status" by adding a negative scalar to the right-hand side of Equation (11).

[7]In Corollary 2 in Appendix A, we show that π and F_ν can also be identified when b_ν is interpreted as the mean borrower valuation.

3.1 Mortgage origination and broker compensation data

Our dataset of mortgage originations and broker compensation is obtained from IPRecovery, Inc. and contains detailed records of all broker-originated loans funded by New Century Financial Corporation. New Century was founded in the mid-1990s in California and quickly grew into a nationwide lender that originated, retained, sold and serviced residential mortgages designed for subprime borrowers. Berndt, Hollifield, and Sandas (2012) compare New Century's origination, broker compensation and loan performance statistics to those for the overall subprime market. They document that prior to the subprime crisis, the New Century loan pool was representative of the broader subprime market.[8] We therefore believe that our empirical analysis is not simply a case study of New Century, and that our findings shed light on the pre-crisis subprime market in general.

We focus on free-standing first-lien loans for single-unit primary residences originated during the three years leading up to the 2007 subprime crisis, 2004–2006. The loan types considered are 30-year fixed mortgages (fixed loans) and 2/28 loans (hybrid loans). The latter are 30-year hybrid loans that have a two-year fixed-interest-rate period after which the interest rate on the mortgage begins to float based on LIBOR plus a margin. The sample is restricted to loans originated in large metropolitan areas for which socio-economic variables such as house price appreciation, property taxes and price-to-rent ratios can be observed. The final dataset is comprised of almost 47,000 loans.[9]

We stratify the data by loan purpose and product type. Nearly half of the loans in our sample are hybrid refinance loans. Fixed refinance, hybrid purchase and fixed purchase

---

[8]Following its bankruptcy filing in 2007, New Century received widespread attention in the popular press, mainly because it was the largest subprime lender to default by that date. By 2009, however, virtually all of New Century's main competitors had either declared bankruptcy, had been absorbed into other lenders, or had otherwise unwound their lending activities.

[9]In Texas, cash-out refinance mortgages are subject to the "Texas Home Equity 3% Fee Cap" which limits broker fees to 3% of the loan amount (see Section 50(a)(6)(E) of the Texas Constitution). Since a binding exogenous constraint on the fee distribution would invalidate the link between broker fees and borrower valuation established in Equation (5), we exclude refinance loans originated in Texas from our sample.

loans account for 25%, 21% and 5% of the sample, respectively. For each stratum, Table 1 reports descriptive statistics for loan and borrower characteristics. Table B.1 in the appendix provides a definition for each characteristic. For hybrid refinance loans, the average loan amount is \$185K. The average loan-to-value (LTV) ratio is close to 80%, the average FICO score is around 580, and the average debt-to-income (DTI) ratio is about 40%. The majority of the loan contracts require full documentation and include a prepayment penalty clause. The average percentage of neighbors with a B.A. degree is 14%.

[Table 1 about here.]

During our sample period, nearly 12,500 different brokerage firms did business with New Century. They earned revenues from two sources: a direct fee paid by the borrower and an indirect fee—the yield spread premium—paid by the lender. Direct fees include all compensation associated with the mortgage origination paid by the borrower directly to the broker, including finance charges such as appraisal and credit report fees. The YSP rewards the broker for originating loans with higher mortgage rates, holding other things equal.[10]

Table 2 shows that average broker revenues per loan, measured as a percentage of loan amount, range between 2.9% and 3.4%, depending on the loan purpose–product type combination.[11] The majority of the revenues stems from fees. For hybrid refinance loans, average percentage fees, YSP and revenues amount to 2.4%, 1.0% and 3.4%, respectively.[12] Aver-

---

[10]New loan originator compensation rules went in effect April 1, 2011 as part of the Federal Reserve Board's Regulation Z. They prohibit mortgage broker compensation to vary based on loan terms other than principal. In particular, brokers can no longer receive yield spread premia from the lender.

[11]Ambrose and Conklin (2014) use New Century data to study a different aspect of mortgage broker compensation. They focus on refinance loans and report the sum of percentage broker revenues and percentage lender fees as 3.8% for their sample.

[12]About 24% of the YSP entries in our data are left blank. All else the same, loans with lower FICO scores, lower risk grades and less documentation are more likely to have a missing YSP entry. Such loans usually have high base rates, leaving less room for brokers to convince borrowers to pay rates in excess of the base rate. Moreover, while an increase in YSP is usually associated with a decrease in direct broker fees, we find no statistical significance for a missing-YSP dummy when regressing broker fees on YSP and other observable covariates. With this in mind, we interpret missing-YSP entries as zero YSP, which brings the percentage of zero-YSP loans in our data to 28%. Our findings are robust, however, to excluding missing-YSP loans.

age dollar fees, YSP and revenues per loan are about $4.0K, $1.8K and $5.7K, respectively. Broker fees tend to be lower for purchase loans than for refinance loans.

[Table 2 about here.]

Figure 1 shows that independent of loan purpose and product type, dollar fees tend to increase as loan size increases while percentage fees tend to decrease as loan size increases.

[Figure 1 about here.]

3.2 Benchmark future fees and cost of waiting

We consider a non-expert borrower with benchmark valuation and assume that he is risk-neutral. Lacking better information, the borrower bases his estimate of discounted average future fees, $b_{\widetilde{f}}$, on his perception of average current fees for the kind of loan he expects to take out m months from now. In other words, we make the assumption that

$$b_{\widetilde{f}} \;\; = \;\; \frac{b_f(\widetilde{X}_m)}{1 + r_{inv}}, \tag{12}$$

where $\widetilde{X}_m$ describes the loan that the borrower expects to take out at time m, $b_f(\widetilde{X}_m)$ is the borrower's perception of average current fees for loans with characteristics $\widetilde{X}_m$, and r_{inv} is the m-month rate of return on investments, adjusted for capital gains taxes.

The borrower computes $b_f(\widetilde{X}_m)$ based on aggregate information about broker fees. Fee data may be available online or via government-sponsored surveys, but are generally not neighborhood-specific.[13] Suppose it is reported that average broker fees for a loan with characteristics $\widetilde{X}_m$ range between $b_f^{low}(\widetilde{X}_m)$ and $b_f^{high}(\widetilde{X}_m)$, depending on the zip code in

---

[13]For example, HSH Associates regularly conducts a national survey of more than 150 lenders and reports national averages for origination fees on a semi-annual basis (see library.hsh.com/articles/more-tools-resources-and-info/mortgage-basics/average-selected-closing-fees-and-charges/). Government-sponsored surveys of mortgage origination fees include Lacko and Pappalardo (2004) and Woodward (2008).

15

which the property is located. We assume that without additional information the borrower is cautious, and sets $b_f(\widetilde{X}_m) = b_f^{high}(\widetilde{X}_m)$.

To compute the borrower's benchmark cost of waiting, b_k, we use empirical measures for $\widetilde{P}_{0 \to m}$, $\widetilde{P}_{m \to T}$ and $P_{0 \to T}$. For purchase loans, $\widetilde{P}_{0 \to m}$ is the value of the payments associated with renting the same or a similar home for m months, $\widetilde{P}_{0 \to m}^{rent}$. It is computed as

$$\widetilde{P}_{0 \to m}^{rent} \;=\; \frac{m \times rent \times (1 + ins_r) + deposit \times r_{inv}}{1 + r_{inv}}, \tag{13}$$

where *rent* denotes the monthly rent for the same or a similar home, ins_r is the rent insurance rate, and *deposit* is the renter's deposit. To obtain an empirical measure for $\widetilde{P}_{0 \to m}^{rent}$, we evaluate the right-hand-side of Equation (13) using the inputs listed in Table 3.

[Table 3 about here.]

For refinance loans, $\widetilde{P}_{0 \to m}$ is the value of the payments associated with keeping the existing mortgage for m more months, $\widetilde{P}_{0 \to m}^{ex\, mtg}$. It is computed as

$$\widetilde{P}_{0 \to m}^{ex\, mtg} \;=\; \frac{\sum_{i=1}^{m} recurring\ costs_i^{ex\, mtg} + \sum_{i=1}^{m} opportunity\ costs_i^{ex\, mtg}}{1 + r_{inv}}. \tag{14}$$

The recurring costs for the existing mortgage for month i, $recurring\ costs_i^{ex\, mtg}$, consists of (i) the mortgage payments, $p_i + i_i(1 - \tau)$, where p_i and i_i are the principal and interest payments in month i and τ is the borrower's effective tax rate, (ii) property taxes, (iii) maintenance and renovation fees, and (iv) home owner insurance. Opportunity costs are tracked for the recurring costs.

The mortgage payments $P_{0 \to T}$ are computed as

$$P_{0 \to T} \;=\; d_0 + l_0 - co_0 + \frac{\sum_{i=1}^{T} rec\ costs_i + \sum_{i=1}^{T} opp\ costs_i + (B_T + selling\ costs_T)}{(1 + r_{inv}^a)^T}, \tag{15}$$

where d_0 is the downpayment on the house at time 0, l_0 are lender points and fees, co_0 is the cash taken out, if any, B_T is the remaining principal on the mortgage at time T, and r_{inv}^a is the annualized tax-adjusted rate of return on investments. Recurring and opportunity costs are defined as above, except that they are tailored to the loan taken out now rather than the existing one. If the home is sold at time T, *selling costs$_T$* includes the closing costs associated with selling the house and the capital gains taxes, if any, on profits from the sale.

The mortgage payments $\widetilde{P}_{m \to T}$ are computed in a similar fashion as $P_{0 \to T}$ in Equation (15), but all the inputs are now for the loan taken out at time m. In particular, $d_0 + l_0 - co_0$ is replaced by

$$\frac{\widetilde{d}_m + \widetilde{l}_m - \widetilde{co}_{m,inv}}{1 + r_{inv}} - \frac{\widetilde{co}_{m,debt}}{1 + r_{debt}},$$

where $\widetilde{d}_m$ is the time-m downpayment on the house, $\widetilde{l}_m$ are the time-m discount points and lender fees, and $\widetilde{co}_{m,inv}$ is the portion of the cash taken out at time m used to pay broker fees and lender points and fees. The remaining portion of the cash taken out, $\widetilde{co}_{m,debt}$, is used to pay off high-interest-rate debt, such as credit card debt. The rate r_{debt} is the m-month interest rate on debt. When computing $\widetilde{P}_{m \to T}$, $\sum_{i=1}^{T}$ in Equation (15) is replaced by $\sum_{i=m+1}^{T}$.

The borrower expects the initial rate and rate margin, if any, for the loan taken out later to be the same as those for the loan taken out today, and that YSP, discount points and lender charges, when measured as a percentage of loan amount, remain the same as well.[14] Empirical measures for $\widetilde{P}_{0 \to m}^{ex\,mtg}$, $P_{0 \to T}$ and $\widetilde{P}_{m \to T}$ are obtained by evaluating Equations (14) and (15) using the inputs listed in Table 3.

For 30-year fixed mortgages, we set T equal to six years.[15] For 2/28 loans, we assume that

---

[14]In robustness checks, we allow the expected rate change to be different from zero (see Section 7).

[15]The choice of six years is based on a report by Credit Sesame, which states that during our sample period, the median tenure of a home seller was six years. The report is available at www.creditsesame.com/blog/how-long-are-americans-staying-in-their-homes/, and is based on data provided by the National Association of Realtors. The National Association of Home Builders (www.nahb.org) estimates the time until half of buyers move from their home to be ten years. We verify that our results are robust to setting T equal to ten years.

the borrower's horizon is the end of the teaser period and hence set T to two years. Results are presented for m equal to one month. Robustness checks are performed in Section 7.

3.3   Choice of conditioning variables

When estimating F_f, and later π and F_ν, a subset of the variables in Tables 1 and 2 serve as conditioning variables X. Two of the conditioning variables—loan purpose and product type—are used to stratify the data. The others form the state vector x. Our choice of the variables in x is guided by the model framework in Sections 1 and 2. It postulates that fees depend not only on standard loan and borrower characteristics, such as the size of the loan and the risk profile of the borrower, but also on borrowers' level of informedness about fees and the costs they associates with a delay in taking out the loan.

To proxy for the level of informedness, consider a non-expert borrower who is about to decide whether or not to consult with a friend. Since it is not in the broker's interest that the borrower consults with a friend, the broker may have an incentive to convince the borrower that he will not be able to locate a friend or that it would be too costly for him to find a friend, or that even if he found a friend, he would not be able to originate at the same fee.

While the borrower's own level of financial sophistication may have an impact on the likelihood that the broker is successful in keeping the borrower from becoming more informed, we do not observe it. Instead, we use the fraction of neighbors with a bachelor's degree as a predictor for the borrower's level of sophistication, and include it in the state vector x.[16] Moreover, independent of the borrower's own level of educational achievement, he may benefit from the advice of educated neighbors to become more informed about fees.

A borrower's private cost of waiting is likely to depend on many of the same inputs as the benchmark cost of waiting, b_k. For refinance loans, b_k depends mainly on the age, rate and balance of the existing loan, the size, rate, rate margin and discount points of the new loan,

---

[16]While monthly household income may be another indicator of borrower sophistication, it is highly correlated with the borrower's risk profile.

expectations about future rate changes, and how any cash taken out is used. We therefore include the loan amount, initial rate, rate margin and discount points of the new loan in x, and allow for yearly fixed effects.[17] The latter may proxy for the temporal variation in the age of the existing loan, the new-to-old-rate ratio, the amount of cash taken out, the interest rate on debt if cash taken out is used to pay off debt, and expected future rate changes.

For purchase loans, b_k depends mainly on the size of the loan, the initial rate, rate margin and discount points, expected future rate changes, the house price or LTV ratio, expected house price appreciation, property taxes and the price-to-rent ratio. Rather than add the LTV ratio and each of the socioeconomic variables to x, which may introduce endogeneity problems, for purchase loans we include b_k itself.

We also control for YSP since it may partially offset the fees paid directly to the broker. We condition on the level of documentation and whether or not the loan has a prepay penalty, so that marginal changes in the rate can be associated with changes in the borrower's risk profile. Location dummies may proxy for geographic differences in broker costs or broker competition.[18] In summary, we set $X = (\text{loan purpose, product type}, x)$, where

$$
\begin{aligned}
x \;=\; & (\text{YSP}, g(\text{Loan amount}), \text{Low documentation, Prepay penalty, Initial rate, Rate} \\
& \text{margin, Discount points, Education, Location, Year, } b_k \text{ (purchase loans)}). \quad (16)
\end{aligned}
$$

The variables in x are defined as in Table B.1. The function $g(\cdot)$ captures the relationship between loan size and fees shown in Figure 1: Dollar fees increase as loan size increases, but percentage fees decrease as loan size increases.

A number of scenarios for x that reflect many of our data's features are listed in Table 4. These scenarios are used to present our empirical results. A second set of scenarios, which

---

[17]In addition to discount points, the lender may charge upfront fees. Since much of the variation in lender fees can be explained by variation in the loan amount, we do not include these fees in the state vector.

[18]As a robustness check, we additionally include the broker competition variable defined in Table B.1 in the state vector and verify that our findings remain nearly unchanged.

is a mirror image of the one below, is shown in Table B.2 and is used for robustness checks.

[Table 4 about here.]

Next we describe our approach to estimating the broker fee distribution F_f as a function of X. Given F_f, we identify the distributions of broker informedness and valuations, following the approach outlined in Section 2.

4. Estimating the Fee Distribution using Non-Crossing Quantile Regressions

The crucial first step in our empirical analysis is the estimation of the broker fee distribution F_f as a function of conditioning variables X. A common approach to estimating conditional distributions is to run quantile regressions for many different quantiles. For each stratum of the data, the qth quantile of the conditional fee distribution is given by $\beta_0(q) + x\beta(q)$, defined via $Prob(f \leq \beta_0(q) + x\beta(q)|X) = q$. Keep in mind that there is a different set of $\beta_0(q)$ and $\beta(q)$ coefficients for each stratum.

A classical quantile regression analysis will estimates the coefficients $\beta_0(q)$ and $\beta(q)$ separately for each quantile level q. The resulting regression functions may cross in finite samples, meaning that the associated conditional quantile curves are not necessarily monotonically increasing in q. This problem is illustrated in Figure 2, which shows estimated conditional quantile curves for hybrid refinance loans. The figure shows that there are specifications for x where $\beta_0(q) + x\beta(q)$ is monotonically increasing in q, and specifications where it is not.

[Figure 2 about here.]

Non-monotonicity of F_f is especially problematic since it implies non-monotonicity of F_ν per Equation (6), and effectively prevents us from obtaining a well-defined empirical borrower valuation distribution. We therefore estimate the quantiles of the conditional fee distribution simultaneously under a non-crossing constraint. This technique is referred to as non-crossing

quantile regression, and, to the best of our knowledge, is new to the finance and real estate literature. The underlying theory is described in Appendix C, and our implementation is based on the R code developed by Bondell, Reich, and Wang (2010).

We perform non-crossing quantile regressions for $q \in \{0.025, 0.1, 0.2, \ldots, 0.9, 0.975\}$. Since fees are often quoted relative to the size of the loan, we set the dependent variable equal to broker fees expressed as a percentage of loan amount. We stratify the data by loan purpose and product type, and specify the state vector as in Equation (16), with $g(\text{Loan amount}) = 100 \log(\text{Loan amount})/\text{Loan amount}$. The results are summarized in Table 5, and additional details are provided in Tables D.1 through D.4 in the appendix.

[Table 5 about here.]

We find that a marginal increase in yield spread premia is only partially offset by a decrease in broker fees, consistent with the findings in Woodward and Hall (2010). As anticipated, the loading on log loan amount is positive, meaning that percentage fees tend to decrease as loan size increases. For example, a marginal increase in the size of the loan from \$100K to \$200K is associated with a decrease in median percentage fees by 85 and 80 basis points for hybrid and fixed refinance loans, and by 53 and 41 basis points for hybrid and fixed purchase loans.[19]

For each stratum of the data, the percentage of educated neighbors has a significant impact on median percentage broker fees. A marginal increase in the percentage of neighbors with a bachelors degree of 10% is associated with a decrease in median percentage fees of 11 to 18 basis points, depending on the loan purpose and product type. For a loan size of \$200K, this translates into a decrease of \$220 to \$360 per loan.

All else the same, loans originated in California tend to be more expensive than loans originated in Florida, but, at least for purchases, less expensive than those originated in

---

[19]For hybrid refinance loans, the impact of a marginal increase in loan amount from \$100K to \$200K is computed as $0.436 \left(100 \log(200)/200 - 100 \log(100)/100\right) = 0.85$. Similarly for the other strata of the data. For purchase loans, the quoted impact is for the specification in columns five and six in Table 5.

Texas. For purchase loans, the benchmark cost of waiting is a significant predictor of median percentage fees. A marginal increase in b_k, as a percentage of loan amount, by 1% is associated with an increase in percentage fees by 3 basis points for hybrid loans and 12 basis points for fixed loans. For the remaining variables, results are more mixed.

We convert the estimated conditional percentage fee distributions into conditional dollar fee distributions F_f. Since only a small fraction of borrowers are fully informed, we set $\underline{f}$ equal to the lowest quantile estimate for F_f. We determine $\overline{\nu}$ so that the sum of $\overline{\nu}$, YSP, discount points and lender fees equals 8% of the loan amount. A loan with total origination charges in excess of 8% is referred to as a "Section 32 mortgage," and is banned from balloon payments, negative amortization, and most prepayment penalties, among other features, according to the Home Ownership and Equity Protection Act of 1994.[20] There are almost no Section 32 mortgages in our sample.

Figure 3 shows the quantile curve F_f for hybrid refinance loans, using piecewise cubic Hermite interpolation to connect the estimated quantiles.[21] The state vector x is equal to the reference state vector in Table 4. For different exogenous choices of π_2, the figure shows the implied distribution of borrower valuations. It reveals that as the fraction of minimally informed borrowers increases, the implied distribution of borrower valuations shifts to the left, meaning towards smaller values. This is intuitive since the given fee distribution is a mixture of $\underline{f}$, $\min\{f_1, \nu\}$ and ν. As more and more weight is shifted towards ν, the implied distribution for ν shifts towards smaller values. As a result, higher values π_2 are associated with smaller median valuations. By pinpointing the median valuation, we are able to identify π_2, and hence π and F_ν. This is the focus of the next section.

[Figure 3 about here.]

---

[20]The cutoff is actually the larger of some small flat fee and 8%. But for all the loans in our sample, 8% is the binding constraint.

[21]Piecewise cubic Hermite interpolation preserves the shape of the data and respects monotonicity.

5. The Implied Distributions of Borrower Informedness and Valuations

For a given specification of X, we use the inverse CDF method (see Glasserman (2004)) to draw independent random samples from the estimated fee distribution F_f. Monte Carlo estimates of expected broker fees are reported in Table 6. We first discuss the results for the reference state vector in Table 4, and then consider other scenarios. For purchase loans, we rely on the results in columns five and six of Table 5.

5.1 Results for the reference state vector

Average fees per loan are higher for refinance loans than for purchase loans, and for hybrid loans than for fixed loans: They are $4.4K and $3.9K for hybrid and fixed refinance loans, and $3.4K and $3.0K for hybrid and fixed purchase loans.

[Table 6 about here.]

According to Equation (12), benchmark future fees $b_{\widetilde{f}}$ are computed as discounted aggregate fees for a loan with characteristics $\widetilde{X}_m$. For a set of characteristics $\widetilde{X}_m$, aggregate fees are defined as average fees for similar loan transactions in neighborhoods with high average fees. For the state vector defined in Equation (16), neighborhood effects are captured through the Education variable. Since a higher fraction of educated neighbors is associated with lower average broker fees (see Tables D.1 through D.4), and since the fraction of educated neighbors rarely falls below 5%, we compute aggregate fees as average fees for loans with similar characteristics originated in zip codes where 5% of the population has a B.A. degree. For hybrid and fixed refinance loans, benchmark future fees are $4.7K and $4.2K, respectively. For hybrid and fixed purchase loans, they are $3.6K and $3.3K.

The benchmark cost of waiting, b_k, is computed as in Section 3.2, using the inputs listed in Tables 3. For refinance loans, b_k is between $0.2K and $0.3K per loan. It is fairly modest in size since by choosing the outside option, the borrower only gives up receiving cash and

23

a lower rate m months earlier. In comparison, a borrower that wants to purchase a house would forego m months of house price appreciation if he delays, and may have to pay a relatively high rent. As a result, b_k is much larger for purchase loans at \$2.9K to \$3.1K.

Equipped with estimates for F_f and b_ν, we interpret b_ν as the median borrower valuation and compute the fraction of minimally informed borrowers π_2 as in Equation (9).[22] Our estimate for π_2 is 55% for hybrid refinance loans and 54% for fixed refinance loans, implying that more than half of the borrowers are minimally informed. The associated distributions of borrower informedness are $\pi = (0.01, 0.43, 0.55)$ for hybrid refinance loans and $\pi = (0.01, 0.44, 0.54)$ for fixed refinance loans. For purchase loans, the benchmark cost of waiting is substantially higher when measured relative to fees, which implies that more non-expert borrowers decide to consult with a friend. The fraction of minimally informed borrowers is only about 4%, both for hybrid and fixed purchase loans, with partially informed borrowers making up much of the rest of the borrower population.

Given estimates for F_f and π, we compute F_ν according to Equation (6) and then sample from F_ν using the inverse CDF method. This allows us to compute expected fees conditional on the level of informedness. They range from the broker's reservation value $\underline{f}$ for fully informed borrowers, to $E(\min\{f_1, \nu\}|\text{non-expert borrower})$ for partially informed borrowers, to $E(\nu|\text{non-expert borrower})$ for minimally informed borrowers. Table 6 reports estimates for the broker's reservation value of \$0.9K and \$0.8K for hybrid and fixed refinance loans, and \$0.5K and \$0.2K for hybrid and fixed purchase loans.[23] All else the same, broker reservation values are estimated to be higher for refinance than for purchase loans.

Expected borrower valuations—which are equal to expected fees for minimally informed borrowers—range from \$4.6K to \$5.0K for refinance loans, and from \$6.8K to \$7.0K for

---

[22]Even though we can compute the distribution of borrower informedness, π, we cannot observe an individual borrowers' level of informedness. In Appendix D we show, however, that once we observe the fee that a borrower has paid, we can update the likelihood of him being fully, partially or minimally informed.

[23]The broker's cost of originating the loan is equal to the sum of the broker's reservation value and YSP (see Equation (1)). For the reference state vector, YSP is \$2,000 per loan.

purchase loans. Expected fees for partially informed borrowers fall between those for fully and minimally informed borrowers. They range between \$3.2K and \$3.6K for refinance loans and between \$2.9K and \$3.3K for purchase loans.

The fraction $\gamma = (\overline{f} - \underline{f})/(\nu - \underline{f})$ of the overall surplus from the mortgage that goes to the broker is zero when the borrower is fully informed and one when he is minimally informed. For partially informed borrowers, the expected value of γ ranges between 71% and 72% for refinance loans, and 46% and 48% for purchase loans. Our model implies that for a given fee distribution, $E(\gamma|\text{partially informed borrower})$ decreases as the level of informedness in the borrower population increases.

5.2   Size and education effects

For each stratum of the data, average fees are higher for larger loans. We refer to this as the size effect on broker fees. According to Equation (5), the size effect on fees translates into a size effect on borrower informedness or a size effect on borrower valuations, or both. The large loan, reference and small loan scenarios in Tables 6 and B.3 reveal that both π and F_ν change with the size of the loan. For each stratum of the data, the fraction of minimally informed borrowers is smaller for larger loans, whereas expected borrower valuations are higher. The reason is that the benchmark cost of waiting nearly scales with the loan amount whereas percentage fees decrease as loan size increases.[24] As a result, benchmark borrower valuations not only increase as the loan size increases, but also increase relative to fees, which implies that the fraction of minimally informed borrowers decreases (see Figure 3).

To give a specific example, consider hybrid refinance loans. Expected fees are \$5.6K for large loans, \$4.4K for reference loans and \$3.0K for small loans. The implied fraction of minimally informed borrowers is 49%, 55% and 70% for large, reference and small loans, and

---

[24]For purchase loans, percentage benchmark costs of waiting are constant across loan size. For refinance loans, we observe a small increase in the percentage cost of waiting with the size of the loan, mainly because for larger loans a smaller fraction of the cash taken out is needed to pay broker fees and lender points and fees, meaning a larger fraction of it can be used to pay off or avoid high-interest-rate debt.

the expected borrower valuation is \$6.8K, \$5.0K and \$3.2K. The impact of the size of the loan on borrower valuations is visualized in the left panel of Figure 4. The expected fraction of the overall surplus from the loan that goes to the broker is smaller for larger loans: $E(\gamma|$ partially informed borrower) equals 71%, 72% and 76% for large, reference and small loans.

[Figure 4 about here.]

As mentioned before, average fees are also higher for loans originated in more educated neighborhoods. We refer to this as the education effect on broker fees. Interestingly, the education effect translates primarily into an effect on borrower informedness—the population of borrowers in neighborhoods with a higher fraction of college graduates is more informed about broker fees—and much less so into an effect on borrower valuations. This is documented in the right panel of Figure 4.

For hybrid refinance loans, expected fees are \$4.0K, \$4.4K and \$4.7K for the educated neighbors, reference and uneducated neighbors scenarios in Tables 6 and B.3. The implied fraction of minimally informed borrowers is 39%, 55% and 77%, respectively, whereas the expected borrower valuation is fairly stable around \$5.0K–\$5.1K. The expected fraction of the overall surplus from the loan that goes to the broker is smaller for loans originated in more educated neighborhoods: $E(\gamma|$ partially informed borrower) is 69%, 72% and 75% for the educated neighbors, reference and uneducated neighbors scenarios.

Benchmark borrower valuations do not depend on the fraction of neighbors with a B.A. degree.[25] According to Equation (9), variations in the broker fee distribution that are associated with variations in education therefore translate into variations in π. The reason that the valuation distribution for a more educated neighborhood is similar to that for a less educated neighborhood is that their medians are the same and their shapes, which are derived from the shapes of the associated fee distributions, are rather similar.

---

[25]In Section 7, we discuss the implications of a neighborhood-specific specification of b_ν.

5.3 House price appreciation and price-to-rent ratio effects

For refinance loans, the benchmark cost of waiting is completely determined by the variables in the state vector x defined in Equation (16). This is because we observe many of the inputs for b_k—such as the percentage of cash taken out, the age of the existing loan, the new-to-old-rate ratio and the interest rate on debt—only at an annual frequency, and therefore can control for them by including yearly dummies in x.[26]

For purchase loans, however, there is significant variation in b_k, even after controlling for all other variables in the state vector. All else the same, higher expected house price appreciation (HPA) is associated with a higher benchmark cost of waiting, as is a lower price-to-rent ratio.[27] For hybrid purchase loans, b_k is \$4.8K for the high-HPA scenario, \$2.9K for the reference scenario and \$0.9K for the low-HPA scenario. The benchmark cost of waiting is \$3.6K, \$2.9K and \$2.6K for the low, reference and high price-to-rent ratio scenario.

While the results for purchase loans in Tables 5, D.3 and D.4 suggest that b_k is a statistically significant predictor of F_f, our estimates in Tables 6 and B.3 reveal that the impact is modest in economic terms: Expected broker fees increase by less than \$400 per loan as expected house price appreciation climbs from 5% to 25%, and by less than \$150 as price-to-rent ratios drop from 20 to 10.

The finding that an increase in b_k as a result of higher house price appreciation or lower price-to-rent ratios is not fully reflected in the fee distribution suggests that high house price appreciation or low price-to-rent ratios motivate borrowers to become more informed about fees. Indeed, for hybrid purchase loans the fraction of minimally informed borrowers is 1%

---

[26]The highest frequency at which these variables are observed is quarterly, and for a larger sample of loans it may be sensible to control for quarterly instead of yearly fixed effects. We caution against separately including each of these variables in the state vector, however, since it would introduce a host of endogeneity problems. For refinance loans, the benchmark cost of waiting does not depend on expected house price appreciation. This is by design as we hold the size of the loan taken out at time m fixed. Different specifications of the loan taken out later would allow for house price appreciation to have an impact on b_k.

[27]Similarly, higher property taxes are associated with lower benchmark costs of waiting. Property taxes and house price appreciation are closely tied together, with a correlation of about -50% in our data.

for the high-HPA scenario compared to 29% for the low-HPA scenario, and 2% when price-to-rent ratios are low compared to 6% when price-to-rent ratios are high. In summary, house price appreciation and price-to-rent ratios have a pronounced effect on the distributions of borrower informedness and valuations, and a more modest effect on broker fees.

5.4 How non-expert borrowers decide whether or not to consult with a friend

A precise specification of how non-expert borrowers decide whether or not to consult with a friend is not necessary when solving our model. The only related assumption that we have made thus far is that non-expert borrowers have to decide whether or not to consult with a friend prior to observing their private valuations. Nevertheless, it may be beneficial to interpret our findings under a set of assumptions that describe non-experts' decision process.

To motivate our analysis, recall the size effect on borrower informedness. We want to understand whether the higher level of borrower informedness for larger loans is due to (i) larger borrowers assigning a greater likelihood to being able to lower their fees by consulting with a friend, or (ii) larger borrowers making different decisions based on the same likelihood.

To that end, we consider the case where non-expert borrowers observe the benchmark valuation b_ν prior to making their decision, together with an estimate—based on aggregate fee data—of the likelihood that a friend's fees do not exceed this benchmark, $F_f^*(b_\nu)$. Given $F_f^*(b_\nu)$, a fraction $\pi_2/(\pi_1 + \pi_2)$ of non-expert borrowers decides against consulting with a friend, perhaps because they perceive efforts to consult with a friend as too costly or because they can be dissuaded by the broker from obtaining additional information about fees.

As before, we condition on a set of loan and borrower characteristics and interpret b_ν as the median valuation of non-expert borrowers. We denote the fraction of non-expert borrowers who, given $F_f^*(b_\nu)$, decide to become partially informed by Δ: $\Delta = \pi_1/(\pi_1 + \pi_2)$. A straightforward manipulation of Equations (7) and (9) yields $\Delta = h(q)$, where $q = F_f(b_\nu)$,

$\underline{q} = F_f(\underline{f})$ is constant at 0.025 (see Section 4), and

$$h(q) \quad = \quad \frac{2q - \underline{q} - 1}{q(1 + \underline{q}) - 2\underline{q}}.$$

We define $q^* = F_f^*(b_\nu)$ and let $d = q - q^*$ denote the difference between the true probability that a friend's fees do not exceed b_ν and the probability as perceived by non-expert borrowers prior to making their decision. With this notation, the fraction of non-expert borrowers that decide to become partially informed is $\Delta = h(q^* + d)$. Holding d the same, the function h is increasing in q^*. This implies that if non-expert borrowers perceive benchmark valuations to be high relative to fees, then a larger fraction of them decide to consult with a friend.

For a given loan purpose–product type combination, consider two different specifications of the state vector: The reference scenario (A) and the large loan scenario (B) in Table 4. From our discussion of the size effect on borrower informedness we know that the fraction of non-expert borrowers that decide to become partially informed is higher for larger loans: $\Delta_{(B)} > \Delta_{(A)}$, where (A) and (B) identify the underlying state vector. Specifically, we have

$$\Delta_{(B)} - \Delta_{(A)} \quad = \quad \underbrace{h(q_{(B)}^* + d_{(B)}) - h(q_{(B)}^* + d_{(A)})}_{\substack{\Delta_{(B)} - \Delta_1\text{: change in } \Delta \text{ due to change in} \\ \text{borrowers' decision making given } F_f^*(b_\nu)}} + \underbrace{h(q_{(B)}^* + d_{(A)}) - h(q_{(A)}^* + d_{(A)})}_{\substack{\Delta_1 - \Delta_{(A)}\text{: change in } \Delta \text{ due to change} \\ \text{in perceived likelihood } F_f^*(b_\nu)}}. \quad (17)$$

The second term on the right-hand side of Equation (17) is the increase in Δ that would result from larger borrowers assigning a higher probability to a friend's fees not exceeding benchmark fees, but otherwise making similar decisions as reference borrowers. What we mean by the latter is that the relationship between realized and perceived fees, as measured by $d = F_f(b_\nu) - F_f^*(b_\nu)$, is assumed to be the same for large borrowers as it is for reference borrowers. The first term in Equation (17) is the increase in Δ due to larger borrowers being more inclined to gather additional information about fees than reference borrowers for a given perceived likelihood of being able to lower their fees by consulting with a friend.

29

Consistent with the measurement of benchmark future fees in Section 5, we set F_f^* equal to the conditional fee distribution for similar loans in zip codes where 5% of the population has a B.A. degree. Table 6 reports that much of the increase from $\Delta_{(A)}$ to $\Delta_{(B)}$ is due to benchmark valuations being higher relative to perceived fees for larger loans. For hybrid refinance loans, $\Delta_1 - \Delta_{(A)}$ accounts for 5% of the 6% difference between $\Delta_{(B)}$ and $\Delta_{(A)}$.

If we replace the large loan scenario (B) in Equation (17) with the educated neighbors scenario (C), we obtain a different result. Since neither b_ν nor F_f^* depend on the fraction of neighbors with a B.A. degree, the increase in $\Delta_{(C)}$ over $\Delta_{(A)}$ is entirely due to borrowers with more educated neighbors being more motivated to gather additional information about fees. This finding is consistent with borrowers benefiting from the advice of educated neighbors to become more informed about fees by consulting with a friend.

Table 6 also reports that the increase in Δ from higher expected house price appreciation or lower price-to-rent ratios is associated with a perceived increase of benchmark valuations relative to fees.

6. The Impact of Borrower Informedness on Broker Fees

We quantify the increase in the broker's fees when doing business with a less informed borrower, and the reduction in fees if a partially informed borrower were to consult one extra friend or if the borrower population suddenly became more informed.

6.1 Increase in broker's fees from doing business with a less informed borrower

The results in Tables 6 allow us to quantify the expected increase in the broker's fees when doing business with a less informed borrower. For refinance loans, expected fees are $2.4K–$2.7K higher when the loan is originated from a partially rather than a fully informed borrower, and $1.4K higher when the loan is originated from a minimally rather than a partially informed borrower. For purchase loans, expected fees are $2.7K–$2.8K higher

when the borrower is partially rather than fully informed, and \$3.8K–\$3.9K higher when he is minimally rather than a partially informed.

The increase in expected fees per loan is greater for larger loans, and for loans originated more educated neighborhoods. For hybrid refinance loans, the increase in fees from doing business with a partially rather than a fully informed borrower is \$3.7K for large and \$1.7K for small loans, and \$2.8K versus \$2.7K for borrowers with a high versus low fraction of educated neighbors. The increase in fees from doing business with an minimally rather than a partially informed borrower is \$2.1K for large loans and \$0.7K for small loans, and \$1.7K compared to \$1.2K when the fraction of educated neighbors is high rather than low. Similarly, the increase in fees from doing business with a less informed borrower is higher in areas with high house price appreciation or low price-to-rent ratios.

Our theoretical model implies that for a given fee distribution, a broker's benefit from interacting with a less informed borrower is higher the higher the level of informedness in the borrower population. This is because a decrease in π_2 is associated with a shift of the borrower valuation distribution towards higher values, as shown in Figure 3.

6.2 Reduction in broker fees and γ if a partially informed borrower consults one extra friend

Given F_f and F_ν, we quantify how the fees that a broker extracts from a partially informed borrower would change in expectation if that borrower were to consult one extra friend. In particular, we compute the difference between the expected value of $\min\{f_1, f_2, \nu\}$ and the expected value of $\min\{f_1, \nu\}$, where f_1 and f_2 are independent draws from F_f that are independent of ν. This difference is clearly negative. Its absolute value is reported in Table 7. For the reference state vector, a partially informed borrower would save an average of about \$0.6K per refinance loan or \$0.8K–\$0.9K per purchase loan if he were to consult with two instead of just one friend.

[Table 7 about here.]

We also compute how the expected fraction of the overall surplus from the mortgage that goes to the broker would change if a partially informed borrower were to consult with one extra friend. In particular, we compute the difference between the expected values of $(\min\{f_1, f_2, \nu\} - \underline{f})/(\nu - \underline{f})$ and $(\min\{f_1, \nu\} - \underline{f})/(\nu - \underline{f})$. Table 7 shows that for the reference state vector, $E(\gamma|$ partially informed borrower$)$ would drop by 13–14 percentage points.

6.3 Reduction in broker fees if borrower population becomes more informed

Given F_f and F_ν, we can also quantify the changes in expected broker fees if the borrower population suddenly became more informed. Holding the implied borrower valuation distribution fixed, we compute

$$E(f|\pi(\theta)) \;=\; \pi_0(\theta)\,\underline{f} + \pi_1(\theta)\,E(\min\{f_1, \nu\}|\text{non-expert}) + \pi_2(\theta)\,E(\nu|\text{non-expert}) \quad (18)$$

for a new distribution of borrower informedness $\pi(\theta) = (\pi_0, \pi_1 + \theta\pi_2, (1 - \theta)\pi_2)$. Clearly, Equation (18) is no longer consistent with the equilibrium condition in Equation (5). We simply use it to answer the question "What would happen to fees, temporarily, if there would be a sudden change in the distribution of borrower informedness?".

Table 8 reports results for $\theta \in \{0, 0.5, 1\}$. For $\theta = 0$, the distribution of borrower informedness remains unchanged. We use $\theta = 0.5$ to describe the scenario where 50% of minimally informed borrowers suddenly become partially informed, and $\theta = 1$ to describe the scenario where all minimally borrowers suddenly become partially informed. For the reference state vector, expected fees for refinance loans decrease by about \$0.4K per loan if $\theta = 0.5$, and \$0.8K if $\theta = 1$. For purchase loan, expected fees decrease by about \$0.1K for $\theta = 0.5$ and \$0.1K–\$0.2K for $\theta = 1$.

[Table 8 about here.]

7. Extensions and Robustness

We describe extensions to our model framework and perform robustness checks.

7.1 Model extension

We extend our model of how borrowers form their reservation value for the fees, $\bar{f}$, by lifting the assumption that a non-expert borrower's valuation for the loan is independent of his level of informedness. This allows for situations where the median valuation is different for partially and minimally informed borrowers. Consider a scenario where $b_\nu = b_k + b_{\tilde{f}}$ defined in Sections 2 and 3 is the median valuation of minimally informed borrowers. A partially informed borrower with median valuation, however, learns about average conditional broker fees in his neighborhood while consulting with a friend. He uses that information to update his expectation for discounted future fees from $b_{\tilde{f}}$—which is based on average fees for similar loans in neighborhoods with high average fees—to a neighborhood-specific estimate $b_{1,\tilde{f}}$. This implies that his updated valuation $b_{1,\nu} = b_k + b_{1,\tilde{f}}$ is potentially lower than the median valuation of minimally informed borrowers, $b_{2,\nu} = b_\nu$.

Let $F_{1,\nu}$ be the cumulative distribution function of partially informed borrowers' valuations and $F_{2,\nu}$ be that of minimally informed borrowers' valuations. Mimicking the approach in Section 1, we show that F_f, π, $F_{1,\nu}$ and $F_{2,\nu}$ are linked in equilibrium:

$$
\begin{aligned}
F_f(a) &= \pi_0 + \pi_1 Prob(\min\{f_1, \nu\} \le a | \text{partially informed borrower}) + \pi_2 F_{2,\nu}(a) \\
&= \pi_0 + \pi_1 \{1 - (1 - F_f(a))(1 - F_{1,\nu}(a))\} + \pi_2 F_{2,\nu}(a),
\end{aligned}
\tag{19}
$$

where $a \in [\underline{f}, \bar{\nu}]$. Equation (19) can be rewritten as

$$
F_f(a)(1 - \widetilde{\kappa}_1(a)) + \widetilde{\kappa}_1(a) - \widetilde{\kappa}_0(a) = 0,
\tag{20}
$$

where $a \in [\underline{f}, \bar{\nu}]$, $\widetilde{\kappa}_1(a) = \pi_1(1 - F_{1,\nu}(a))$ and $\widetilde{\kappa}_0(a) = 1 - \pi_2(1 - F_{2,\nu}(a))$.

In Appendix A, we prove that there is a unique equilibrium to the extended model, and that $F_f \geq \min\{F_{1,\nu}, F_{2,\nu}\}$.

Proposition 4 *Let $F_{1,\nu}$ and $F_{2,\nu}$ be some distribution functions with support on $[\underline{f}, \overline{\nu}]$. For $\pi_0, \pi_1, \pi_2 > 0$, the equilibrium distribution of broker fees is well-defined by Equation (20), has support on $[\underline{f}, \overline{\nu}]$ and is unique.*

Proposition 5 *For $\pi_0, \pi_1, \pi_2 > 0$, the equilibrium distribution of broker fees satisfies $F_f(a) \geq \min\{F_{1,\nu}(a), F_{2,\nu}(a)\}$, for all $a \in [\underline{f}, \overline{\nu}]$.*

In what follows, we show that we can again exploit the equilibrium condition (20) to estimate the implied distributions π, $F_{1,\nu}$ and $F_{2,\nu}$, subject to an over-identifying restriction that links the shape of $F_{1,\nu}$ to that of $F_{2,\nu}$. Specifically, we impose $F_{2,\nu}(\underline{f}) = 0$ and $F_{1,\nu} = w_\alpha(F_{2,\nu})$, where $w_\alpha(\cdot)$ has support on $[0, 1]$ and is defined as

$$w_\alpha(x) = \begin{cases} (1 - \alpha^x)/(1 - \alpha), & \text{if } \alpha > 0 \text{ and } \alpha \neq 1 \\ x, & \text{if } \alpha = 1. \end{cases} \tag{21}$$

Lemma 2 in Appendix A proves that if $F_{2,\nu}$ is a well-defined distribution function, so is $F_{1,\nu}$.

The relationship between $F_{1,\nu}$ and $F_{2,\nu}$ as a function of α is displayed in Figure B.1 in Appendix D. Different choices for α capture different relations between $F_{1,\nu}$ and $F_{2,\nu}$. For $\alpha < 1$, $F_{1,\nu} \geq F_{2,\nu}$, whereas for $\alpha > 1$, $F_{1,\nu} \leq F_{2,\nu}$. Since we are interested in scenarios where the median valuation of a partially informed borrowers is smaller than that of a minimally informed borrower, we expect α to be less than one.

Combining $F_{1,\nu} = w_\alpha(F_{2,\nu})$ with Equation (19) yields

$$F_f(a) = \pi_0 + \pi_1 \{1 - (1 - w_\alpha(F_{2,\nu}(a)))(1 - F_f(a))\} + \pi_2 F_{2,\nu}(a), \tag{22}$$

which allows us to compute $F_{2,\nu}(a)$ as a function of $F_f(a)$ and π, for $a \in [\underline{f}, \overline{\nu}]$. The

next proposition states that the same conditions on the model parameters that ensure the existence of F_ν in Section 2 are sufficient to ensure that $F_{2,\nu}$ is well-defined:

Proposition 6 *Let F_f be some distribution function with support on $[\underline{f}, \overline{\nu}]$. For $\pi_0, \pi_1, \pi_2 > 0$ and $F_f(\underline{f}) \geq \pi_0/(1-\pi_1)$, $F_{2,\nu}$ given by Equation (22) is a well-defined distribution function with support on $[\underline{f}, \overline{\nu}]$.*

Based on estimates for F_f, we identify π, $F_{1,\nu}$ and $F_{2,\nu}$ as follows. First, we use $\pi_0 + \pi_1 + \pi_2 = 1$ and $F_{2,\nu}(\underline{f}) = 0$ in Equation (22) to express π_0 and π_1 as a function of π_2. We choose a value for α and evaluate Equation (22) at $a = b_{2,\nu}$ to solve for π_2, using the assumption $F_{2,\nu}(b_{2,\nu}) = 0.5$. Second, using the implied π, we compute $F_{2,\nu}$ and $F_{1,\nu}$. Third, we repeat the first two steps for different values of α until the median of $F_{1,\nu}$ matches $b_{1,\nu}$.[28]

Our estimates are summarized in Table 9. For the reference state vector, the implied value for α ranges from 0.51 for hybrid refinance loans to 0.81 for fixed purchase loans. The fraction of minimally informed borrowers is slightly higher than its counterpart in Table 6. It is nearly 60% for refinance loans, and 4% for purchase loans.

[Table 9 about here.]

Table 9 also reports expected fees conditional on the borrower's level of informedness.[29] As anticipated, expected fees paid by partially informed borrowers are somewhat lower than their counterparts in Table 6, reflecting the fact that median valuations are no longer based on upward biased benchmarks for future fees. This translates into a lower expected fraction of the overall surplus from the loan that goes not the broker. For the extended model, $E(\gamma|$ partially informed borrower) is 68% for hybrid refinance loans compared to 72% in Table 6, and 67% for fixed refinance loans compared to 71%. The adjustments in $E(\gamma|$

---

[28]The search for α converges since lower values for α are associated with lower median valuations for partially informed borrowers.

[29]Since the fee distribution is the same as in Table 6, fees paid by fully informed borrowers are the same.

partially informed borrower) are less pronounced for purchase loans for whom the vast majority of borrowers are partially informed, both in the original and the extended model framework. For each stratum of the data, partially informed borrowers leave a large fraction of the overall surplus form the mortgage on the table.

Figure 5 shows the implied valuation distributions for partially and minimally informed borrowers. As expected, $F_{1,\nu}$ lies to the left of $F_{2,\nu}$. Note that the implied valuation distributions are fairly similar for partially and minimally informed borrowers. As a result, the implications of the extended model are consistent with those discussed in Sections 5 and 6.

[Figure 5 about here.]

7.2 Model generalization: Partially informed borrowers consult with N friends

Both the original and the extended model framework can be generalized by allowing partially informed borrowers to consult with $N \geq 1$ friends rather than $N = 1$ friend. A partially informed borrower who consults with N friends sets his reservation value as $\overline{f} = \min\{f_1, \ldots, f_N, \nu\}$, where $f_1, \ldots, f_N$ are independent draws from F_f, independent of ν. The generalized versions of Propositions 1 through 6 and Corollary 1 are stated and proven in Appendix A.

7.3 Robustness checks

We have implemented a number of robustness checks. Our findings are summarized below, and additional details are available upon request.

First, we consider the case where perceived average broker fees b_f in Equation (12) are equal to true average fees. Our findings remain the same qualitatively, with the exception that education now has a direct impact on borrower valuations and a more limited impact on borrower informedness. Moreover, in the setting of Section 5.4, a consistent assumption would be $d = F_f(b_\nu) - F_f^*(b_\nu) = 0$, so that the first term on the right-hand side of Equation (17) is equal to zero. This implies that changes in the fraction of non-expert borrowers

36

that decide to become partially informed are entirely due to changes in the (perceived) likelihood that a friend's fees do not exceed benchmark fees.

Second, we verify that for the state vectors specified in Tables 4 and B.2, the benchmark cost of waiting increases as the period of waiting m increases. As a result, the implied fraction of minimally informed borrowers decreases as m increases. For purchase loans, the benchmark cost of waiting is a fairly flat function of the time T for which the mortgage is kept, except as T approaches the horizon at which the capital gains tax on proceeds from selling the home starts to kick in.

Third, we re-compute benchmark valuations while allowing for changes in mortgage rates over the next month. Specifically, we use the historical mortgage rates by product type provided by HSH Associates (see www.hsh.com/mtghst.html) to obtain monthly forecasts of rate changes. Since rates are persistent over the short run, predicted future rates changes are generally very small. As a result, our results remain largely the same.

8. Conclusion

We offer three contributions. Our first contribution is to develop a potentially realistic model of the mortgage origination process that explains, for the first time, the link between the distributions of broker fees, borrower valuations and borrower informedness. We believe this is an important step forward as our framework provides consumer advocates and regulators with a tool that allows them to analyze how mortgage origination charges may change in response to *(i)* better educating borrowers about how to evaluate loan proposals and outside options, or to *(ii)* better educating borrowers about the origination process and urging them to shop from several brokers or consult with friends.

Our second contribution is to estimate the distribution of broker fees, as a function of loan and borrower characteristics, using non-crossing quantile regressions. The common approach to estimating conditional distributions is to run separate quantile regressions for

37

a large number of quantiles. The drawback of this approach is that the resulting regression functions may cross in finite samples, meaning the estimated quantile curves may not be monotonically increasing. To ensure that the estimated distribution is well defined, we implement non-crossing quantile regressions, a technique that is new to the real estate and the finance literature. Using data from, formerly, one of the largest subprime lenders, we document that conditional fees tend to be higher for larger loans, and when the fraction of neighbors with a B.A. degrees is lower.

Our third contribution is to provide estimates for the fraction of the borrower population that is fully, partially and minimally informed about broker fees, and to use these estimates to compute the implied distribution of borrower valuations. We show that in situations where borrowers perceive the cost of waiting to take out the loan to be high relative to fees, they tend to be more inclined to gather additional information about fees. This is the case for large loans, and for purchase loans when the expected house price appreciation is high or the price-to-rent ratio is low. And even if borrowers' perception of the cost of waiting and the fees are the same, those with more educated neighbors may benefit from better advice and become more informed about fees.

We quantify the fraction of the overall surplus from the mortgage that goes to the broker. While fully informed borrowers pocket the entire spread between their value for the loan and the minimum fee for which the broker is willing to originate it, minimally informed borrowers are left empty handed. Partially informed borrowers, on the other hand, split the overall surplus from the mortgage with the broker. For reference refinance loans, borrowers retain less than 30% of the overall surplus, whereas for purchase loans that fraction is higher at about 50% due to higher borrower valuations. If a partially informed borrower were to consult with one extra friend, the fraction of the overall surplus retained by the borrower would increase by more than ten percentage points.

A. Proofs

In this appendix, we prove Propositions 1 through 6 and Corollary 1 under the generalized assumption that a partially informed borrower consults $N \geq 1$ friends who are entirely honest about the broker fees f_i they paid in a similar recent mortgage transaction. He forms his reservation value as $\overline{f} = \min\{f_1, \ldots, f_N, \nu\}$, where $f_1, \ldots, f_N$ are independent draws from F_f, independent of ν. Under this generalization, the equilibrium fee distribution satisfies

$$F_f(a) + \kappa_1(a)(1 - F_f(a))^N - \kappa_0(a) = 0 \qquad (A.1)$$

for $a \in [\underline{f}, \overline{\nu}]$, where $\kappa_1(a) = \pi_1(1 - F_\nu(a))$ and $\kappa_0(a) = 1 - \pi_2(1 - F_\nu(a))$ are defined as in Section 1. Equation (A.1) can be inverted to express F_ν as a function of the equilibrium distribution F_f and the borrower informedness distribution π:

$$F_\nu(a) = 1 - \frac{1 - F_f(a)}{\pi_1(1 - F_f(a))^N + \pi_2} \qquad (A.2)$$

for $a \in [\underline{f}, \overline{\nu}]$, and $F_\nu(a) = 0$ for $a < \underline{f}$.

When computing F_ν via Equation (6), we generally impose the condition $F_\nu(\underline{f}) = 0$. Together with $\pi_0 + \pi_1 + \pi_2 = 1$, it implies

$$\pi_1 = \frac{1 - F_f(\underline{f}) - \pi_2}{(1 - F_f(\underline{f}))^N}. \qquad (A.3)$$

Lemma 1 *For $N \geq 1$ and $0 < \kappa_1 < \kappa_0 < 1$, $p(x) = x + \kappa_1(1 - x)^N - \kappa_0$ has exactly one root in $[0, 1]$, and it is in $(0, 1)$.*

Proof. A proof of Lemma 1 can be found in Woodward and Hall (2009). To make our paper self-contained, we reproduce it here. The function $p(x)$ is continuous, and $p(0) = \kappa_1 - \kappa_0 < 0$

and $p(1) = 1 - \kappa_0 > 0$. So there exists at least one root between 0 and 1. To rule out the possibility of more than one root, consider

$$p'(x) = 1 - \kappa_1 N (1 - x)^{N-1}$$

and

$$p''(x) = \begin{cases} \kappa_1 N(N-1)(1-x)^{N-2} \geq 0, & \text{if } N > 1 \\ 0, & \text{if } N = 1. \end{cases}$$

If $N = 1$ or if $N > 1$ and $\kappa_1 N \leq 1$, then $p'(x) > 0$ in $(0, 1)$, meaning there is only one root of $p(x)$ in $[0, 1]$, and it is in $(0, 1)$. If $N > 1$ and $\kappa_1 N > 1$, then $p'(x^*) = 0$ if and only if

$$x^* = 1 - (\kappa_1 N)^{-\frac{1}{N-1}}.$$

Because $p''(x)$ is non-negative, $p'(x)$ is positive for $x^* < x \leq 1$ and negative for $0 \leq x < x^*$. So again $p(x)$ has only one root in $[0, 1]$, and it is in $(0, 1)$. $\qquad\square$

Proposition 1 *Let F_ν be some distribution function with support on $[\underline{f}, \bar{\nu}]$. For $N \geq 1$, $\pi_0, \pi_1, \pi_2 > 0$ and $\pi_1 N \leq 1$, the equilibrium distribution of broker fees is well-defined by Equation (A.1), has support on $[\underline{f}, \bar{\nu}]$ and is unique.*

Proof. Equation (A.1) implies that $F_\nu(\bar{\nu}) = 1$ if and only if $F_f(\bar{\nu}) = 1$. The broker's participation constraint implies that $F_f(a) = 0$ for $a < \underline{f}$. Hence F_f has support on $[\underline{f}, \bar{\nu}]$. Let $y = F_f(a)$ and $x = F_\nu(a)$. According to Equation (A.1),

$$y + \pi_1 (1 - x)(1 - y)^N - \pi_0 - \pi_1 - \pi_2 x = 0.$$

Differentiation w.r.t. x yields $y' - \pi_1 (1 - y)^N - \pi_1 N (1 - x)(1 - y)^{N-1} y' - \pi_2 = 0$. For

40

$x \in [0,1]$ and $\pi_1 N \leq 1$, we have $\pi_1 N(1-x)(1-y)^{N-1} < 1$. (Note that for $N = 1$, $\pi_1 N < 1$. For $N > 1$ and $x = 0$, y solves Equation (5) with $\kappa_1 = \pi_1$ and $\kappa_0 = \pi_0 + \pi_1$. According to Lemma 1, $y \in (0,1)$.) Hence

$$y' = \frac{\pi_1(1-y)^N + \pi_2}{1 - \pi_1 N(1-x)(1-y)^{N-1}} > 0,$$

which implies that $F_f(a)$ is monotonically increasing in $a \in [\underline{f}, \bar{\nu}]$. Uniqueness follows immediately from Lemma 1. $\qquad \square$

Proposition 2 *For $N \geq 1$, $\pi_0, \pi_1, \pi_2 > 0$ and $\pi_1 N \leq 1$, the equilibrium distribution of broker fees satisfies $F_f(a) \geq F_\nu(a)$ for all $a \in [\underline{f}, \bar{\nu}]$.*

Proof. For any $a \in [\underline{f}, \bar{\nu}]$, Equation (A.2) implies

$$Prob(f > a) = \lambda \, Prob(\nu > a),$$

where $0 < \lambda = \pi_1(1 - F_f(a))^N + \pi_2 < 1$. Therefore, $Prob(f > a) \leq P(\nu > a)$ and hence $F_f(a) \geq F_\nu(a)$. $\qquad \square$

Proposition 3 *Let F_f be some distribution function with support on $[\underline{f}, \bar{\nu}]$. For $N \geq 1$, $\pi_0, \pi_1, \pi_2 > 0$ and $\pi_1 N \leq 1$, F_ν given by Equation (A.2) is a well-defined distribution function as long as $F_f(\underline{f}) \geq x$, where $x \in (0,1)$ solves $x + \pi_1(1-x)^N - (1-\pi_2) = 0$. F_ν has support on $[\underline{f}, \bar{\nu}]$ and is unique.*

Proof. What we have to show is that

(a) F_ν is unique and $F_\nu(\bar{\nu}) = 1$ (immediate)

(b) $F_\nu(\underline{f}) \geq 0$

(c) $F_\nu(a)$ is monotonically increasing in a for $a \in [\underline{f}, \bar{\nu}]$.

41

With regard to (b), $F_\nu(\underline{f}) \geq 0$ if $p(x^*) \leq 1$, where $x^* = F_f(\underline{f})$ and

$$p(x) = \frac{1-x}{\pi_1(1-x)^N + \pi_2}.$$

We have $p(0) = \frac{1}{\pi_1+\pi_2} > 1$ and $p(1) = 0$, so there is at least some $x \in (0,1)$ for which $p(x) = 1$. Taking derivatives of $p(x)$ w.r.t. x, for $x \in (0,1)$, yields

$$
\begin{aligned}
p'(x) &= \frac{\pi_1(N-1)(1-x)^N - \pi_2}{(\pi_1(1-x)^N + \pi_2)^2} \\
&= \frac{-\lambda + \pi_1 N(1-x)^N}{\lambda^2} \\
&= \frac{-\lambda + \pi_1 N \lambda (1-y)(1-x)^{N-1}}{\lambda^2} \\
&= -\frac{1 - \pi_1 N(1-y)(1-x)^{N-1}}{\lambda},
\end{aligned}
$$

where $\lambda = \pi_1(1-x)^N + \pi_2$. As long as $\pi_1 N \leq 1$, $p'(x) < 0$. (Note that for $N = 1$, $\pi_1 N < 1$. For $N > 1$ and $x \in (0,1)$, $(1-y)(1-x)^{N-1} < 1$.) Hence $p(x^*) \leq 1$ if and only if $x^* \geq x$, where $x + \pi_1(1-x)^N - (1-\pi_2) = 0$. According to Lemma 1, $x \in (0,1)$.

With regard to (c), let $y = F_\nu(a)$ and $x = F_f(a)$ for some $a \in [\underline{f}, \overline{\nu}]$. Equation (A.2) implies $y = 1 - p(x)$. Differentiation w.r.t. x yields $y' = 1 - p'(x)$, which implies that $y' \geq 0$ as long as $\pi_1 N \leq 1$. □

Corollary 1 *Let F_f be some distribution function with support on $[\underline{f}, \overline{\nu}]$ and $F_f(\underline{f}) \in (0,1)$. Set $G_f(a) = 1 - F_f(a)$ for all $a \in [\underline{f}, \overline{\nu}]$. For $N \geq 1$, let $\pi_2 \in \left(\frac{G_f(\underline{f}) - G_f^N(\underline{f})}{1 - G_f^N(\underline{f})}, G_f(\underline{f}) \right)$ with $\pi_2 \geq G_f(\underline{f}) - \frac{G_f^N(\underline{f})}{N}$, and let π_1 be defined as in Equation (A.3). Then F_ν given by Equation (A.2) is a well-defined distribution function with support on $[\underline{f}, \overline{\nu}]$. For each $a \in [\underline{f}, \overline{\nu}]$,*

$$F_\nu(a) = 1 - \frac{G_f^N(\underline{f}) \, G_f(a)}{G_f(\underline{f}) \, G_f^N(a) + (G_f^N(\underline{f}) - G_f^N(a)) \, \pi_2}. \tag{A.4}$$

Proof: First, we show that the conditions of Proposition 3—$\pi_0, \pi_1, \pi_2 > 0$, $\pi_1 N \leq 1$ and $F_f(\underline{f}) \geq x$ where $x \in (0, 1)$ solves $x + \pi_1(1 - x)^N - (1 - \pi_2) = 0$—are satisfied. They imply that F_ν given by Equation (A.2) is a well-defined distribution function with support on $[\underline{f}, \bar{\nu}]$.

Since $G_f(\underline{f}) \in (0, 1)$ and $\pi_2 > \frac{G_f(\underline{f}) - G_f^N(\underline{f})}{1 - G_f^N(\underline{f})}$, $\pi_2 > 0$. Equation (A.3) states

$$\pi_1 = \frac{G_f(\underline{f}) - \pi_2}{G_f^N(\underline{f})}, \tag{A.5}$$

which, with $\pi_2 < G_f(\underline{f})$, implies $\pi_1 > 0$. And since $\pi_2 > \frac{G_f(\underline{f}) - G_f^N(\underline{f})}{1 - G_f^N(\underline{f})}$,

$$\begin{aligned}
\pi_0 &= 1 - \pi_1 - \pi_2 \\
&= 1 - \frac{G_f(\underline{f}) - \pi_2}{G_f^N(\underline{f})} - \pi_2 \\
&= \frac{-(G_f(\underline{f}) - G_f^N(\underline{f})) + \pi_2(1 - G_f^N(\underline{f}))}{G_f^N(\underline{f})} \\
&> 0.
\end{aligned}$$

Moreover, with $\pi_2 \geq G_f(\underline{f}) - \frac{G_f^N(\underline{f})}{N}$,

$$\begin{aligned}
\pi_1 N &= \frac{G_f(\underline{f}) - \pi_2}{G_f^N(\underline{f})} N \\
&\leq \frac{G_f(\underline{f}) - G_f(\underline{f}) + \frac{G_f^N(\underline{f})}{N}}{G_f^N(\underline{f})} N \\
&= 1.
\end{aligned}$$

Lastly, Equation (A.3) is equivalent to $x + \pi_1(1 - x)^N - (1 - \pi_2) = 0$ with $x = F_f(\underline{f}) \in (0, 1)$.

To derive Equation (A.4), we rewrite Equation (A.2) as

$$F_\nu(a) = 1 - \frac{G_f(a)}{\pi_1 G_f^N(a) + \pi_2}$$

and substitute in Equation (A.5) to obtain

$$F_\nu(a) = 1 - \frac{G_f(a)}{\frac{G_f(\underline{f})-\pi_2}{G_f^N(\underline{f})}G_f^N(a) + \pi_2},$$

or

$$F_\nu(a) = 1 - \frac{G_f^N(\underline{f})\,G_f(a)}{G_f(\underline{f})\,G_f^N(a) + (G_f^N(\underline{f}) - G_f^N(a))\,\pi_2}.$$

$\square$

Corollary 2 *In the setting of Corollary 1, there is a strictly monotonic decreasing rela-tionship between the fraction of minimally informed borrowers, π_2, and expected borrower valuations, $E(\nu|\text{borrower is non-expert})$. Hence, given $E(\nu|\text{non-expert borrower})$ there ex-ists at most one π_2 that satisfies Equation (A.4).*

Proof: For each $a \in [\underline{f}, \overline{\nu}]$ with $F_f(a) = F_f(\underline{f})$, Equation (A.4) implies $F_\nu(a) = 0$. And for each $a \in (\underline{f}, \overline{\nu}]$ with $F_f(a) > F_f(\underline{f})$, Equation (A.4) implies a strictly monotonic increasing relationship between π_2 and $F_\nu(a)$. The larger (smaller) the fraction of minimally informed borrowers, the further the distribution of borrower valuations shifts to the left (right). As a result, there is a strictly monotonic decreasing relationship between π_2 and $E(\nu)$, where $E(\nu)$ is used as a short-cut for $E(\nu|\text{non-expert borrower})$. $\square$

For the extended model in Section 7.1, the equilibrium fee distribution satisfies

$$F_f(a) + \widetilde{\kappa}_1(a)(1 - F_f(a))^N - \widetilde{\kappa}_0(a) = 0, \tag{A.6}$$

for $a \in [\underline{f}, \overline{\nu}]$, where $\widetilde{\kappa}_1(a) = \pi_1(1 - F_{1,\nu}(a))$ and $\widetilde{\kappa}_0(a) = 1 - \pi_2(1 - F_{2,\nu}(a))$. Substituting in $w_\alpha(F_{2,\nu}(a))$ for $F_{1,\nu}(a)$, Equation (A.6) can be rewritten as

$$F_f(a) = \pi_0 + \pi_1\left\{1 - (1 - w_\alpha(F_{2,\nu}(a)))(1 - F_f(a))^N\right\} + \pi_2 F_{2,\nu}(a). \tag{A.7}$$

44

Proposition 4 *Let $F_{1,\nu}$ and $F_{2,\nu}$ be some distribution functions with support on $[\underline{f}, \overline{\nu}]$. For $N \geq 1$, $\pi_0, \pi_1, \pi_2 > 0$ and $\pi_1 N \leq 1$, the equilibrium distribution of broker fees is well-defined by Equation (A.6), has support on $[\underline{f}, \overline{\nu}]$ and is unique.*

Proof. Equation (A.6) implies that $F_\nu(\overline{\nu}) = 1$ if and only if $F_f(\overline{\nu}) = 1$. The broker's participation constraint implies that $F_f(a) = 0$ for $a < \underline{f}$. Hence F_f has support on $[\underline{f}, \overline{\nu}]$. Let $y = F_f(a)$, $x_1 = F_{1,\nu}(a)$ and $x_2 = F_{2,\nu}(a)$. According to Equation (A.6),

$$y + \pi_1(1 - x_1)(1 - y)^N - \pi_0 - \pi_1 - \pi_2 x_2 = 0.$$

Differentiation w.r.t. x_1 yields $y' - \pi_1(1 - y)^N - \pi_1 N(1 - x_1)(1 - y)^{N-1} y' - \pi_2 x_2' = 0$. For $x_1 \in [0, 1]$ and $\pi_1 N \leq 1$, we have $\pi_1 N(1 - x_1)(1 - y)^{N-1} < 1$. (Note that for $N = 1$, $\pi_1 N < 1$. For $N > 1$ and $x_1 = 0$, y solves Equation (A.1) with $\kappa_1 = \pi_1$ and $\kappa_0 = \pi_0 + \pi_1 + \pi_2 x_2$. According to Lemma 1, $y \in (0, 1)$.) Hence

$$y' = \frac{\pi_1(1 - y)^N + \pi_2 x_2'}{1 - \pi_1 N(1 - x_1)(1 - y)^{N-1}},$$

which implies that $F_f(a)$ is monotonically increasing in $a \in [\underline{f}, \overline{\nu}]$. Uniqueness follows immediately from Lemma 1. $\qquad\square$

Proposition 5 *For $N \geq 1$, $\pi_0, \pi_1, \pi_2 > 0$ and $\pi_1 N \leq 1$, the equilibrium distribution of broker fees satisfies $F_f(a) \geq \min\{F_{1,\nu}(a), F_{2,\nu}(a)\}$, for all $a \in [\underline{f}, \overline{\nu}]$.*

Proof. First consider the case where $F_{1,\nu}(a) \leq F_{2,\nu}(a)$ for some $a \in [\underline{f}, \overline{\nu}]$. Equation (A.6) implies $F_f(a) + \pi_1(1 - F_{1,\nu}(a))(1 - F_f(a))^N - 1 + \pi_2 - \pi_2 F_{1,\nu}(a) \geq 0$, or

$$\begin{aligned} F_{1,\nu}(a) &\leq \frac{F_f(a) + \pi_1(1 - F_f(a))^N - 1 + \pi_2}{\pi_1(1 - F_f(a))^N + \pi_2} \\ &= 1 + \frac{F_f(a) - 1}{\pi_1(1 - F_f(a))^N + \pi_2}. \end{aligned}$$

45

We obtain $1 - F_f(a) \leq \lambda(1 - F_{1,\nu}(a))$, where $0 < \lambda = \pi_1(1 - F_f(a))^N + \pi_2 < 1$. Therefore, $1 - F_f(a) \leq 1 - F_{1,\nu}(a)$ and hence $F_f(a) \geq F_{1,\nu}(a)$. A similar argument proves the result for $F_{1,\nu}(a) \geq F_{2,\nu}(a)$. $\square$

Lemma 2 *The function $w_\alpha(x)$ defined in Equation (21) increases monotonically from zero at $x = 0$ to one at $x = 1$. For $\alpha < 1$, $F_{1,\nu} \geq F_{2,\nu}$, and for $\alpha > 1$, $F_{1,\nu} \leq F_{2,\nu}$.*

Proof. For a given α, we differentiate $w_\alpha(x)$ w.r.t. x:

$$w'_\alpha(x) = \begin{cases} -1/(1-\alpha)\,\alpha^x \ln(\alpha), & \text{if } \alpha \neq 1 \\ 1, & \text{if } \alpha = 1. \end{cases}$$

Since $-\ln(\alpha)/(1-\alpha)$ is negative for $\alpha > 1$ and for $\alpha < 1$, $w'_\alpha(x)$ is always positive.

For a given $x \in [0, 1]$, we also differentiate $w_\alpha(x)$ w.r.t. $\alpha \neq 1$:

$$\frac{dw_\alpha(x)}{d\alpha} = \frac{-x\alpha^{x-1}}{1-\alpha} + \frac{1-\alpha^x}{(1-\alpha)^2}.$$

Hence $dw_\alpha(x)/d\alpha \leq 0$ if and only if $d_x(\alpha) \equiv -x\alpha^{x-1}(1-\alpha) + 1 - \alpha^x \leq 0$. Since $d_x(1) = 0$, $dw_\alpha(x)/d\alpha \leq 0$ if and only if $d'_x(\alpha) \geq 0$ for $\alpha \leq 1$ and $d'_x(\alpha) \leq 0$ for $\alpha \geq 1$. Since

$$d'_x(\alpha) = x(1-x)\alpha^{x-2}(1-\alpha),$$

the latter is indeed the case. $\square$

Proposition 6 *Let F_f be some distribution function with support on $[\underline{f}, \overline{\nu}]$. For $N \geq 1$, $\pi_0, \pi_1, \pi_2 > 0$ and $\pi_1 N \leq 1$, $F_{2,\nu}$ given by Equation (A.7) is a well-defined distribution function as long as $F_f(\underline{f}) \geq x$, where $x \in (0, 1)$ solves $x + \pi_1(1-x)^N - (1 - \pi_2) = 0$. $F_{2,\nu}$ has support on $[\underline{f}, \overline{\nu}]$ and is unique.*

Proof. What we have to show is that

(a) $F_{2,\nu}(\bar{\nu}) = 1$ (immediate)

(b) $F_{2,\nu}(\underline{f}) \geq 0$

(c) $F_{2,\nu}(a)$ is monotonically increasing in a for $a \in [\underline{f}, \bar{\nu}]$

(d) $F_{2,\nu}$ is unique.

With regard to (c), let $y = F_{2,\nu}(a)$ and $x = F_f(a)$. According to Equation (A.7),

$$x = \pi_0 + \pi_1 \left\{1 - (1 - w_\alpha(y))(1 - x)^N\right\} + \pi_2 y. \tag{A.8}$$

Differentiation w.r.t. x yields $1 = \pi_1 \left(w'_\alpha(y)y'(1 - x)^N + (1 - w_\alpha(y))N(1 - x)^{N-1}\right) + \pi_2 y'$, or

$$y' = \frac{1 - \pi_1(1 - w_\alpha(y))N(1 - x)^{N-1}}{\pi_1 w'_\alpha(y)(1 - x)^N + \pi_2}.$$

Because $w'_\alpha(y) > 0$ (see proof of Lemma 2), we have $\pi_1 w'_\alpha(y)(1 - x)^N + \pi_2 > 0$. Hence $y' \geq 0$ as long as $\pi_1 N \leq 1$.

With regard to (b), note that for any given x, the derivative of the right-hand side of Equation (A.8) w.r.t. y is strictly positive. (Recall that $w'_\alpha(y) > 0$.) Hence for any given x, there is at most one value of y that satisfies Equation (A.8). According to Lemma 1 there exists a unique $x \in (0, 1)$ such that $x + \pi_1(1 - x)^N - (1 - \pi_2) = 0$. For this value of x, y equals zero. And since $y' \geq 0$, we obtain $F_{2,\nu}(\underline{f}) \geq 0$ as long as $F_f(\underline{f}) \geq x$.

With regard to (d), for a given $x = F_f(a)$ suppose that $y < \bar{y}$. Because $w'_\alpha(y) > 0$, we have $-\pi_1(w_\alpha(\bar{y}) - w_\alpha(y))(1 - x)^N < \pi_2(\bar{y} - y)$ and hence

$$\pi_1 w_\alpha(y)(1 - x)^N + \pi_2 y < \pi_1 w_\alpha(\bar{y})(1 - x)^N + \pi_2 \bar{y},$$

meaning that there can only be one solution to Equation (A.8). $\square$

B.   Additional Tables and Figures

Table B.1: **List of loan and borrower characteristics**

Variable	Description
Loan purpose	Purchase or refinance loan
Product type	2/28 (hybrid) or 30-year fixed (fixed) mortgage
Loan amount	Amount borrowed in $1,000
Low documentation	Indicator for a loan with low (i.e., limited or stated) rather than full documentation
Loan-to-value ratio (LTV)	Value of the loan divided by that of the house
Prepay penalty	Indicator for a loan with a prepayment penalty
Initial rate	Initial mortgage rate
Rate margin for hybrid loans	Rate margin that is added to LIBOR to determine the floating rate
Discount points	Discount points charged on the loan
Lender charges	Upfront charges by New Century in addition to points and broker fees
FICO	Fair, Isaac and Company's borrower credit score at origination
Debt-to-income ratio (DTI)	All monthly debt payments divided by monthly gross income
Education	Fraction of population in zip code with B.A. degree, based on 2000 census data
Broker competition	For a given month and zip code, broker competition is the number of brokers who submitted loan applications to New Century divided by the number of housing units (in thousands)
House price appreciation	Annualized rate estimated for each Core Based Statistical Area using quarterly FHFA All-Transaction Index data available at www.fhfa.gov. Forecasts are based on an AR(6) model, which offers appealing trade-offs between simplicity and forecasting accuracy (Larson (2010)).
Property tax	Average tax rates on owner-occupied housing for 2005–2009 are available by county at taxfoundation.org/article/property-taxes-owner-occupied-housing-county-ranked-taxes-paid-2005-2009-five-year-average
Price-to-rent ratio	The price-to-rent ratio is defined as the ratio of the house price to the annual rent. Aggregate data on price-to-rent ratios are available for large metropolitan areas, and are based on data provided by Fiserv, PPR Inc. and Moody's Economy.com (see money.cnn.com/real_estate/storysupplement/price_to_rent/). We use 15-year averages for 1989–2003.

Table B.2: **Alternative specifications of state vector** The table lists additional scenarios for the state vector: (A)=(a) Reference scenario, (b) Small loan, (c) Uneducated neighbors, (d) Low house price appreciation, and (e) High price-to-rent ratio. All other parameters are as in Table 4.

	(a) Reference	(b) Small loan	(c) Uneducated nbrs	(d) Low HPA	(e) High prc/rent
Loan amount (in $1,000)	200	100	200	200	200
Education (%)	15	15	5	15	15
House price appreciation (%)	15	15	15	5	15
Price-to-rent ratio	15	15	15	15	20

Table B.3: **Implied borrower informedness and valuation distributions, for additional state vectors** The table summarizes our findings regarding the implied borrower informedness and valuation distributions. We use $E_{ne}(\min\{f_1,\nu\})$ and $E_{ne}(\nu)$ as shortcuts for $E(\min\{f_1,\nu\}|\text{non-expert})$ and $E(\nu|\text{non-expert})$, and $E_p(\gamma)$ as a shortcut for $E(\gamma|\text{partially informed})$. Δ is the fraction of non-expert borrowers that are partially informed. Δ_1 is computed as $\Delta_{(A)}$, except that the perceived likelihood that a friend's fees do not exceed benchmark valuations is for the given state vector. Dollar values are in \$1,000. The results are based on the estimates in Tables 5 and D.1 through D.4. The state vector specifications are as in Table B.2.

	(a)	(b)	(c)	(a)	(b)	(c)	(d)	(e)
	Refinance: Hybrid			*Purchase: Hybrid*				
$E(f)$	4.353	2.963	4.696	3.413	2.224	3.621	3.356	3.403
$b_{\widetilde{f}}$	4.688	3.129	4.688	3.643	2.340	3.643	3.567	3.633
b_k	0.266	0.126	0.266	2.941	1.470	2.941	0.947	2.589
π_2	0.55	0.70	0.77	0.04	0.08	0.05	0.29	0.06
$\underline{f}$	0.874	0.778	1.156	0.450	0.378	0.595	0.416	0.444
$E_{ne}(\min\{f_1,\nu\})$	3.612	2.482	3.822	3.275	2.097	3.459	2.857	3.219
$E_{ne}(\nu)$	5.028	3.212	5.025	7.045	3.799	6.932	4.616	6.510
$E_p(\gamma)$	0.72	0.76	0.75	0.48	0.55	0.50	0.65	0.51
Δ	0.44	0.29	0.21	0.96	0.92	0.95	0.70	0.94
$\Delta - \Delta_{(A)}$	0.00	-0.15	-0.23	0.00	-0.04	-0.01	-0.26	-0.02
$\Delta_1 - \Delta_{(A)}$	0.00	-0.10	0.00	0.00	-0.04	0.00	-0.32	-0.02
	Refinance: FRM			*Purchase: FRM*				
$E(f)$	3.944	2.675	4.221	3.031	1.939	3.255	2.789	2.990
$b_{\widetilde{f}}$	4.211	2.807	4.211	3.274	2.061	3.274	3.014	3.233
b_k	0.293	0.141	0.293	3.128	1.564	3.128	1.011	2.776
π_2	0.54	0.69	0.69	0.04	0.06	0.05	0.23	0.05
$\underline{f}$	0.781	0.649	1.045	0.195	0.304	0.350	0.043	0.170
$E_{ne}(\min\{f_1,\nu\})$	3.212	2.211	3.396	2.873	1.825	3.076	2.344	2.801
$E_{ne}(\nu)$	4.620	2.920	4.659	6.802	3.672	6.743	4.339	6.380
$E_p(\gamma)$	0.71	0.75	0.73	0.46	0.50	0.48	0.62	0.49
Δ	0.45	0.29	0.30	0.96	0.94	0.95	0.77	0.95
$\Delta - \Delta_{(A)}$	0.00	-0.15	-0.15	0.00	-0.02	-0.01	-0.19	-0.01
$\Delta_1 - \Delta_{(A)}$	0.00	-0.13	0.00	0.00	-0.02	0.00	-0.24	-0.02

Table B.4: **Reduction in fees and γ if a partially informed borrower consults one extra friend, for additional state vectors** The table reports the reduction in expected broker fees (in \$1,000) and in the expected value of γ in Equation (3) if a partially informed borrower were to consult with one extra friend. The results are based on the estimates in Tables 5 and D.1 through D.4. The state vector specifications are as in Table B.2. The Monte Carlo estimates are based on 100,000 scenarios.

	(a)	(b)	(c)	(a)	(b)	(c)	(d)	(e)
	Refinance: Hybrid			*Purchase: Hybrid*				
Reduction in fees	0.599	0.347	0.547	0.852	0.495	0.847	0.630	0.820
Reduction in γ	0.13	0.13	0.13	0.14	0.15	0.14	0.15	0.14
	Refinance: FRM			*Purchase: FRM*				
Reduction in fees	0.579	0.335	0.544	0.817	0.463	0.812	0.629	0.794
Reduction in γ	0.14	0.13	0.14	0.13	0.14	0.13	0.15	0.14

Table B.5: **Expected broker fees if borrowers become more informed, for additional state vectors** The table reports expected broker fees $E(f|\pi(\theta))$ defined in Equation (18) (in \$1,000), for $\pi(\theta) = (\pi_0, \pi_1 + \theta\pi_2, (1-\theta)\pi_2)$ and $\theta \in \{0, 0.5, 1\}$. The results are based on the estimates in Tables 5 and D.1 through D.4. The state vector specifications are as in Table 4. The Monte Carlo estimates are based on 100,000 scenarios.

θ	(a)	(b)	(c)	(a)	(b)	(c)	(d)	(e)
	Refinance: Hybrid			*Purchase: Hybrid*				
0	4.353	2.963	4.696	3.413	2.224	3.621	3.356	3.403
0.5	3.964	2.707	4.233	3.343	2.159	3.539	3.098	3.310
1	3.573	2.452	3.769	3.272	2.094	3.455	2.838	3.215
	Refinance: FRM			*Purchase: FRM*				
0	3.944	2.675	4.221	3.031	1.939	3.255	2.789	2.990
0.5	3.562	2.429	3.788	2.951	1.881	3.164	2.561	2.895
1	3.179	2.183	3.355	2.870	1.823	3.072	2.330	2.798
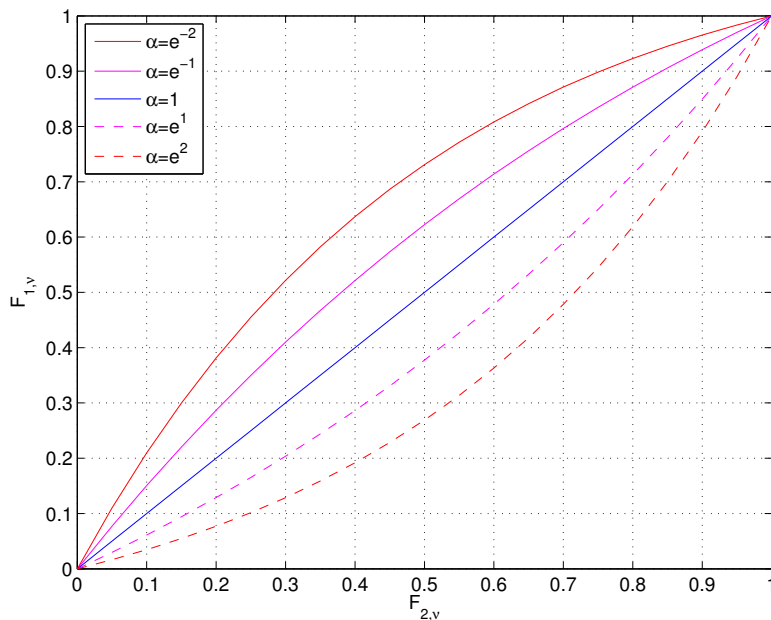
Figure B.1: **Relation between conditional borrower valuation distributions for different α** The figure shows $F_{1,\nu} = w_\alpha(F_{2,\nu})$ as a function of $F_{2,\nu}$, for different values of α.

C. Non-Crossing Quantile Regressions

The classical estimator of the regression coefficients for the quantile function is

$$\widehat{\beta}(q) \;=\; \text{argmin}_{\beta_0,\beta} \left(\sum_{f_i \geq \beta_0 + x_i \beta} q(f_i - \beta_0 - x_i \beta) - \sum_{f_i < \beta_0 + x_i \beta} (1-q)(f_i - \beta_0 - x_i \beta) \right).$$

This optimization problem can be rewritten as a standard linear programming problem:

$$\widehat{\beta}(q) \;=\; \text{argmin}_{\beta_0,\beta,u^+,u^-} \left(q\,u^+ + (1-q)\,u^- \right), \tag{C.1}$$

subject to

$$f - \beta_0 - x\beta \;=\; u^+ - u^-, \quad u^+, u^- \geq 0 \text{ and only one is non-zero.} \tag{C.2}$$

Suppose that the state vector consists of m variables, each of which has bounded support, and that $\mathcal{D} \subset \mathcal{R}^m$ is a closed convex polytope such that $x \in \mathcal{D}$ for all x. For $0 < q_1 < \ldots < q_k < 1$, we replace the optimization problem in (C.1) and (C.2) with

$$\widehat{\beta} \;=\; \text{argmin}_{\beta_0(q_j),\beta(q_j),u_i^+(q_j),u_i^-(q_j)} \sum_{j=1}^{k} \sum_{i=1}^{n} \left\{ q_j\, u_i^+(q_j) + (1-q_j)\, u_i^-(q_j) \right\},$$

where n is the number of observations, subject to

$$f - \beta_0(q_j) - x_i \beta(q_j) \;=\; u_i^+(q_j) - u_i^-(q_j), \quad u_i^+(q_j), u_i^-(q_j) \geq 0,$$

and the non-crossing restriction

$$\beta_0(q_{j-1}) + x\beta(q_{j-1}) \;\leq\; \beta_0(q_j) + x\beta(q_j), \tag{C.3}$$

for all $x \in \mathcal{D}$ and $j = 2, \ldots, k$.

Since the control variables have bounded support, we can represent $\mathcal{D}$ as the convex hull of 2^m points. It therefore suffices to enforce the non-crossing restriction (C.3) at each of the 2^m vertices of $\mathcal{D}$. Without loss of generality, suppose that $\mathcal{D} = [0, 1]^m$.[30] For $\eta_0(q_1) = \beta_0(q_1)$ and $\eta(q_1) = \beta(q_1)$, and $\eta_0(q_j) = \beta_0(q_j) - \beta_0(q_{j-1})$ and $\eta(q_j) = \beta(q_j) - \beta(q_{j-1})$, the constraint in (C.3) is equivalent to

$$\eta_0(q_j) + x\eta(q_j) \geq 0, \quad \text{for all } x \in \mathcal{D} \text{ and } j = 2, \ldots, k. \tag{C.4}$$

For $\eta(q) = (\eta_1(q), \ldots, \eta_m(q))'$, break each component $\eta_l(q)$ into its positive and negative parts: $\eta_l(q) = \eta_l^+(q) - \eta_l^-(q)$, where both $\eta_l^+(q)$ and $\eta_l^-(q)$ are non-negative and only one is non-zero. Using this parameterization, the "worst case scenario" in (C.4) obtains for $x^{wc} = (x_1^{wc}, \ldots, x_m^{wc})$, where $x_l^{wc} = 1$ when $\eta_l(q) < 0$ and $x_l^{wc} = 0$ when $\eta_l(q) > 0$. Hence the constraint in (C.4) becomes

$$\eta_0(q_j) - \sum_{l=1}^{m} \eta_l^-(q_j) \geq 0, \quad \text{for all } j = 2, \ldots, k.$$

As a result, only a total of $k - 1$ additional constraints are needed to ensure non-crossing, rather than the $n \times (k-1)$ constraints in (C.3). After reparameterization, the problem is thus a straightforward linear programming problem, which can be solved via standard software. As noted by Bondell, Reich, and Wang (2010), the linear program is extremely sparse and the use of a sparse matrix representation via the SparseM package of Koenker and Ng (2003) is recommended.

---

[30] As long as each of the control variables in x has bounded support, there exists an invertible affine transformation that maps $\mathcal{D}$ to $[0, 1]^m$. The transformation is performed before the estimation, and then transformed back after the estimation, while retaining the non-crossing property.

D. Material for Online Appendix

D.1 *Detailed non-crossing quantile regression results*

Table D.1: **Non-crossing quantile regression results: Hybrid refinance loans** The table reports the results for the non-crossing quantile regressions of broker fees (as % of loan amount) on the state vector x, for select quantiles. The data include all 23,246 hybrid refinance loans.

	Quantiles						
	2.5%	10%	30%	50%	70%	90%	97.5%
Constant	-0.870	-0.422	-0.251	-0.143	0.199	0.923	1.988
	(0.185)	(0.168)	(0.126)	(0.123)	(0.142)	(0.194)	(0.307)
YSP (% of loan amt)	-0.175	-0.298	-0.338	-0.368	-0.401	-0.431	-0.365
	(0.022)	(0.018)	(0.014)	(0.014)	(0.014)	(0.019)	(0.032)
Log loan amt (% of loan amt)	0.174	0.267	0.359	0.436	0.481	0.498	0.460
	(0.014)	(0.011)	(0.008)	(0.009)	(0.011)	(0.016)	(0.039)
Low documentation	-0.064	-0.064	-0.056	-0.038	-0.036	-0.006	-0.080
	(0.027)	(0.023)	(0.019)	(0.019)	(0.022)	(0.031)	(0.049)
Prepay penalty	0.272	0.296	0.330	0.406	0.428	0.437	0.398
	(0.039)	(0.032)	(0.027)	(0.027)	(0.033)	(0.048)	(0.077)
Initial rate (%)	0.007	0.007	0.035	0.040	0.039	0.033	0.007
	(0.014)	(0.014)	(0.011)	(0.011)	(0.012)	(0.017)	(0.022)
Margin for hybrid loans (%)	0.165	0.165	0.176	0.194	0.201	0.200	0.190
	(0.033)	(0.030)	(0.022)	(0.022)	(0.026)	(0.035)	(0.048)
Discount points (% of loan amt)	0.035	0.061	0.096	0.096	0.078	-0.022	-0.022
	(0.073)	(0.036)	(0.030)	(0.028)	(0.038)	(0.046)	(0.103)
Education (%)	-0.014	-0.016	-0.018	-0.018	-0.019	-0.018	-0.013
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)
FL	-0.987	-1.050	-0.552	-0.413	-0.186	0.136	0.497
	(0.068)	(0.046)	(0.037)	(0.039)	(0.042)	(0.059)	(0.106)
West w/o CA	-0.152	-0.152	-0.216	-0.225	-0.123	0.210	0.596
	(0.036)	(0.031)	(0.025)	(0.026)	(0.032)	(0.053)	(0.086)
MidWest	-0.341	-0.341	-0.372	-0.374	-0.290	-0.065	0.239
	(0.047)	(0.037)	(0.029)	(0.033)	(0.038)	(0.061)	(0.129)
South w/o FL & TX	-0.028	0.026	0.027	0.030	0.123	0.388	0.555
	(0.049)	(0.039)	(0.034)	(0.035)	(0.040)	(0.060)	(0.095)
NorthEast	-0.247	-0.247	-0.247	-0.253	-0.257	-0.233	-0.141
	(0.080)	(0.059)	(0.049)	(0.048)	(0.048)	(0.066)	(0.150)
2005	0.093	0.113	0.102	0.108	0.142	0.204	0.269
	(0.036)	(0.027)	(0.021)	(0.021)	(0.024)	(0.035)	(0.055)
2006	0.051	0.051	-0.002	-0.053	-0.114	-0.248	-0.418
	(0.053)	(0.040)	(0.031)	(0.032)	(0.035)	(0.048)	(0.077)

Table D.2: **Non-crossing quantile regression results: Fixed refinance loans** The table reports the results for the non-crossing quantile regressions of broker fees (as % of loan amount) on the state vector x, for select quantiles. The data include all 11,565 fixed refinance loans.

				Quantiles			
	2.5%	10%	30%	50%	70%	90%	97.5%
Constant	-0.142	0.174	0.373	0.527	1.051	1.738	2.391
	(0.169)	(0.139)	(0.123)	(0.125)	(0.136)	(0.203)	(0.302)
YSP (% of loan amt)	-0.338	-0.440	-0.440	-0.498	-0.528	-0.522	-0.418
	(0.040)	(0.030)	(0.025)	(0.022)	(0.024)	(0.033)	(0.050)
Log loan amt (% of loan amt)	0.132	0.214	0.345	0.407	0.433	0.433	0.386
	(0.016)	(0.014)	(0.011)	(0.011)	(0.014)	(0.022)	(0.051)
Low documentation	-0.158	-0.142	-0.083	-0.078	-0.073	-0.113	-0.113
	(0.043)	(0.037)	(0.031)	(0.031)	(0.033)	(0.044)	(0.071)
Prepay penalty	0.240	0.240	0.315	0.374	0.374	0.452	0.452
	(0.093)	(0.062)	(0.050)	(0.054)	(0.057)	(0.088)	(0.134)
Initial rate (%)	0.068	0.068	0.068	0.086	0.082	0.082	0.082
	(0.022)	(0.020)	(0.018)	(0.016)	(0.018)	(0.026)	(0.039)
Discount points (% of loan amt)	-0.046	0.001	0.042	0.025	-0.014	-0.129	-0.151
	(0.042)	(0.030)	(0.025)	(0.025)	(0.029)	(0.039)	(0.081)
Education (%)	-0.013	-0.013	-0.013	-0.013	-0.016	-0.017	-0.009
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.005)
FL	-0.824	-0.847	-0.364	-0.218	-0.099	0.185	0.493
	(0.072)	(0.062)	(0.049)	(0.047)	(0.054)	(0.082)	(0.125)
West w/o CA	-0.038	-0.082	-0.150	-0.145	-0.119	0.309	0.621
	(0.056)	(0.039)	(0.035)	(0.040)	(0.046)	(0.091)	(0.108)
MidWest	-0.385	-0.385	-0.426	-0.435	-0.365	-0.136	0.315
	(0.105)	(0.061)	(0.048)	(0.051)	(0.061)	(0.102)	(0.184)
South w/o FL & TX	-0.009	-0.009	-0.036	-0.030	0.042	0.228	0.359
	(0.076)	(0.057)	(0.049)	(0.050)	(0.058)	(0.088)	(0.125)
NorthEast	-0.184	-0.181	-0.193	-0.215	-0.271	-0.242	-0.077
	(0.115)	(0.088)	(0.061)	(0.061)	(0.061)	(0.109)	(0.167)
2005	0.155	0.265	0.208	0.218	0.240	0.324	0.451
	(0.050)	(0.034)	(0.028)	(0.029)	(0.033)	(0.048)	(0.074)
2006	0.155	0.195	0.126	0.114	0.080	0.027	-0.051

Table D.3: **Non-crossing quantile regression results: Hybrid purchase loans** The table reports the results for the non-crossing quantile regressions of broker fees (as % of loan amount) on the state vector x, for select quantiles. The data include all 9,727 hybrid purchase loans.

				Quantiles			
	2.5%	10%	30%	50%	70%	90%	97.5%
Constant	-0.490	-0.284	0.202	0.288	0.450	0.869	1.872
	(0.242)	(0.252)	(0.213)	(0.235)	(0.309)	(0.377)	(0.501)
YSP (% of loan amt)	-0.159	-0.166	-0.181	-0.191	-0.218	-0.277	-0.290
	(0.021)	(0.025)	(0.020)	(0.019)	(0.022)	(0.030)	(0.049)
Log loan amt (% of loan amt)	0.078	0.137	0.204	0.273	0.329	0.400	0.464
	(0.012)	(0.014)	(0.011)	(0.011)	(0.014)	(0.019)	(0.036)
Low documentation	-0.052	-0.052	-0.018	-0.018	0.000	0.011	0.025
	(0.028)	(0.032)	(0.026)	(0.026)	(0.032)	(0.045)	(0.083)
Prepay penalty	0.189	0.189	0.237	0.260	0.321	0.389	0.383
	(0.041)	(0.042)	(0.036)	(0.035)	(0.042)	(0.060)	(0.108)
Initial rate (%)	0.047	0.047	0.048	0.060	0.080	0.104	0.104
	(0.016)	(0.019)	(0.014)	(0.015)	(0.019)	(0.028)	(0.048)
Margin for hybrid loans (%)	0.042	0.042	0.042	0.042	0.042	0.022	-0.054
	(0.041)	(0.044)	(0.037)	(0.042)	(0.053)	(0.064)	(0.064)
Discount points (% of loan amt)	-0.061	-0.056	-0.056	-0.056	-0.056	-0.056	-0.102
	(0.063)	(0.057)	(0.053)	(0.053)	(0.048)	(0.130)	(0.279)
Education (%)	-0.007	-0.007	-0.010	-0.011	-0.012	-0.010	-0.011
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.005)
FL	-0.513	-0.733	-0.794	-0.490	-0.332	-0.108	0.072
	(0.053)	(0.059)	(0.059)	(0.052)	(0.064)	(0.092)	(0.154)
TX	0.079	0.163	0.163	0.222	0.248	0.349	0.494
	(0.073)	(0.085)	(0.057)	(0.065)	(0.076)	(0.103)	(0.176)
West w/o CA	0.008	0.008	-0.065	-0.134	-0.182	-0.118	0.053
	(0.044)	(0.057)	(0.037)	(0.038)	(0.051)	(0.076)	(0.156)
MidWest	-0.135	-0.135	-0.127	-0.072	-0.072	0.070	0.249
	(0.058)	(0.067)	(0.054)	(0.052)	(0.066)	(0.095)	(0.165)
South w/o FL & TX	0.054	0.091	0.056	0.051	0.015	0.013	0.101
	(0.056)	(0.070)	(0.050)	(0.052)	(0.065)	(0.088)	(0.147)
NorthEast	-0.120	-0.120	-0.055	0.036	0.146	0.533	1.029
	(0.086)	(0.085)	(0.069)	(0.072)	(0.092)	(0.151)	(0.276)
2005	-0.002	0.082	0.081	0.037	0.037	0.042	0.031
	(0.040)	(0.043)	(0.029)	(0.031)	(0.040)	(0.057)	(0.101)
2006	-0.008	0.085	0.044	0.035	-0.039	-0.282	-0.454
	(0.060)	(0.061)	(0.044)	(0.047)	(0.057)	(0.074)	(0.128)
b_k (% of loan amt)	0.017	0.017	0.031	0.031	0.031	0.031	0.031
	(0.017)	(0.022)	(0.014)	(0.015)	(0.020)	(0.024)	(0.051)

Table D.4: **Non-crossing quantile regression results: Fixed purchase loans** The table reports the results for the non-crossing quantile regressions of broker fees (as % of loan amount) on the state vector x, for select quantiles. The data include all 2,258 fixed purchase loans.

	Quantiles						
	2.5%	10%	30%	50%	70%	90%	97.5%
Constant	-0.196	0.153	0.660	0.635	0.577	0.897	1.986
	(0.336)	(0.282)	(0.285)	(0.306)	(0.337)	(0.511)	(1.031)
YSP (% of loan amt)	-0.139	-0.139	-0.194	-0.204	-0.204	-0.296	-0.237
	(0.060)	(0.053)	(0.048)	(0.046)	(0.048)	(0.074)	(0.152)
Log loan amt (% of loan amt)	0.105	0.115	0.183	0.212	0.245	0.339	0.362
	(0.029)	(0.027)	(0.022)	(0.023)	(0.027)	(0.046)	(0.069)
Low documentation	0.031	0.049	0.049	0.052	0.084	0.139	-0.009
	(0.065)	(0.064)	(0.063)	(0.064)	(0.072)	(0.099)	(0.186)
Prepay penalty	0.182	0.103	0.103	0.152	0.233	0.461	0.307
	(0.152)	(0.141)	(0.126)	(0.119)	(0.122)	(0.181)	(0.390)
Initial rate (%)	-0.003	-0.003	-0.003	0.038	0.081	0.081	0.081
	(0.044)	(0.037)	(0.035)	(0.035)	(0.039)	(0.059)	(0.110)
Margin for hybrid loans (%)	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Discount points (% of loan amt)	-0.210	-0.210	-0.210	-0.176	-0.058	-0.058	-0.022
	(0.066)	(0.067)	(0.059)	(0.058)	(0.057)	(0.069)	(0.190)
Education (%)	-0.008	-0.008	-0.012	-0.012	-0.012	-0.012	-0.012
	(0.004)	(0.003)	(0.003)	(0.004)	(0.004)	(0.007)	(0.013)
FL	-0.485	-0.619	-0.408	-0.216	-0.057	0.103	0.327
	(0.127)	(0.127)	(0.128)	(0.116)	(0.128)	(0.204)	(0.334)
TX	0.315	0.704	0.570	0.570	0.666	0.741	1.027
	(0.181)	(0.146)	(0.131)	(0.143)	(0.164)	(0.239)	(0.382)
West w/o CA	0.087	0.087	0.087	0.002	-0.096	-0.085	-0.085
	(0.114)	(0.132)	(0.097)	(0.100)	(0.120)	(0.211)	(0.339)
MidWest	-0.149	-0.261	-0.261	-0.261	-0.138	0.088	0.622
	(0.183)	(0.190)	(0.143)	(0.140)	(0.153)	(0.243)	(0.341)
South w/o FL & TX	0.068	0.362	0.362	0.362	0.362	0.585	0.902
	(0.180)	(0.157)	(0.152)	(0.151)	(0.155)	(0.222)	(0.353)
NorthEast	-0.091	-0.044	-0.021	-0.009	-0.009	0.137	0.137
	(0.177)	(0.196)	(0.164)	(0.181)	(0.204)	(0.295)	(0.432)
2005	0.098	0.135	0.135	0.101	0.093	0.072	-0.152
	(0.091)	(0.080)	(0.067)	(0.070)	(0.077)	(0.120)	(0.236)
2006	0.149	0.174	0.106	0.039	-0.048	-0.129	-0.710
	(0.128)	(0.105)	(0.092)	(0.091)	(0.093)	(0.141)	(0.258)
b_k (% of loan amt)	0.072	0.111	0.118	0.118	0.118	0.118	0.118
	(0.049)	(0.039)	(0.036)	(0.041)	(0.043)	(0.069)	(0.114)

D.2 Borrower informedness conditional on broker fees

Given conditioning characteristics X, $\pi = \pi(X)$ is the ex-ante distribution of informedness across the borrower population. Even though we can compute π, we cannot observe an individual borrower's level of informedness. But once we observe the fee that borrower has paid, we can update the likelihood of him being fully, partially or minimally informed:

$$
\begin{aligned}
\pi_0^{ex,-}(a) &= Prob(\text{fully informed borrower}|f \le a) = \frac{\pi_0}{F_f(a)} \\
\pi_1^{ex,-}(a) &= Prob(\text{partially informed borrower}|f \le a) = \frac{F_f(a) - \pi_0 - \pi_2 F_\nu(a)}{F_f(a)} \\
\pi_2^{ex,-}(a) &= Prob(\text{minimally informed borrower}|f \le a) = \frac{\pi_2 F_\nu(a)}{F_f(a)},
\end{aligned}
$$

for $a \in [\underline{f}, \overline{\nu}]$.

For hybrid refinance loans, the left panel of Figure D.1 displays $\pi_0^{ex,-}(a)$, $\pi_1^{ex,-}(a)$ and $\pi_2^{ex,-}(a)$ as a function of a. As the fee threshold a increases, the ex-post probability that a borrower whose fees do not exceed a is fully informed decreases whereas the ex-post probability that he is minimally informed increases. The ex-post probability of the borrower being partially informed is a hump shaped function of a.

Let $\pi_0^{ex,+}(a) = Prob(\text{fully informed borrower}|f > a)$ for $a \in [\underline{f}, \overline{\nu})$, and let $\pi_1^{ex,+}(a)$ and $\pi_2^{ex,+}(a)$ denote similar ex-post probabilities for the borrower being partially or minimally informed. Bayes' rule implies

$$
\pi_i^{ex,+}(a) = \frac{\pi_i - \pi_i^{ex,-}(a) F_f(a)}{1 - F_f(a)}, \quad i = 0, 1, 2.
$$

For $\underline{f} \le a < b < \overline{\nu}$ and $\pi_0^{ex}(a,b) = Prob(\text{fully informed borrower}|a < f \le b)$, and similar definitions for $\pi_1^{ex}(a,b)$ and $\pi_2^{ex}(a,b)$, we have

$$
\pi_i^{ex}(a,b) = \frac{\pi_i - \pi_i^{ex,-}(a) F_f(a) - \pi_i^{ex,+}(b)(1 - F_f(b))}{F_f(b) - F_f(a)}, \quad i = 0, 1, 2.
$$
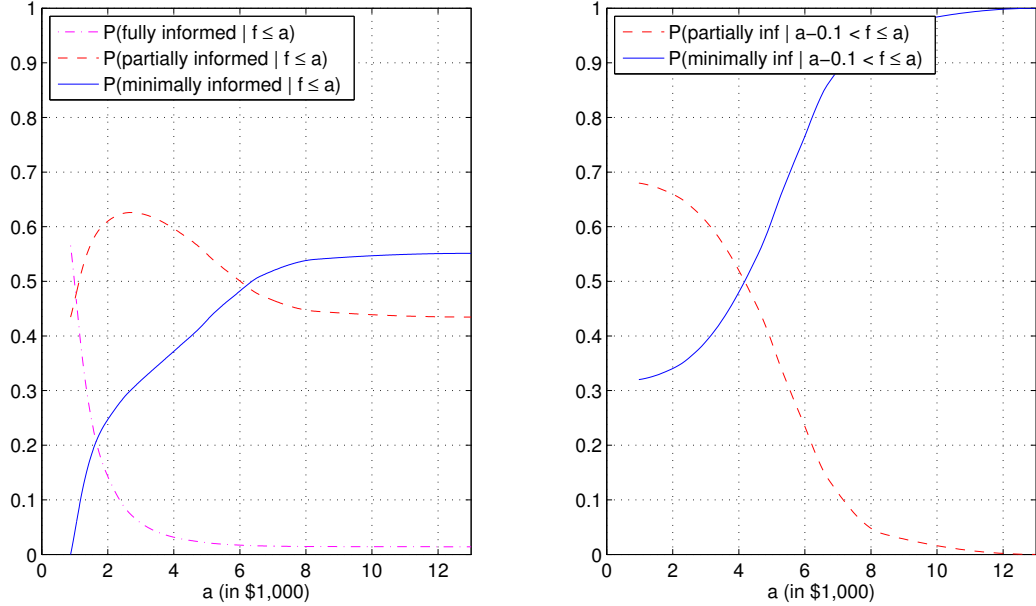
Figure D.1: **Ex-post distribution of borrower informedness given fee threshold** The left panel shows the ex-post distribution of borrower informedness given that broker fees do not exceed a certain threshold. The right panel shows the ex-post distribution of borrower informedness given that broker fees are within the interval $(a - \$100, a]$, for $a \in [\underline{f}, \overline{\nu})$. The estimates are based on the fee distribution for hybrid refinance loans described in Tables 5 and D.1. The state vector is the reference state vector in Table 4.

The right panel of Figure D.1 displays $\pi_1^{ex}(a - \$100, a)$ and $\pi_2^{ex}(a - \$100, a)$ as a function of a, again for hybrid refinance loans and the reference state vector. As observed fees increase, more and more likelihood is attributed to the borrower being minimally informed. Since fully informed borrowers always pay $\underline{f}$, $\pi_0^{ex}(a - \$100, a) = 0$ for all $a \in [\underline{f}, \overline{\nu})$.

References

Ambrose, B., J. Conklin, 2014. Mortgage Brokers, Origination Fees, Price Transparency and Competition. Real Estate Economics 41, 363–421.

Berndt, A., B. Hollifield, P. Sandas, 2012. What Broker Charges Reveal about Mortgage Credit Risk. Working paper, Carnegie Mellon University.

Bondell, H., B. Reich, H. Wang, 2010. Non-crossing Quantile Regression Curve Estimation. Biometrika 97, 825–838.

Engel, K., P. McCoy, 2011. The Subprime Virus: Reckless Credit, Regulatory Failure, and Next Steps. Oxford University Press, Oxford, England.

Glasserman, P., 2004. Convergence of Probability Measures. Springer, New York, NY.

ICF Macro, 2009. Design and Testing of Truth-in-Lending Disclosures for Closed-end Mortgages. Available at www.federalreserve.gov.

Koenker, R., P. Ng, 2003. SparseM: A Sparse Matrix Package for R. Journal of Statistical Software 8, available at www.jstatsoft.org/v08/i06.

Lacko, J., J. Pappalardo, 2004. The Effect of Mortgage Broker Compensation Disclosures on Consumers and Competition: A Controlled Experiment. Federal Trade Commission Staff Report.

——— , 2007. Improving Consumer Mortgage Disclosure. Federal Trade Commission Staff Report.

Larson, W., 2010. Evaluating Alternative Methods of Forecasting House Prices: A Post-Crisis Reassessment. Working paper, George Washington University.

Woodward, S., 2008. A Study of Closing Costs for FHA Mortgages. U.S. Department of Housing and Urban Development Report.

Woodward, S., R. Hall, 2009. The Equilibrium Distribution of Prices Paid by Imperfectly Informed Customers: Theory and Evidence from the Mortgage Market. Working paper, Stanford University.

——— , 2010. Consumer Confusion in the Mortgage Market: Evidence of Less Than a Perfectly Transparent and Competitive Market. American Economic Review, Papers and Proceedings 100, 511–515.

Table 1: **Descriptive statistics** The table reports average statistics for brokered loans funded by New Century, by loan purpose and product type. Our data include 46,796 free-standing first-lien loans for single-unit primary residences originated between 2004 and 2006.

	Refinance		Purchase	
	Hybrid	Fixed	Hybrid	Fixed
No of loans ($\times$ 1,000)	23,246	11,565	9,727	2,258
No of brokers ($\times$ 1,000)	8,063	4,824	5,126	1,547
Loan and borrower characteristics				
Loan amount ($\times$ \$1,000)	185	176	167	145
Low documentation	0.37	0.23	0.51	0.32
LTV	0.79	0.75	0.87	0.82
Prepay penalty	0.86	0.92	0.85	0.93
Initial rate (%)	7.69	7.24	8.18	8.10
Rate margin for hybrid loans (%)	5.83	–	5.74	–
Discount points (% of loan amount)	0.07	0.23	0.04	0.24
Lender fees (% of loan amount)	0.55	0.60	0.64	0.74
FICO	578	617	597	612
DTI	0.39	0.39	0.39	0.38
Neighborhood characteristics				
Education	0.14	0.13	0.15	0.14
Broker competition	0.96	0.93	0.82	0.69
House price appreciation (%)	17	17	11	10
Property taxes (% of house price)	0.84	0.83	1.10	1.30
Price-to-rent ratio	15	15	14	14
Market interest rates				
2-year Treasury rate (%)	3.30	3.47	3.37	3.62
6-year Treasury rate (%)	3.98	4.09	4.03	4.18
Location and year dummies				
CA	0.31	0.35	0.16	0.16
FL	0.14	0.19	0.14	0.15
TX	–	–	0.12	0.25
West w/o CA	0.19	0.14	0.13	0.07
MidWest	0.17	0.12	0.22	0.20
South w/o FL & TX	0.16	0.15	0.17	0.12
NorthEast	0.04	0.05	0.06	0.04
2004	0.49	0.42	0.47	0.38
2005	0.34	0.38	0.31	0.31
2006	0.17	0.20	0.22	0.31

Table 2: **Broker charges** The table reports average broker fees, yield spread premia and total revenues per loan, by loan purpose and product type. Our data include 46,796 free-standing first-lien loans for single-unit primary residences originated between 2004 and 2006.

| | Refinance | | Purchase | | Refinance | | Purchase | |
	Hybrid	Fixed	Hybrid	Fixed	Hybrid	Fixed	Hybrid	Fixed
	Percent of loan amount				*Dollar per loan (×1,000)*			
Fees	2.41	2.51	1.94	2.02	3.963	3.885	2.796	2.534
YSP	0.99	0.59	1.24	0.84	1.752	0.975	1.905	1.130
Revenues	3.40	3.10	3.18	2.86	5.714	4.859	4.701	3.663

Table 3: **Input parameters for the benchmark cost of waiting** Unless otherwise noted, the chosen parameter values match those for the New York Times Rent versus Buy Calculator calculator, available at www.nytimes.com/interactive/2014/upshot/buy-rent-calculator.html?, which in turn relies on data from Moody's Analytics, the Federal Reserve Bank of St. Louis, and Miller Samuel Incorporated. Freddie Mac Cash-Out Refinance Reports are available at www.freddiemac.com/finance/refi_archives.html.

Variable	Estimate
Homeownership	
House price appreciation	Annualized rate based on AR(6) forecasting model, using quarterly FHFA All-Transaction Index data for Core Based Statistical Areas. See Table B.1.
Property tax	Average tax rate on owner-occupied housing by county. See Table B.1.
Renovation costs	0.5% of house price per year, then increasing with inflation
Maintenance costs	0.5% of house price per year, then increasing with inflation
Homeowner insurance	0.5% of house price
Cost of selling home	6% of house price
Capital gains exclusion	$500K, assuming borrower is filing jointly
Renting	
Monthly rent	House price divided by twelve times the price-to-rent ratio, which is available for large metropolitan areas. See Table B.1.
Rent deposit	One month of rent
Renter insurance	1.32% of rent
Additional inputs	
Age, rate and balance of existing loan	For refinance loans, the age of the existing loan is set to 2, 2.5 and 3 years in 2004, 2005 and 2006. The new-to-old-rate ratio is set to 0.84, 0.92 and 1.04 in 2004, 2005, and 2006. The balance of the existing loan divided by the size of the loan taken out at time zero is 0.870, 0.785 and 0.707 in 2004, 2005 and 2006. Source: Freddie Mac Cash-Out Refinance Reports.
Loan amount for loan taken out later	For purchase loans, the size of the loan taken our at time m is equal to the loan amount at time zero times the expected house price appreciation over the next m months. For refinance loans, it is equal to the loan amount at time zero.
Loan-to-value ratio for loan taken out later	For purchase loans, LTV is the same whether or not the loan is taken out now or later. For refinance loans, it may change slightly, depending on the house price appreciation.
Downpayment	For purchase loans, the downpayment is the product of the loan amount and one minus the loan-to-value ratio. For refinance loans, it is set to zero.
Interest rate on debt	Annual commercial bank interest rate on credit card plans, accounts assessed interest, as reported by the St. Louis Fed
Effective tax rate	15%. For the average subprime borrower, the combined annual income tends to range between $74K in 2004 and $86K in 2006 (Berndt, Hollifield and Sandas (2012)). Historical effective tax rates by income are available from the Congressional Budget Office (see, for example, www.nytimes.com/interactive/2012/01/18/us/effective-income-tax-rates.html).
Capital gains tax	15%
Rate of investment	T-month constant maturity Treasury rate available from the St. Louis Fed
Inflation	3% per year

Table 4: **Specifications of state vector** The table lists several scenarios for the state vector: (A) Reference scenario, (B) Large loan, (C) Educated neighbors, (D) High house price appreciation, and (E) Low price-to-rent ratio.

	(A) Reference	(B) Large loan	(C) Educated nbrs	(D) High HPA	(E) Low prc/rent
Loan amount (in $1,000)	200	300	200	200	200
Education (%)	15	15	25	15	15
House price appreciation (%)	15	15	15	25	15
Price-to-rent ratio	15	15	15	15	10

Additional parameters

Low documentation: 0, Prepay penalty: 1, LTV: 0.8, YSP: 1% of loan amount, Initial rate: 7%, Rate margin for hybrid loans: 5.5%, Discount points: 0% of loan amount, Lender fees: 0.5% of loan amount, Expected rate change: 0, Property taxes: 1%, Cash taken out (refi): 13% of loan amount, Age of existing loan (refi): 2 years, New-to-old-rate ratio (refi): 0.84, Interest on debt (refi): 13.22%, 2-yr Treasury rate: 2.5%, 6-yr Treasury rate: 3.5%, Location: CA, Year: 2004

Table 5: **Non-crossing quantile regressions** The table reports the results for the non-crossing quantile regressions of broker fees (as % of loan amount) for the median, for different strata of the data. Standard errors are provided in parentheses. The data include 46,796 free-standing first-lien loans for single-unit primary residences originated between 2004 and 2006.

	Refinance		Purchase		Purchase	
	Hybrid	Fixed	Hybrid	Fixed	Hybrid	Fixed
Constant	-0.143	0.527	0.319	1.055	0.288	0.635
	(0.123)	(0.125)	(0.236)	(0.281)	(0.235)	(0.306)
YSP (% of loan amt)	-0.368	-0.498	-0.188	-0.193	-0.191	-0.204
	(0.014)	(0.022)	(0.019)	(0.046)	(0.019)	(0.046)
Log loan amt (% of loan amt)	0.436	0.407	0.271	0.225	0.273	0.212
	(0.009)	(0.011)	(0.011)	(0.023)	(0.011)	(0.023)
Low documentation	-0.038	-0.078	-0.017	0.083	-0.018	0.052
	(0.019)	(0.031)	(0.026)	(0.064)	(0.026)	(0.064)
Prepay penalty	0.406	0.374	0.266	0.117	0.260	0.152
	(0.027)	(0.054)	(0.035)	(0.119)	(0.035)	(0.119)
Initial rate (%)	0.040	0.086	0.059	0.012	0.060	0.038
	(0.011)	(0.016)	(0.015)	(0.035)	(0.015)	(0.035)
Margin for hybrid loans (%)	0.194	–	0.050	–	0.042	–
	(0.022)		(0.042)		(0.042)	
Discount points (% of loan amt)	0.096	0.025	-0.059	-0.178	-0.056	-0.176
	(0.028)	(0.025)	(0.054)	(0.057)	(0.053)	(0.058)
Education (%)	-0.018	-0.013	-0.011	-0.012	-0.011	-0.012
	(0.001)	(0.002)	(0.002)	(0.004)	(0.002)	(0.004)
FL	-0.413	-0.218	-0.500	-0.250	-0.490	-0.216
	(0.039)	(0.047)	(0.052)	(0.116)	(0.052)	(0.116)
TX	–	–	0.162	0.329	0.222	0.570
			(0.060)	(0.116)	(0.065)	(0.143)
West w/o CA	-0.225	-0.145	-0.158	-0.134	-0.134	0.002
	(0.026)	(0.040)	(0.037)	(0.092)	(0.038)	(0.100)
MidWest	-0.374	-0.435	-0.121	-0.497	-0.072	-0.261
	(0.033)	(0.051)	(0.047)	(0.113)	(0.052)	(0.140)
South w/o FL & TX	0.030	-0.030	0.010	0.127	0.051	0.362
	(0.035)	(0.050)	(0.047)	(0.131)	(0.052)	(0.151)
NorthEast	-0.253	-0.215	-0.013	-0.235	0.036	-0.009
	(0.048)	(0.061)	(0.068)	(0.165)	(0.072)	(0.181)
2005	0.108	0.218	0.042	0.134	0.037	0.101
	(0.021)	(0.029)	(0.031)	(0.071)	(0.031)	(0.070)
2006	-0.053	0.114	0.031	0.074	0.035	0.039
	(0.032)	(0.043)	(0.047)	(0.091)	(0.047)	(0.091)
b_k (% of loan amt)					0.031	0.118
					(0.015)	(0.041)

Table 6: **Implied borrower informedness and valuation distributions** The table summarizes our findings regarding the implied borrower informedness and valuation distributions. We use $E_{ne}(\min\{f_1,\nu\})$ and $E_{ne}(\nu)$ as shortcuts for $E(\min\{f_1,\nu\}|\text{non-expert borrower})$ and $E(\nu|\text{non-expert borrower})$, and $E_p(\gamma)$ as a shortcut for $E(\gamma|\text{partially informed borrower})$. Δ is the fraction of non-expert borrowers that are partially informed. Δ_1 is computed as $\Delta_{(A)}$, except that the perceived likelihood that a friend's fees do not exceed benchmark valuations is for the given state vector. Dollar values are in \$1,000. The results are based on the estimates in Tables 5 and D.1 through D.4. The state vector specifications are as in Table 4.

	(A)	(B)	(C)	(A)	(B)	(C)	(D)	(E)
		Refinance: Hybrid			*Purchase: Hybrid*			
$E(f)$	4.353	5.628	4.010	3.413	4.526	3.206	3.466	3.433
$b_{\widetilde{f}}$	4.688	6.132	4.688	3.643	4.871	3.643	3.714	3.664
b_k	0.266	0.407	0.266	2.941	4.411	2.941	4.779	3.644
π_2	0.55	0.49	0.39	0.04	0.03	0.03	0.01	0.02
$\underline{f}$	0.874	0.921	0.593	0.450	0.498	0.304	0.481	0.461
$\overline{E}_{ne}(\min\{f_1,\nu\})$	3.612	4.637	3.392	3.275	4.364	3.086	3.398	3.335
$E_{ne}(\nu)$	5.028	6.761	5.061	7.045	10.354	7.160	8.402	7.703
$E_p(\gamma)$	0.72	0.71	0.69	0.48	0.44	0.45	0.40	0.44
Δ	0.44	0.50	0.61	0.96	0.97	0.97	0.99	0.98
$\Delta - \Delta_{(A)}$	0.00	0.06	0.17	0.00	0.01	0.01	0.02	0.01
$\Delta_1 - \Delta_{(A)}$	0.00	0.05	0.00	0.00	0.01	0.00	0.03	0.02
		Refinance: Fixed			*Purchase: Fixed*			
$E(f)$	3.944	5.109	3.666	3.031	4.059	2.807	3.188	3.111
$b_{\widetilde{f}}$	4.211	5.512	4.211	3.274	4.424	3.274	3.449	3.355
b_k	0.293	0.446	0.293	3.128	4.692	3.128	4.505	3.832
π_2	0.54	0.49	0.41	0.04	0.03	0.03	0.02	0.03
$\underline{f}$	0.781	0.876	0.518	0.195	0.056	0.040	0.294	0.245
$\overline{E}_{ne}(\min\{f_1,\nu\})$	3.212	4.111	3.039	2.873	3.853	2.667	3.092	2.995
$E_{ne}(\nu)$	4.620	6.229	4.646	6.802	9.924	6.863	7.974	7.527
$E_p(\gamma)$	0.71	0.69	0.68	0.46	0.44	0.44	0.41	0.43
Δ	0.45	0.50	0.59	0.96	0.97	0.97	0.98	0.97
$\Delta - \Delta_{(A)}$	0.00	0.05	0.14	0.00	0.01	0.01	0.02	0.02
$\Delta_1 - \Delta_{(A)}$	0.00	0.05	0.00	0.00	0.01	0.00	0.03	0.02

Table 7: **Reduction in fees and γ if a partially informed borrower consults one extra friend** The table reports the reduction in expected broker fees (in $1,000) and in the expected value of γ in Equation (3) if a partially informed borrower were to consult one extra friend. The results are based on the estimates in Tables 5 and D.1 through D.4. The state vector specifications are as in Table 4. The Monte Carlo estimates are based on 100,000 scenarios.

	(A)	(B)	(C)	(A)	(B)	(C)	(D)	(E)
	Refinance: Hybrid			*Purchase: Hybrid*				
Reduction in fees	0.599	0.844	0.651	0.852	1.188	0.855	0.908	0.883
Reduction in γ	0.13	0.14	0.14	0.14	0.13	0.13	0.12	0.13
	Refinance: FRM			*Purchase: FRM*				
Reduction in fees	0.579	0.813	0.620	0.817	1.157	0.819	0.867	0.850
Reduction in γ	0.14	0.14	0.14	0.13	0.13	0.13	0.12	0.13

Table 8: **Expected broker fees if borrowers become more informed** The table reports expected broker fees $E(f|\pi(\theta))$ defined in Equation (18) (in \$1,000), for $\pi(\theta) = (\pi_0, \pi_1 + \theta\pi_2, (1-\theta)\pi_2)$ and $\theta \in \{0, 0.5, 1\}$. The results are based on the estimates in Tables 5 and D.1 through D.4. The state vector specifications are as in Table 4. The Monte Carlo estimates are based on 100,000 scenarios.

θ	(A)	(B)	(C)	(A)	(B)	(C)	(D)	(E)
	Refinance: Hybrid			*Purchase: Hybrid*				
0	4.353	5.628	4.010	3.413	4.526	3.206	3.466	3.433
0.5	3.964	5.110	3.687	3.343	4.445	3.146	3.432	3.384
1	3.573	4.590	3.364	3.272	4.361	3.084	3.397	3.334
	Refinance: FRM			*Purchase: FRM*				
0	3.944	5.109	3.666	3.031	4.059	2.807	3.188	3.111
0.5	3.562	4.590	3.340	2.951	3.956	2.737	3.140	3.053
1	3.179	4.070	3.013	2.870	3.850	2.664	3.090	2.993

Table 9: **Implied borrower informedness and valuations distributions for extended model**
The table summarizes our findings with regard to the implied borrower informedness and valuation distributions under the extended model assumptions. We use $E_p(\min\{f_1, \nu\})$ and $E_m(\nu)$ as shortcuts for $E(\min\{f_1, \nu\}|$partially informed borrower) and $E(\nu|$minimally informed borrower), and $E_p(\gamma)$ as a shortcut for $E(\gamma|$partially informed borrower). $E_p(\min\{f_1, \nu\})$ and $E_m(\nu)$ are in \$1,000. The results are based on the estimates in Tables 5 and D.1 through D.4. The state vector is the reference state vector in Table 4.

| | Refinance | | Purchase | |
	Hybrid	Fixed	Hybrid	Fixed
α	0.51	0.60	0.73	0.81
π_2	0.59	0.57	0.04	0.04
$E_p(\min\{f_1, \nu\})$	3.494	3.125	3.263	2.863
$E_m(\nu)$	5.025	4.617	6.973	6.741
$E_p(\gamma)$	0.68	0.67	0.47	0.46
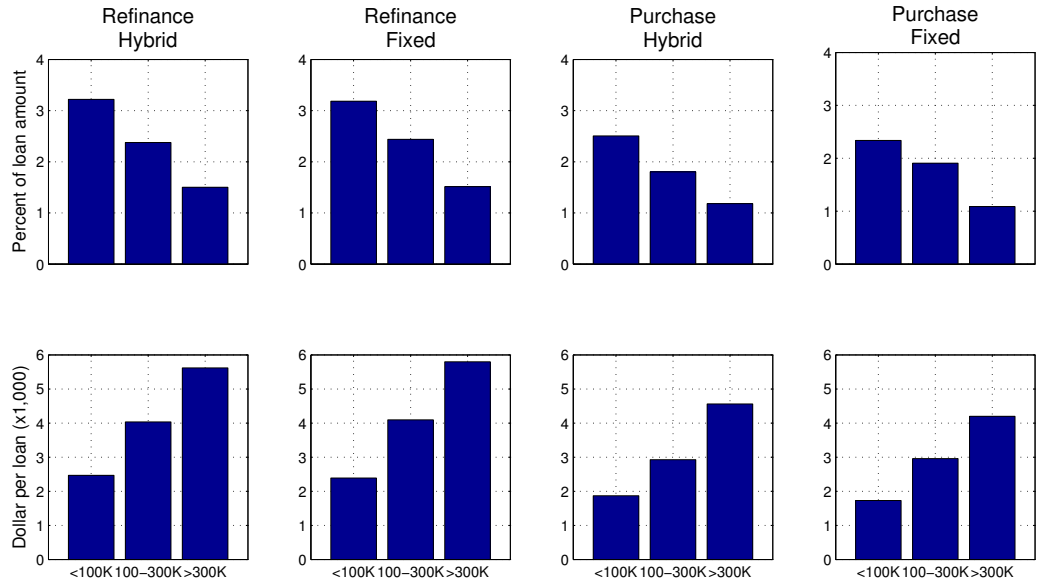
Figure 1: **Broker fees by loan amount** The top panel shows average percentage broker fees for loans of less than $100K (<100K), between $100K and $300K (100-300K) and more than $300K (>300K), by loan purpose and product type. The bottom panel shows average dollar fees per loan. Our data include 46,796 free-standing first-lien loans for single-unit primary residences originated between 2004 and 2006.
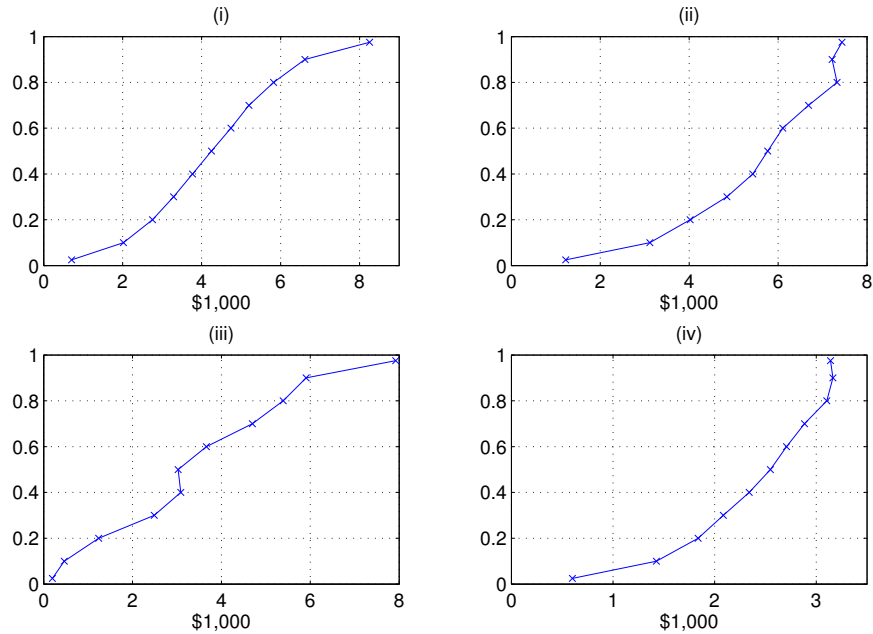
Figure 2: **Conditional quantile curves without non-crossing constraint** The figure shows the results of a classical quantile regression analysis, for $q = 0.025, 0.1, 0.2, \ldots, 0.9, 0.975$. The data include all 23,246 hybrid refinance loans. Panel (i) shows the quantile curve for the reference state vector in Table 4. In Panel (ii), we change the state vector by setting the rate margin to 7.5% and discount points to 4%. In Panel (iii), we change x by setting YSP to 2%, loan amount to $400K, prepay penalty to zero, rate margin to 3%, discount points to 3% and education to 5%. In Panel (iv), we change x by setting the loan amount to $50K, initial rate to 10%, rate margin to 7.5%, and discount points to 2%.
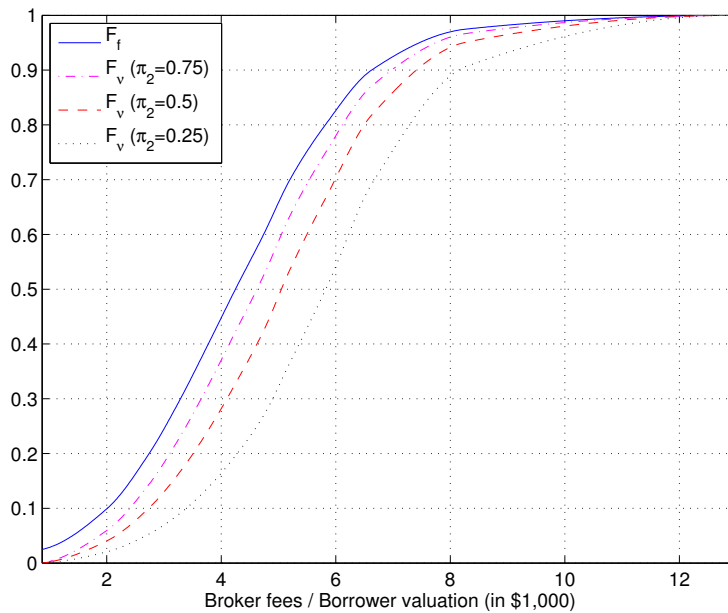
Figure 3: **Broker fee and π_2-dependent borrower valuation distributions** The figure shows the estimated broker fee distribution, and the associated borrower valuation distributions for different choices of π_2. The estimates are based on the fee distribution for hybrid refinance loans described in Tables 5 and D.1, and the reference state vector in Table 4.
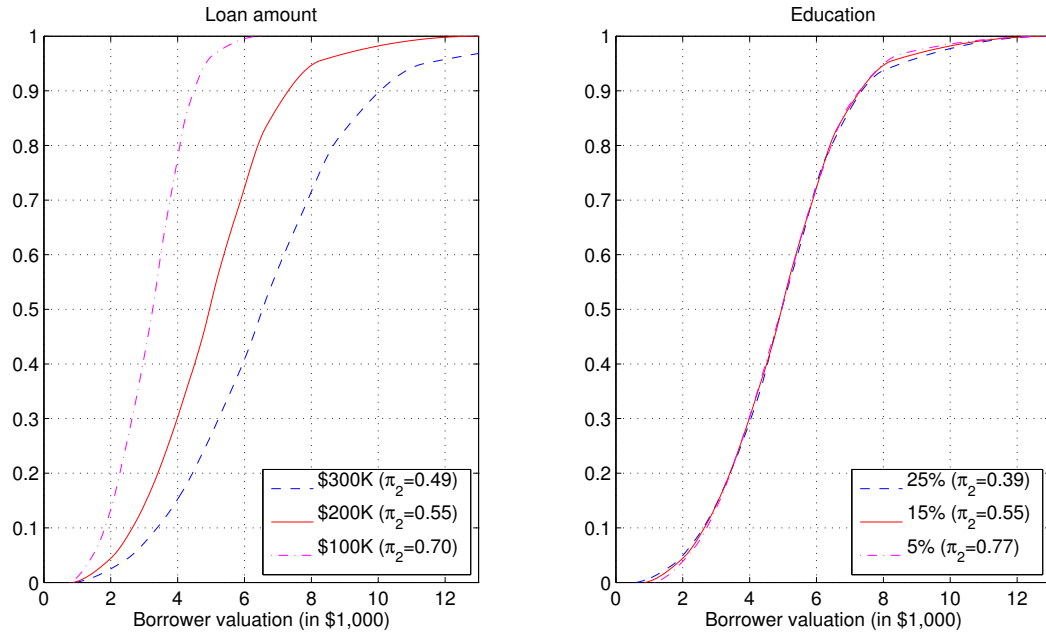
Figure 4: **Implied borrower valuation distributions** The left panel shows the conditional distribution of borrower valuations as a function of the loan amount. The estimates are based on the fee distribution for hybrid refinance loans described in Tables 5 and D.1. The state vector for the plots marked "$300K", "$200K" and "$100K" is equal to the large, reference and small loan scenarios in Tables 4 and B.2, respectively. Similarly, the right panel shows the conditional distribution of borrower valuations as a function of the fraction of neighbors with a B.A. degree. The state vector for the plots marked "25%", "15%" and "5%" is equal to the educated neighbors, reference and uneducated neighbors scenarios.
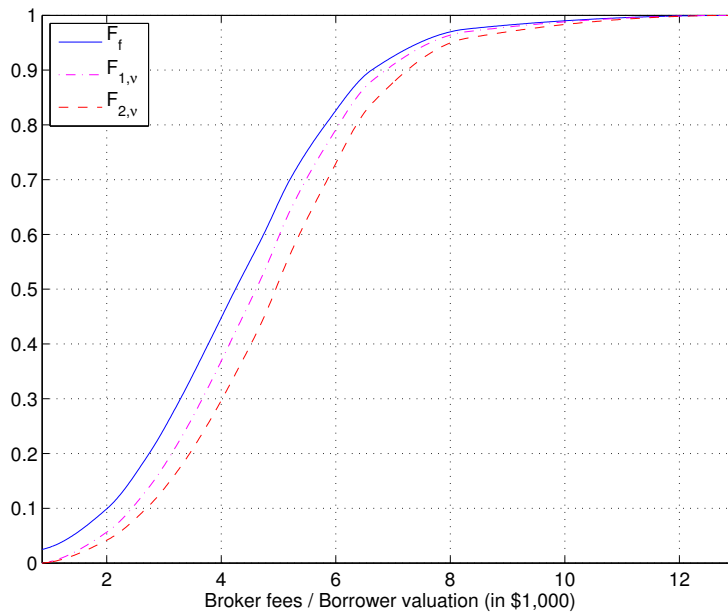
Figure 5: **Implied valuation distributions for partially and minimally informed borrowers** The figure shows the conditional distribution of borrower valuations for partially and minimally informed borrowers, $F_{1,\nu}$ and $F_{2,\nu}$. The estimates are based on the fee distribution for hybrid refinance loans described in Tables 5 and D.1. The state vector is the reference state vector in Table 4.

Earlier Working Papers:

For a complete list of Working Papers published by Sveriges Riksbank, see www.riksbank.se