

The Equilibrium Distribution of Prices Paid by Imperfectly
Informed Customers: Theory, and Evidence from the Mortgage
Market *

Susan E. Woodward

Sand Hill Econometrics

swoodward@sandhillecon.com

Robert E. Hall

Hoover Institution and Department of Economics,

Stanford University

National Bureau of Economic Research

rehall@stanford.edu; website: Google "Bob Hall"

June 19, 2009

Abstract

*A file containing the calculations will be available at Hall's website.

1 Introduction

Consumers often transact with imperfect information about the best price available for a product. Some examples are mundane—a grocery shopper does not know the price of kleenex at every other nearby grocery store. Others are more substantial, such as the best available price for a particular new car or the best expense ratio for an S&P 500-indexed mutual fund. We consider a large expenditure that the majority of consumers make every 5 or 10 years, the payment at the closing for a mortgage. These payments range from zero to \$30,000 for mortgages of normal size. They are called “origination fees” or “points” or a myriad of other terms; often the borrower pays for several different categories. Occasionally a borrower receives money back as negative “points.” Because consumers enter the mortgage market infrequently and because of features of the market that make it difficult to learn the best price, mortgage pricing is a leading example of a market where many consumers pay well above the best price.

Earlier research has shown that mortgage charges are higher for less-educated borrowers, members of minorities, borrowers who pay high interest rates, and those who borrow larger principal amounts. The research has not shown whether the borrowers paying higher charges did so because arranging the mortgage cost more or because those borrowers suffered exploitation of their lack of knowledge of the best available charge, which should be little higher than cost. We use a novel identification strategy to achieve that separation. We show that the groups of borrowers who pay higher charges are more costly to serve, but that part of the higher charges they pay go to mortgage brokers as profit rather than as compensation for higher cost.

Our findings are relevant for policies on mortgage disclosures. Borrowers with better information can bargain for lower charges, but only down to the level of cost. Disclosure policies should alert borrowers to the benefits of learning about the best deals available. We develop a model that demonstrates the potential amplification of the favorable influence of better-informed borrowers—if borrowers find out better deals from their friends, each borrower who shops more effectively conveys an external benefit on future borrowers who learn from the borrower about better deals.

We are concerned mainly with closing charges rather than with mortgage interest rates. The reason is that we lack an observable benchmark for the rate that a consumer should choose. Cash-constrained borrowers should pick a higher interest rate, as we demonstrate.

We develop an Edgeworth-box analysis of the bargain between a lender and a borrower. There is a unique interest rate that defines the location of the contract curve, but the cash paid at closing is a lump sum that depends on the bargaining powers of the two parties. Our analysis is the standard one for two-part pricing, adapted to the specifics of the mortgage transaction.

We focus on mortgages arranged through brokers, because the wholesale value of a mortgage is known for these mortgages. We take the minimum upfront payment for a mortgage to be the broker's transaction cost of around \$2,400 less the difference between the proceeds from the loan and its wholesale value. The broker receives that difference as the so-called yield-spread premium, which is often the majority of the broker's compensation. Thus, if a borrower pays \$1,500 as a closing payment, receives \$100,000 from the lender at the closing, and the wholesale value of the loan (the present value of the borrower's expected payments) is \$97,700, with a yield-spread premium to the broker of $\$100,000 - \$97,700 = \$2,300$, she has paid $\$1,500 + \$2,300 = \$3,800$ for a loan when she could have pressed the broker to do the loan for only \$2,400. A savvy borrower could have insisted that the broker charge her only \$100, which, together with the yield-spread premium of \$2,300, would have just covered his cost of \$2,400. The numbers in this example are representative of those in our data.

Our data come from a sample of mortgages guaranteed by the Federal Housing Administration (FHA) during a six-week period in 2001. The FHA guarantees mortgages of fairly creditworthy borrowers for modestly priced houses. The data report the amount that the borrower paid for mortgage-related charges at the closing, the principal amount that the lender paid into the closing, the yield-spread premium, and the interest rate for the mortgage. FHA mortgages are always 30-year fixed-rate instruments.

We start by estimating a model that divides the broker's reported revenue (borrower payment plus yield-spread premium) into cost and profit. The basic idea is simple—brokers won't make deals with borrowers below cost, but the best-informed borrowers can push the brokers close to cost. We characterize cost as a random variable fairly tightly and symmetrically distributed around a mean of about \$2,400 and profit as a random variable that is always positive and quite asymmetrically distributed above zero. This characterization of the two components of the total charge for a mortgage—cost and broker profit—identifies the components separately.

We interpret the distribution of profit within a model of the bargain between the broker

and the borrower. In the model, the broker makes a take-it-or-leave-it offer to the borrower after learning the borrower's reservation price. Naturally, the broker's offer is at the borrower's reservation price. Borrowers formulate their reservation prices in three ways. (1) They check with a few friends who have recently taken out mortgages and consider the minimum of the closing charges they paid, (2) they are influenced by the broker, whom they regard as well informed about mortgage terms but not necessarily a reliable source of information, and (3) a small fraction of borrowers are experts, who know how to determine the best available closing charge by shopping aggressively among brokers and other mortgage sources.

We show how the equilibrium distribution of closing payments depends on the number of friends consulted, the distribution of the influence of the broker, and the fraction of experts in the borrower population. Because borrowers learn from their friends, who are under the influence of brokers, as well as under the direct influence of brokers, the equilibrium amplifies the price-raising effect of the brokers. The payoff to a policy that reduces the effect of brokers on reservation prices is similarly amplified. Raising the fraction of experts in the borrower population also has a high payoff, as the equilibrium also amplifies their favorable influence.

2 Model of the Mortgage Transaction

The model presumes that the borrower has picked a house and the principal amount of a mortgage, P . The borrower bargains with one or more brokers over the interest rate, r , and the closing charge L , the amount the borrower has to pay in cash at closing for the mortgage. The closing charge can be negative. The payments for the mortgage are $p(r)$ in future periods t . All of the calculations are in nominal terms.

The borrower decides on the mortgage at time $t = 0$ and formulates a corresponding consumption plan c_t . The borrower's preferences are

$$U = \mathbb{E} \sum_{\tau} \beta^{\tau} u(c_{t+\tau}). \quad (1)$$

The borrower's cash holdings at the beginning of period t are A_t . Earnings and other cash receipts are y_t . The budget constraint is

$$A_1 = (1 + s)(A_0 + y_0 - c_0 - L) \quad (2)$$

and

$$A_{t+1} = (1 + s)(A_t + y_t - c_t - p(r)) \text{ for } t > 0. \quad (3)$$

We solve the budget constraint for consumption and substitute into the utility function:

$$U = u\left(A_0 + y_0 - L + \frac{A_1}{1+s}\right) + \mathbb{E} \sum_{t>0} \beta^t u\left(A_t + y_t - p(r) + \frac{A_{t+1}}{1+s}\right). \quad (4)$$

To derive the marginal (dis)utility of the closing charge L , we take the borrower's current and future savings, A_t , to be constant. We also take income y_t to be exogenous and constant. Then we interpret the derivative of U with respect to L to be the desired marginal utility. The Envelope Theorem implies that the assumption of constant A_t is appropriate and rigorous—it observes that rearrangements of consumption over time would have no effect on utility, because the borrower optimizes consumption within the budget constraint in the first place. Thus

$$\frac{\partial U}{\partial L} = -u'\left(A_0 + y_0 - L + \frac{A_1}{1+s}\right) = -u'(c_0) \quad (5)$$

and

$$\frac{\partial U}{\partial r} = -\mathbb{E} \sum_{t>0} \beta^t u'\left(A_t + y_t - p(r) + \frac{A_{t+1}}{1+s}\right) p'(r) = -\mathbb{E} \sum_{t>0} \beta^t u'(c_t) p'(r). \quad (6)$$

The competitive wholesale market values a mortgage at $P \cdot W(r)$. Here $W(r)$ is the amount that a wholesale lender pays for a loan per dollar of notional principal, in the sense that the lender delivers P to the loan closing and $Y(r) = P \cdot W(r) - P$, the yield-spread premium, to the broker; the sum of the two payments is $P \cdot W(r)$. The broker incurs cost k to handle the loan transaction, including the value of his own time. The broker's profit from the loan is

$$\pi(L, r) = P \cdot W(r) - P + L - k. \quad (7)$$

A bedrock principle of bargaining with full information is efficiency—the deal has the property that no alteration, with suitable transfer, could improve the outcomes of both parties. An efficient bargain maximizes the weighted sum of the objective functions of the parties:

$$\max_{L, r} J(L, r) = U(L, r) + \psi \pi(L, r) \text{ for any } \psi > 0. \quad (8)$$

The first-order conditions for the maximization are

$$\frac{\partial J}{\partial L} = \frac{\partial U}{\partial L} - \psi = 0 \quad (9)$$

and

$$\frac{\partial J}{\partial r} = \frac{\partial U}{\partial r} - \psi P \cdot W'(r) = 0. \quad (10)$$

Thus

$$\frac{\partial U/\partial r}{\partial U/\partial L} = P \cdot W'(r). \quad (11)$$

With an efficient bargain, the borrower equates her marginal rate of substitution between upfront charge L and later charge r to the tradeoff in the wholesale lending market. This finding has a standard interpretation in the literature on two-part pricing. An intermediary should not put a surcharge on the price of a supplier (double marginalization), but rather pass that price on to the customer and extract profit with a lump-sum charge. Here, the broker passes on the benefit available in the wholesale market from a higher interest payment r to the borrower and extracts profit in the form of the closing charge L .

From equations (5) and (6), we have

$$\frac{\mathbb{E} \sum_{t>0} \beta^t u'(c_t) p'_t(r)}{u'(c_0)} = P \cdot W'(r). \quad (12)$$

The left side of this equation is the borrower's personal present value of the incremental cost of a higher interest rate on the mortgage. The efficiency condition calls for the present value to equal the immediate benefit of the higher rate, in the form of more cash $P \cdot W(r)$ at the disposal of the borrower.

The marginal value in the wholesale market, $W'(r)$, is proportional to r for an interest-only loan with zero default probability and no pre-payment option. In that case, the lender receives payments proportional to r , the present value is correspondingly proportional to r , and $W'(r)$ is a constant, independent of r . For real-world mortgages, with amortization, positive default probabilities that rise with r , and pre-payment probabilities that also rise with r , the value of a mortgage is a more complicated function of r . We discuss the evidence on $W(r)$ in the next section. Because defaults and prepayments rise with r , $W'(r)$ is a declining function of r .

Equation (12) establishes one restriction in the L, r space, based on the principle of efficiency. In terms of an Edgeworth-box view of bargaining, it places the parties on their contract curve.

Note that efficiency does *not* require that the borrower receive an interest rate equal to the wholesale rate. As our examples will show, the satisfaction of the efficiency condition occurs at a rate above the wholesale rate for the typical liquidity-constrained borrower.

We illustrate the efficiency condition with a simple example. Suppose the mortgage is constant in real terms and lasts forever; the borrower is immortal as well. The financial

discount rate is β , so

$$W(r) = \frac{r}{1 - \beta} \quad (13)$$

and

$$W'(r) = \frac{1}{1 - \beta}. \quad (14)$$

The borrower consumes $c - (L - \bar{L})$ in period zero. \bar{L} is an amount of cash left after making the down payment on a house. Consumption in future periods is $c - (r - \bar{r})P$, where $\bar{r}P$ is income available for continuing housing expense, for example, from income currently spent on rent. The borrower's marginal rate of substitution from equation (12) is

$$\frac{\frac{1}{1-\beta}u'(c - (r - \bar{r})P)}{u'(c - (L - \bar{L}))}. \quad (15)$$

Suppose that the utility function has the constant-elastic form often assumed in applied work:

$$u(c) = \frac{u^{1-1/\sigma}}{1 - 1/\sigma} \quad (16)$$

with

$$u'(c) = c^{-1/\sigma}. \quad (17)$$

Then the efficiency condition is

$$\frac{[c - (r - \bar{r})P]^{-1/\sigma}}{[c - (L - \bar{L})]^{-1/\sigma}} = 1 \quad (18)$$

so

$$c - (r - \bar{r})P = c - (L - \bar{L}) \quad (19)$$

and

$$(r - \bar{r})P = L - \bar{L}. \quad (20)$$

The efficiency relation has slope $dL/dr = P$ in the L, r space.

The other restriction comes from the division of the surplus from the mortgage transaction. In the Edgeworth box, this relation determines the point on the contract curve. In one simple case, where the borrower has all the bargaining power, the relation assigns zero profit to the broker:

$$\pi(L, r) = P \cdot W(r) - P + L - k = 0. \quad (21)$$

Figure 1 shows the equilibrium with the example efficiency condition and surplus division. Parameter values are: Notional principal $P = \$100,000$, broker cost $k = \$2000$, market

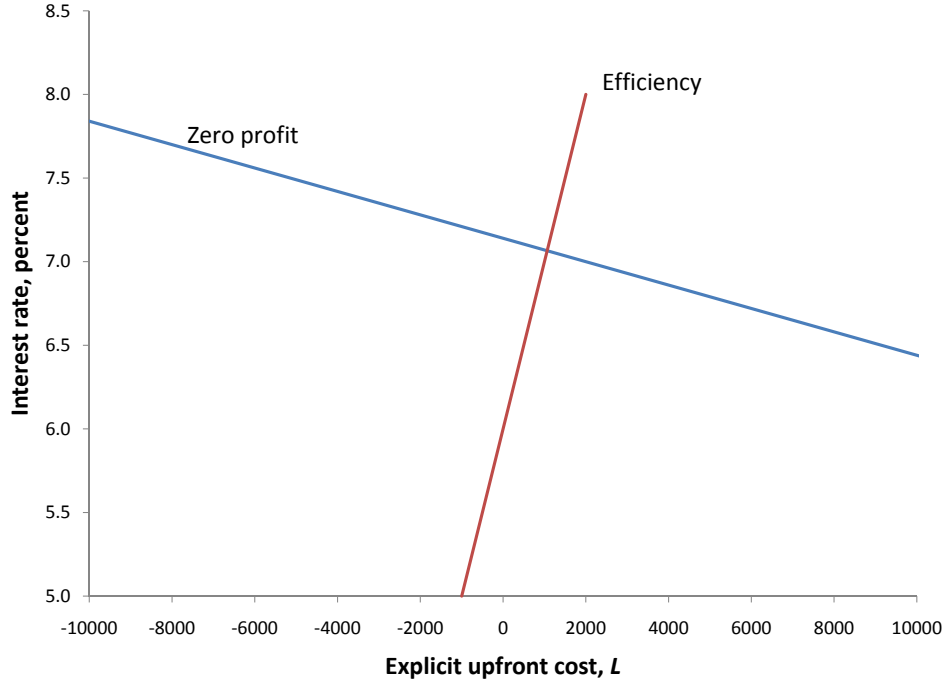


Figure 1: Equilibrium When the Borrower Has All the Bargaining Power and the Broker Makes Zero Pure Profit

discount $\beta = 0.93$, available savings $\bar{L} = \$1000$, available future income $\bar{r} = 7$ percent. The zero-profit line is quite flat—unless the interest rate is close to the rate in the wholesale market (here, $1 - \beta = 7$ percent), the broker makes or loses large amounts of money if the upfront payment L is in the range required by the efficiency condition.

In the equilibrium in Figure 1, the interest rate is 7.1 percent. The wholesale value of the mortgage per dollar of notional principal is $W(r) = r/(1 - \beta) = .071/.07 = 1.009$, so the mortgage is worth $P \cdot W(r) = \$100,935$ and the broker receives a yield-spread premium of $P \cdot W(r) - P = \$935$. The borrower pays the broker $L = \$1065$, so the broker's total earnings from the loan are \$2000, the same as his cost, $k = \$2000$.

The data show that borrowers often burden themselves with mortgages less favorable than the one described by the equilibrium in Figure 1. Sometimes the burden takes the form of a much higher value of the upfront payment L , which is occasionally above \$20,000 in our FHA sample. More often, the borrower agrees to an interest rate well above the wholesale rate. The extra payment can be measured as the broker's pure profit $\pi(r, L)$, the broker's compensation from the borrower, L , plus the yield-spread premium $P \cdot W(r) - P$ less the cost k . Figure 2 shows the broker's profit as iso-profit lines in the L, r space.

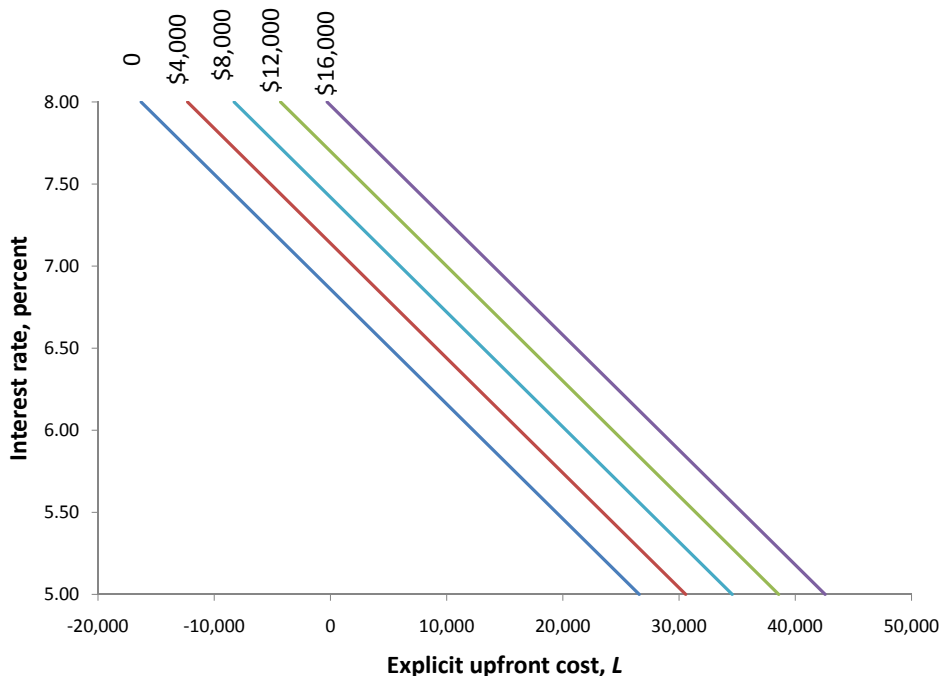


Figure 2: Lines Showing Combinations of the Upfront Charge and Interest Rate Providing Given Levels of Profit to the Broker

3 Data

Table 1 describes the relevant variables in our sample of 1,525 FHA brokered loans. Interest rates are fairly tightly clustered around 7 1/2 percent. Principal is generally around \$100,000 and rarely exceeds \$200,000. The total charges—cash closing charge plus yield-spread premium—average just above \$4,000 but have substantial dispersion. The cash component is typically a little under half of the total charge and the YSP a little more than half. The fractions of the borrowers who are members of minorities are close to the U.S. average in the population, at 11 percent African-American and 13 percent Latino. The last statistic is the fraction of the adult population in the borrower’s census tract who hold a BA degree. The average BA density is 21 percent.

Table 2 shows the distribution of the loans by interest rate. The 7 1/2 percent loan is by far the most popular.

Figure 3 shows the distribution of total closing charge for the 7 1/2-percent loans. The figure illustrates a property of the distribution that is the key to our identification strategy for separating cost from profit when we observe only the sum. The red diamonds show the actual distribution of total charge (cash plus YSP), in bins \$1,000 wide. The distribution

	<i>Mean</i>	<i>Standard deviation</i>
Interest rate, percentge points	7.40	0.45
Principal, dollars	111,269	40,051
Total closing charge, dollars	4,024	2,309
Cash charge, dollars	1,762	1,716
Yield-spread premium, dollars	2,262	1,381
Percent African-American	11	
Percent Latino	13	
Prcent white	76	
Percent of neighbors with BA degrees	21	12

Table 1: Descriptive Statistics for Brokered Loans

<i>Interest rate</i>	<i>Percent of sample</i>
Below 7	7.2
7	14.0
7 1/8 to 7 3/8	20.9
7 1/2	36.0
7 5/8 to 7 7/8	11.6
8 and higher	10.4

Table 2: Frequency Distribution of Interest Rates

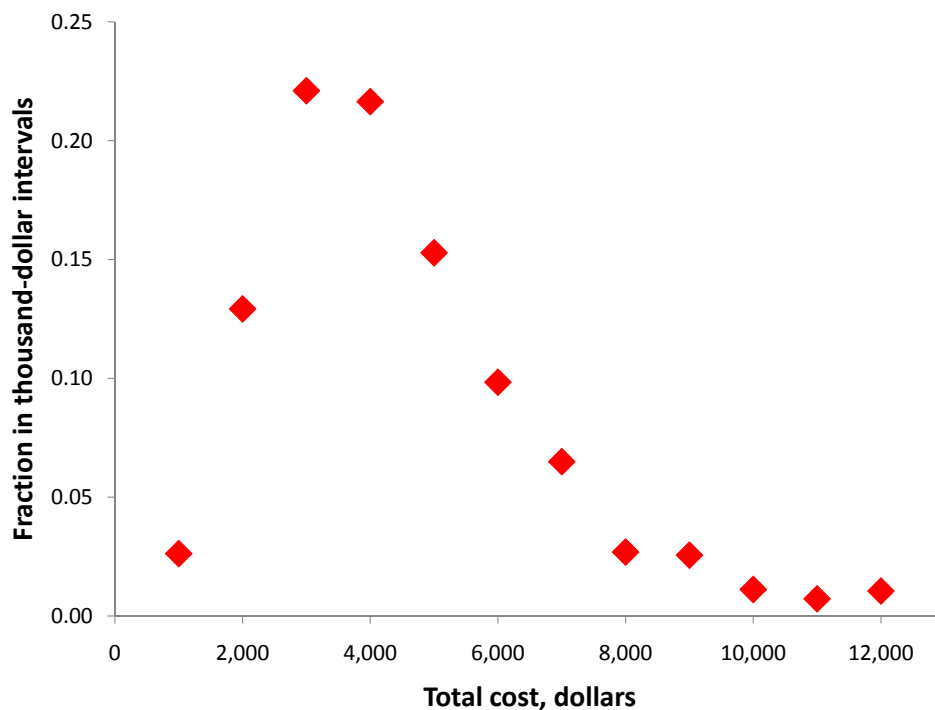


Figure 3: Distribution of Total Charge

reaches a conspicuous peak in the \$2,000 to \$2,999 bin.

4 The Yield-Spread Premium

As we discussed in the previous section, the yield-spread premium is the difference between the market value of a mortgage and the amount of the principal that the lender delivers at closing. The lender pays the premium to the broker outside the closing. The determination of the market value of a mortgage is a complex financial calculation because it depends on the borrower’s propensity to pay the mortgage off before maturity, either because of a sale of the property or because of refinancing. Most mortgages grant the borrower a free or low-cost option to pre-pay before maturity. The higher the interest rate on a mortgage, the higher is the likelihood of pre-payment.

To our knowledge, the yield-spread premiums that wholesale lenders offer to mortgage brokers have never been tabulated in a systematic way. They do not appear to be aggregated and published by any organization. Brokers receive daily rate sheets from lenders laying out the premiums in matrix form. One force behind the sudden explosion of independent brokers in the 1990s was the proliferation of inexpensive fax machines that made it practical for a

<i>Interest rate in category, percent</i>	<i>Average rate in category, percent</i>	<i>Average yield-spread premium per \$100 principal</i>	<i>Standard error</i>
Below 7	6.33	1.31	0.06
7 to 7 3/8	7.17	1.33	0.03
7 1/2	7.50	2.32	0.03
7 5/8 to 7 7/8	7.78	2.68	0.06
8 and higher	8.11	3.18	0.09

Table 3: Average Yield-Spread Premiums in the FHA Data, by Interest Rate

lender to communicate mortgage terms to thousands of brokers.

One source of evidence about yield-spread premiums comes from the disclosure of the premium on the HUD-1 form that federal law requires a broker to give to a borrower before closing a mortgage. This source is potentially imperfect, because there seems to be relatively little monitoring of the accuracy of a broker’s disclosure and the broker may try to conceal a large premium from a borrower. Table 3 shows the average yield-spread premium for the brokered loans in our sample, by interest rate. We place the interest rates in five categories.

The other source of information is a group of rate sheets from a dozen large lenders for May 2000. A rate sheet tells brokers the premium that a lender will pay depending on the loan’s interest rate and on the duration of the period that the offer of a loan at the specified terms remains open (the lock period). Figure 4 gives averages for May 31, 2000, for a 30-day lock period. The error bars are one standard error—they become larger at both ends of the line because fewer lenders quoted premiums so far from the popular interest rates. The figure also shows the premiums from Table 3 and the theoretical premium if all mortgages prepaid at 7 years despite their 30-year notional maturity and were valued by the standard present-value formula for a constant interest rate of 6.5 percent.

The curve in Figure 4 for the rate sheets lies to the right of those in the FHA data because mortgage rates were higher in general in 2000. The slope of the premium is generally lower in the FHA data, especially for the highest and lowest interest rates. The difference in the slope could reflect changes in expectations about pre-payments and changes in the slope of

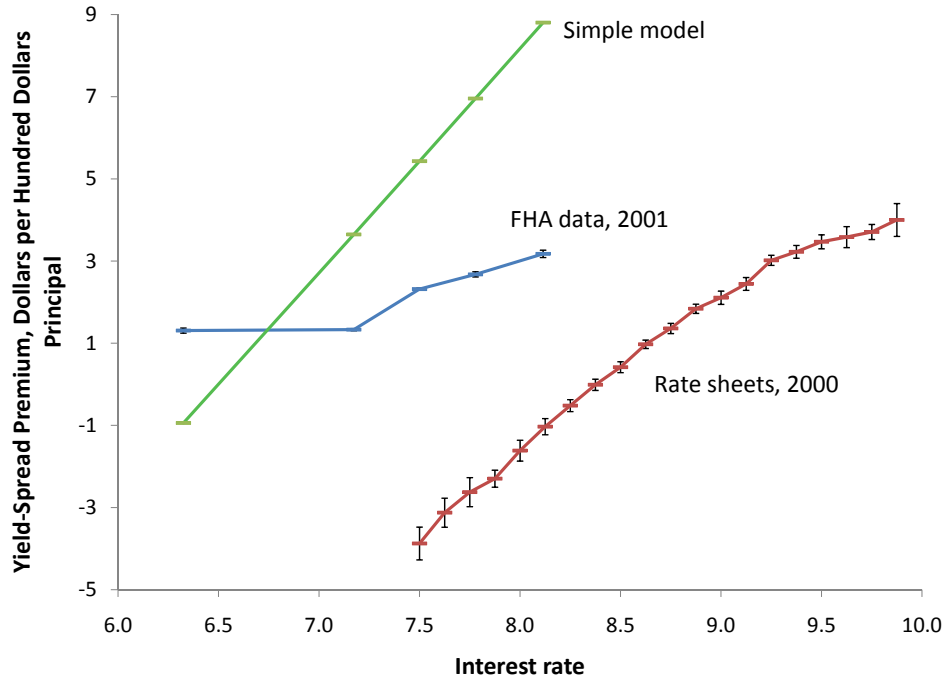


Figure 4: Yield-Spread Premiums in Rate Sheets in 2000, in the FHA Data, and in a Simple Model

the yield curve between 2000 and 2001. It could also reflect some tendency for brokers to understate their actual premiums when they are high, though why they should overstate premiums when they are low or negative is unclear.

Both the rate-sheet data and the FHA data make it clear that actual yield-spread premiums are not nearly as responsive to the interest rate as suggested by the simple exogenous pre-payment model. Lenders severely discount high-rate mortgages because they pre-pay sooner and possibly for other reasons. Although we believe that the wholesale lending market is substantially competitive, lenders may still be able to capture part of the rather large amount that borrowers typically pay for high-rate mortgages, rather than passing all of it on to the broker. We do not pursue the issue of the determination of the yield-spread premium in this paper. We take the premium as a given feature of the market. Brokers are price-takers with respect to lenders, with the price stated in the rate sheets.

Because the yield-spread premiums reported in the FHA data are not obviously at odds with those in the rate sheets, and because we are unable to adjust the curve from 2000 to improve measurement of the premiums actually paid for the FHA mortgages, we will accept the reported premiums for the rest of our analysis.

5 Estimation of the Distributions of Costs and Broker Profits

A broker receives revenue

$$\tau = L + Y(r) \tag{22}$$

from arranging a mortgage at rate r . In this section, we take the interest rate as given, and examine these variables separately by interest-rate categories, so we usually suppress r in our notation. We break the revenue τ into two components: k , the broker's cost, and z , the broker's profit or share of the surplus:

$$\tau = k + z \tag{23}$$

Our hypothesis is that k is symmetrically distributed around a mean κ and z is non-negative with an asymmetric distribution. Specifically, we take k to be normal with mean κ and standard deviation σ and z to be exponential, with mean $1/\lambda$.

The log-likelihood of a sample of observations on τ is

$$\mathcal{L} = \sum_i \log h(\tau_i), \tag{24}$$

where h is the density of τ , the convolution

$$h(\tau) = \int_{-\infty}^{\infty} f(k)g(\tau - k)dk \tag{25}$$

and f and g are the densities of k and z . These are

$$f(k) = \frac{1}{\sigma(2\pi)^{1/2}} \exp\left(-\frac{(k - \kappa)^2}{2\sigma^2}\right) \tag{26}$$

and

$$g(z) = \lambda \exp(-\lambda z). \tag{27}$$

We treat σ and θ as parameters and κ and λ as functions of a row vector of observable characteristics, x , of the borrower:

$$\kappa = x\gamma \tag{28}$$

and

$$\lambda = -x\beta, \tag{29}$$

where γ and β are column vectors of parameters analogous to coefficients of the regression of broker cost on the x and the coefficients of an exponential model fitted to profit data. We

<hr/> <hr/>		
σ : standard deviation of broker cost		0.85 (0.04)
<hr/>		
<i>Determinant</i>	<i>Broker cost</i>	<i>Broker revenue in excess of cost</i>
<hr/>		
Intercept	2.83 (0.16)	-0.60 (0.06)
Interest rate	0.73 (0.10)	0.32 (0.05)
Principal	1.80 (0.11)	0.39 (0.03)
White	-0.26 (0.14)	-0.15 (0.04)
Latino	0.07 (0.17)	-0.09 (0.05)
Fraction of neighbors with BAs	-1.24 (0.35)	-0.17 (0.15)
<hr/> <hr/>		

Table 4: Estimation Results

put the minus sign in the formula for λ so that higher values of x with positive values of β shift the distribution of z toward higher values.

We estimate the parameters σ , θ , γ , and β by maximum likelihood. We estimate the covariance matrix of the estimates as the inverse of the Hessian matrix of the log-likelihood.

For the borrower and loan characteristics, x , we use an intercept, the loan's interest rate, indicator variables for Latino and white borrowers, the fraction of neighbors with BAs, and the amount of the principal of the loan, as a ratio to \$100,000. Table 4 shows the results for the 1,525 brokered loans.

Table 5 describes five cases that we use to illustrate the implications of the results. Each case perturbs the base specification along one dimension of the explanatory variables. We do not include a case for a Latino borrower because our results show little difference between African-Americans and Latinos.

Our estimates of both cost and the broker's profit rise with the interest rate. Second line from the bottom shows the implied value of the mean of profit, calculated as the reciprocal of the exponential slope λ .

	Case				
	1	2	3	4	5
	<i>Base</i>	<i>White borrower</i>	<i>High interest rate</i>	<i>Highly educated neighbors</i>	<i>High principal</i>
Interest rate, percent	7 1/2	7 1/2	8	7 1/2	7 1/2
Principal, dollars	100,000	100,000	100,000	100,000	200,000
Race	African-American	White	African-American	African-American	African-American
Percent of neighbors with BA degree	21	21	21	45	21
Estimated broker cost, dollars	2,568	2,307	2,932	2,271	4,367
Estimated closing payment to broker above cost	1,570	1,275	2,091	1,478	4,114
Estimated total charge	4,138	3,583	5,023	3,749	8,481

Table 5: Five Illustrative Cases

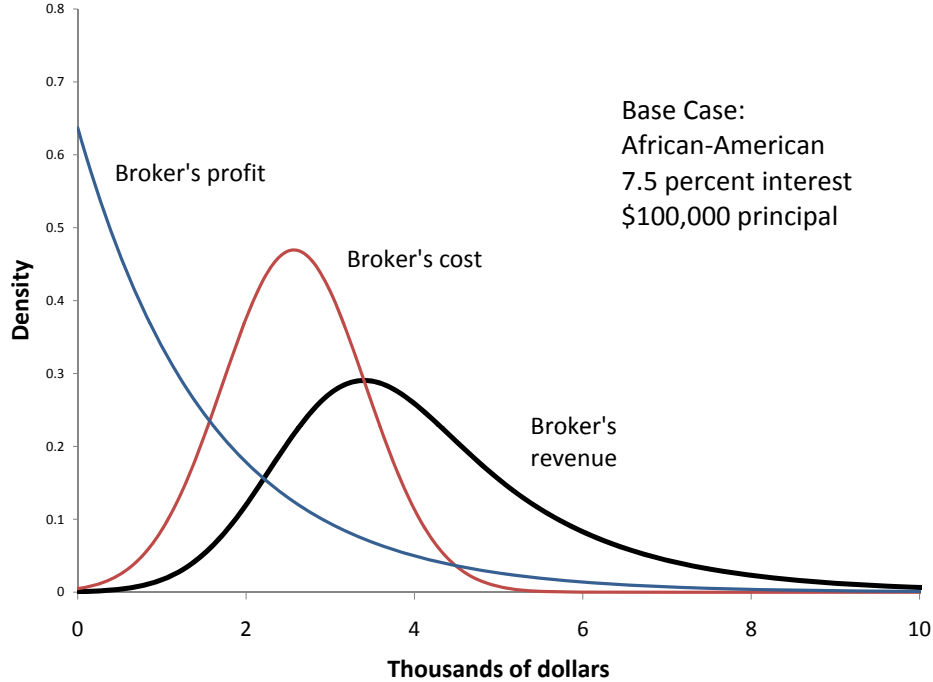


Figure 5: Distributions of Broker's Cost, Profit, and Revenue for the Base Case

Figure 5 through Figure 9 plot the estimated distributions for the five illustrative cases.

6 Model of the Equilibrium Distribution of the Closing Charge

This section considers a simple model of the equilibrium distribution of the closing charge, L . Recall that L is the amount of cash the borrower pays the broker at the closing, not including the yield-spread premium the broker receives outside the closing from the lender. We let $\bar{L} = k - Y(r)$, the minimum possible value of L that just covers the broker's cost (the broker receives L from the borrower and $Y(r)$ from the lender and the two together cover at least the cost, k).

We start with a simple model of bargaining: The broker learns the borrower's reservation value of L and offers the borrower a loan with that level of closing charge, provided that L and the yield-spread premium cover the broker's cost: $L \geq \bar{L}$. The broker has all the bargaining power. By this, we mean only that the surplus associated with the broker's contact with the borrower goes entirely to the broker. We do not mean that the broker can necessarily appropriate much of the surplus from the underlying house purchase or mortgage

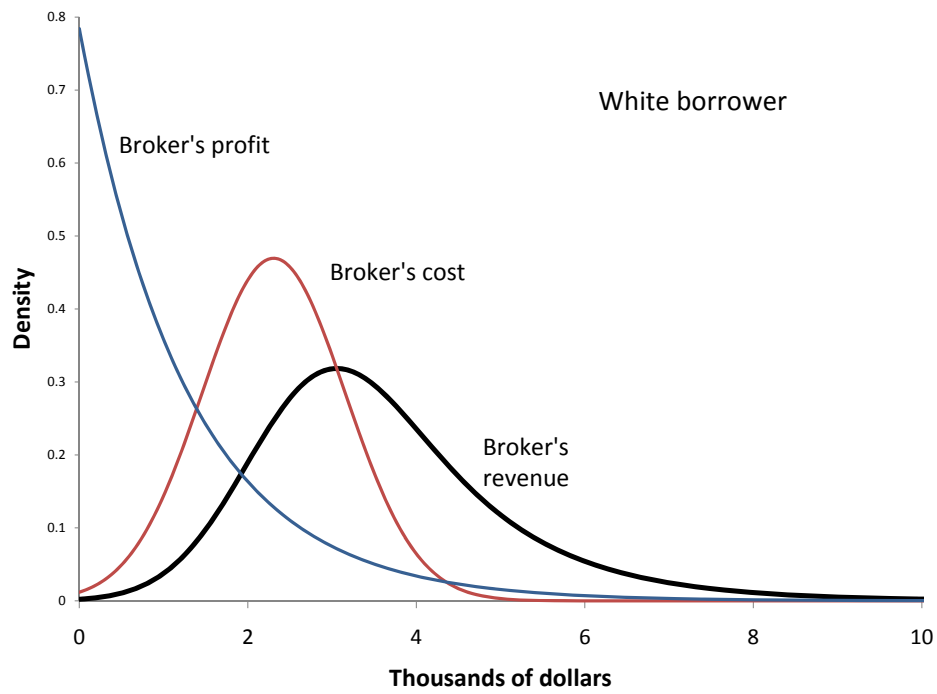


Figure 6: Distributions of Broker's Cost, Profit, and Revenue for a White Borrower

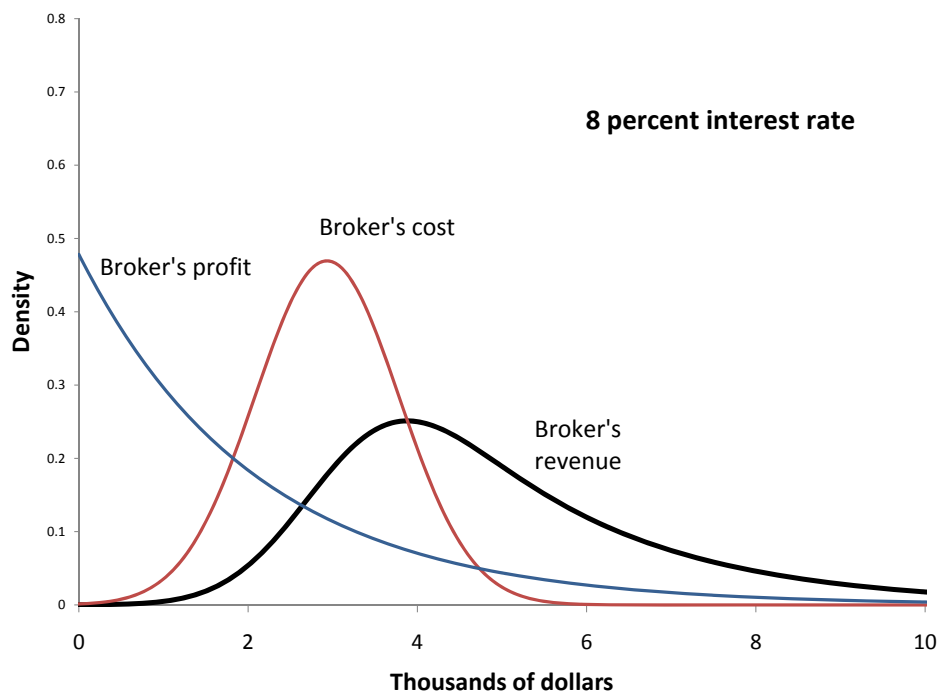


Figure 7: Distributions of Broker's Cost, Profit, and Revenue for a Higher Interest Rate

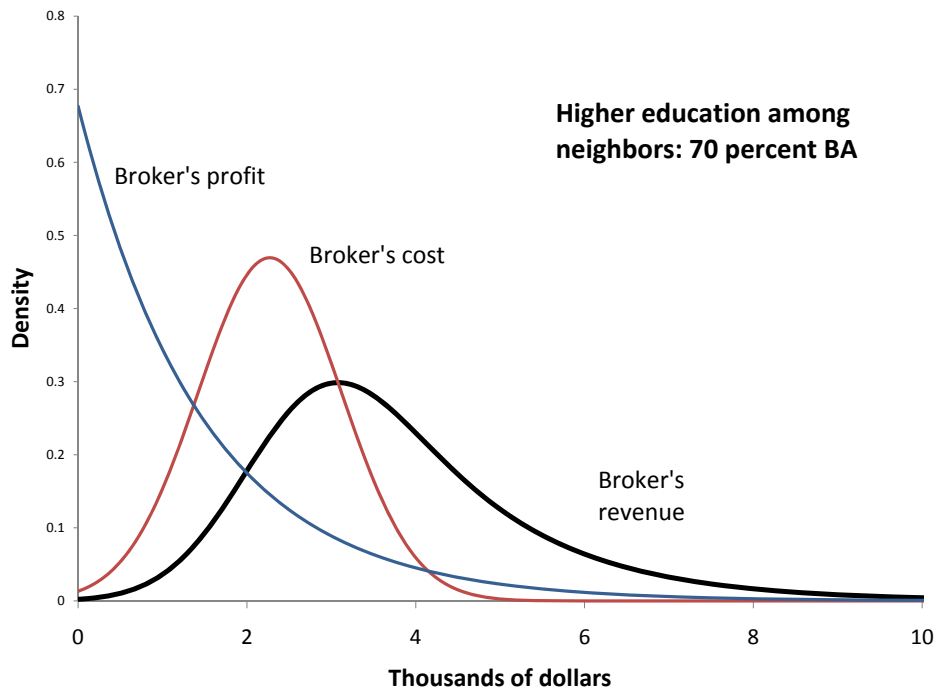


Figure 8: Distributions of Broker's Cost, Profit, and Revenue for a Borrower in a Better-Educated Neighborhood

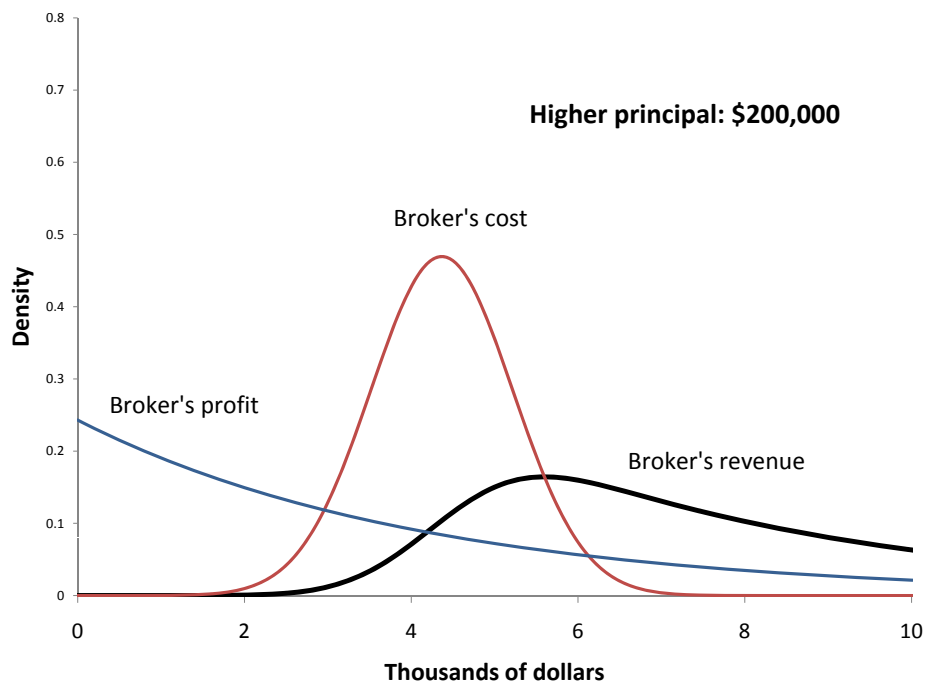


Figure 9: Distributions of Broker's Cost, Profit, and Revenue for a Larger Loan

refinance. A savvy borrower will have a low reservation value for L because she knows that good deals may be available by shopping other brokers.

Next we construct a simple but potentially realistic model of the way that borrowers determine their reservation values. Most borrowers (a fraction $1 - \chi$) gather information from two sources: (1) Friends who are entirely honest in providing information but who know only the closing charge L_i they paid in a recent mortgage transaction at the same interest rate, and (2) a broker, whom they regard as better informed but not necessarily trustworthy. These borrowers consult N friends and one broker and forms a reservation value as

$$L = \min_{i \leq N}(L_i) + b. \quad (30)$$

Here $b > 0$ is a random variable representing the upward effect of the broker on the reservation value. Picking the minimum of the amounts paid by several friends is a natural response, because the borrower knows that all of them are acceptable to the broker and that the friends who paid more than the minimum paid more than they needed to.

The remaining borrowers, a fraction χ , are experts in the mortgage market. Their reservation values are the minimum feasible, \bar{L} , which is the amount they pay for closing charges.

We assume that the friends of the non-expert borrowers are random draws from the population of recent borrowers and that no change in mortgage-market conditions has occurred since they borrowed. We look for the equilibrium distribution of reservation and transaction values. An equilibrium distribution is one where the distribution of reservation values L (and subsequent transaction values) is the same as the distribution among the earlier borrowers. If $F(L)$ is the distribution among the earlier borrowers and $\tilde{F}(L)$ is the distribution among current borrowers induced by their method for gathering information, then an equilibrium occurs when

$$\tilde{F}(L) = F(L) \text{ for all } L. \quad (31)$$

To avoid some mathematical complications, we will take the distribution of the broker influence b to be discrete, with probabilities π_n on a grid of equally spaced values $b_n = \bar{L} + n\delta$ with $n \in [0, 1, \dots, M]$. The probability that $\min(L_i) + b_n$ is less than L is the probability that $\min(L_i)$ is less than $L - b_n$, which is $1 - (1 - F(L - b_n))^N$. Let

$$I(L) = 0 \text{ for } L < \bar{L} \quad (32)$$

$$= 1 \text{ for } L \geq \bar{L}, \quad (33)$$

the distribution among experts. The distribution of induced reservation values is the mixture of the distributions,

$$\tilde{F}(L) = \chi I(L) + (1 - \chi) \sum_n \pi_n [1 - (1 - F(L - b_n))^N]. \quad (34)$$

The simplest equilibrium occurs when $N = 1$, $b = 0$, and $\chi = 0$. New borrowers replicate the distribution among their friends. Because a borrower does not pick the lowest value among those reported by several friends, there is no downward pressure from that source. And by assumption there is no upward pressure from broker persuasion. Any distribution is self-replicating—the model imposes no restrictions on the equilibrium distribution of L . As we will discuss shortly, indeterminacy of the equilibrium occurs in all cases where there are no experts in the population. Friends and the broker exert pressure relative to the overall level, but do not pin down the level.

We express analytical findings about the equilibrium distribution in a series of propositions. Their proofs reside in the appendix. The first proposition says that any heterogeneity in the equilibrium distribution arises from the heterogeneity in the broker’s influence:

Proposition 1 *With experts ($\chi > 0$), the equilibrium distribution is discrete; it has positive probability on the grid $\bar{L} + n\delta$ and zero density elsewhere.*

The second proposition establishes that the smallest positive influence of experts is enough to pin down the equilibrium distribution:

Proposition 2 *The equilibrium distribution of closing payments L is unique provided there are experts: $\chi > 0$*

The next proposition describes the indeterminacy when there are no experts at all; though the shape of the distribution is determinate, its central tendency is indeterminate:

Proposition 3 *With no experts, if F is an equilibrium distribution, so is another distribution in which all values of L are higher by δ : the distribution \hat{F} with $\hat{F}(\bar{L}) = 0$ and $\hat{F}(\bar{L} + n\delta) = F(\bar{L} + (n - 1)\delta)$ is also an equilibrium distribution.*

The number of friends consulted in forming a borrower’s reservation value controls the downward force from that source. As the number becomes larger and larger, the equilibrium distribution shrinks toward the minimum, \bar{L} . The shrinkage is limited by the direct effect of the broker and enhanced by the presence of experts:

Proposition 4

$$\lim_{N \rightarrow \infty} F(L) = \chi + (1 - \chi)\Pi(L - \bar{L}), \quad (35)$$

where Π is the cumulative distribution of the broker influence, π .

Next we consider equilibrium without any upward broker influence. With a positive fraction of experts, no matter how small, and no upward pressure from brokers, the unique equilibrium gives every borrower the full benefit of the experts' knowledge. The intuition of this proposition is pretty clear: If new borrowers tend to get better deals than the friends they consult, and those friends have not all received the same deal, their distribution cannot be the same as the earlier distribution—it must shift toward lower values. There could be an equilibrium above \bar{L} where every borrower pays the same L , but that is ruled out by the presence of experts:

Proposition 5 *If $b = 0$, $N > 1$, and $\chi > 0$, the equilibrium distribution of L places all probability at the single value \bar{L} : $F(\bar{L}) = 1$.*

The following result is a curiosity about equilibrium without experts. It rules out an equilibrium unless the distribution of the broker effect gives a substantial probability to zero effect. By contrast, with even the smallest fraction of experts, as Proposition 2 demonstrates, equilibrium is possible for any distribution of the broker effect.

Proposition 6 *With no experts, equilibrium is impossible unless $\pi_0 \geq 1/N$.*

From this point on, we will consider the equilibrium under all three forces: Borrowers check with more than one friend and so benefit from taking the minimum ($N > 1$), brokers have some influence ($\pi_0 < 1$), and some borrowers are experts ($\chi > 0$). To take the simplest example, suppose $N = 2$, the minimum number of consultations with friends to give a downward effect, and only two values of the broker's persuasive effect, $b = 0$ with probability π and $b = \delta$ with probability $1 - \pi$. Then

$$x_0 = \chi + (1 - \chi)\pi[1 - (1 - x_0)^2], \quad (36)$$

which has solution

$$x_0 = \frac{2\pi - 1}{\pi}, \quad (37)$$

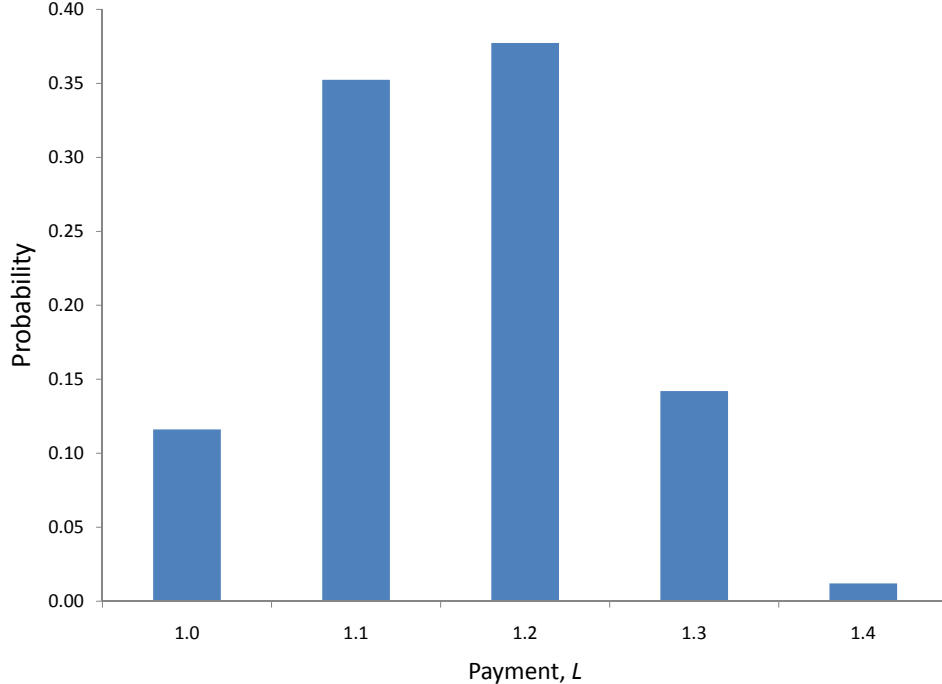


Figure 10: Example of Equilibrium Distribution of Closing Charge

the fraction of borrowers who get the best possible deal with $L = \bar{L}$. To explain the data showing that only a small fraction of borrowers receive such a good deal, we need a value of π just above one half. The rest of the distribution follows from the difference equation

$$x_i = \pi[1 - (1 - x_i)^2] + (1 - \pi)[1 - (1 - x_{i-1})^2]. \quad (38)$$

Figure 10 shows the equilibrium distribution of L for $\bar{L} = 1$, $\delta = 0.1$, and $\pi = 0.51$.

Figure 11 through Figure 15 shows the values of the distribution of broker influence, π , corresponding to the distribution of closing payments about the minimum, $L - \bar{L} = \tau$, for the five cases in Table 5. We calculated the distribution by solving equation (34) for π given the points on the cumulative distribution of τ , x_i . We used the estimated exponential distribution from the previous section, Table 4, for the distribution of τ .

All the distributions assign substantial probability to the lowest category of broker influence, π_0 . As noted earlier, in the case where the borrower consults with two friends ($N = 2$), the probability must exceed one half, as it does here, but only by a modest margin.

Figure 16 shows the distribution of broker influence for $N = 2, 3$, and 4. For higher values of N , the probability of zero influence is lower, because the stronger force toward getting a better deal by consulting more friends requires a stronger offset from broker influence, to

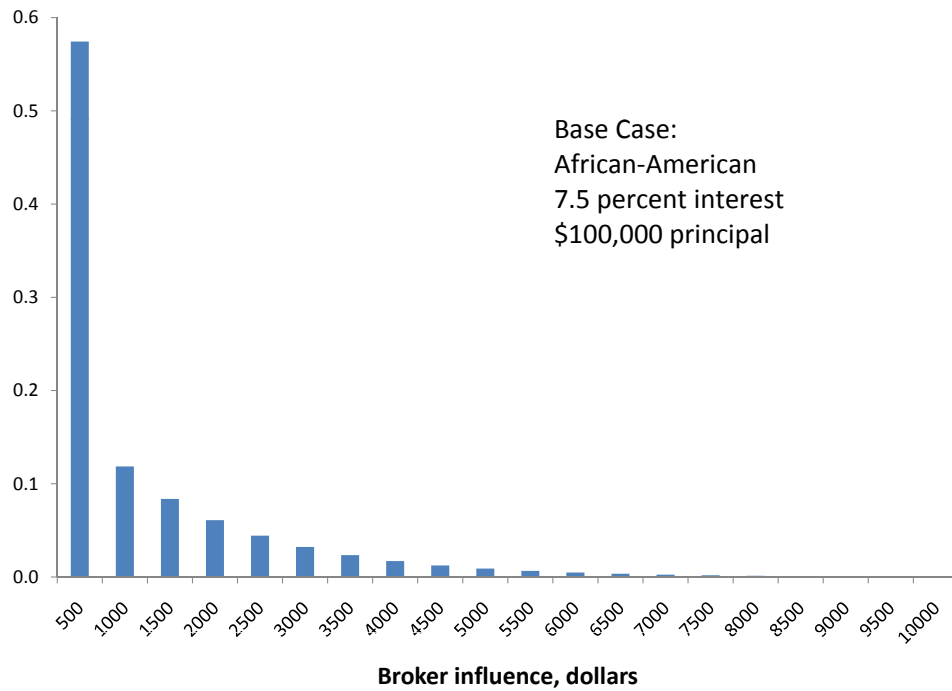


Figure 11: Estimated Distribution of Broker Influence on Closing Charge, Base Case

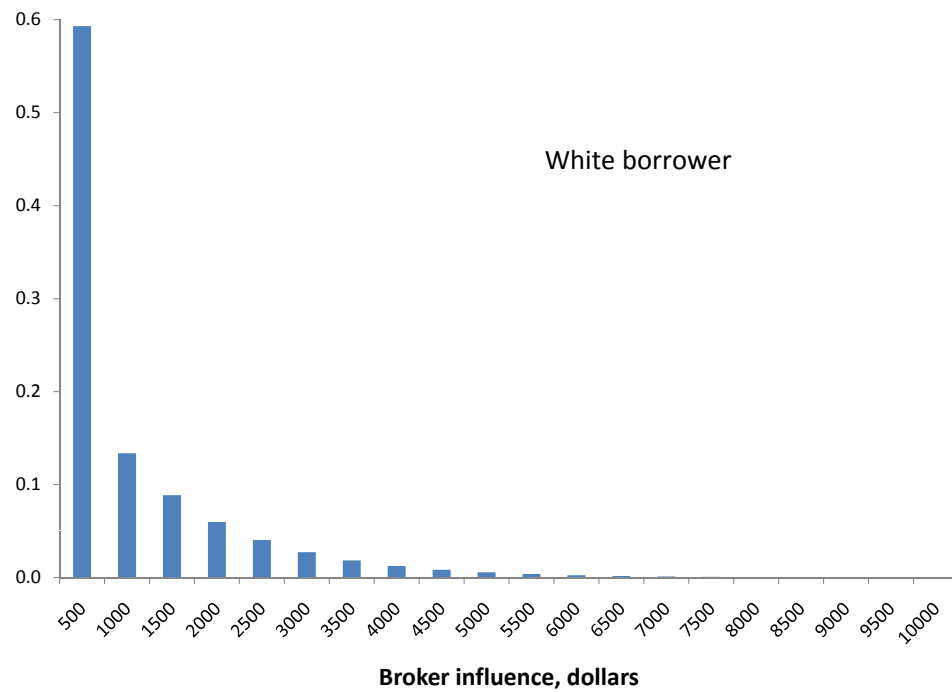


Figure 12: Estimated Distribution of Broker Influence on Closing Charge, White Borrower

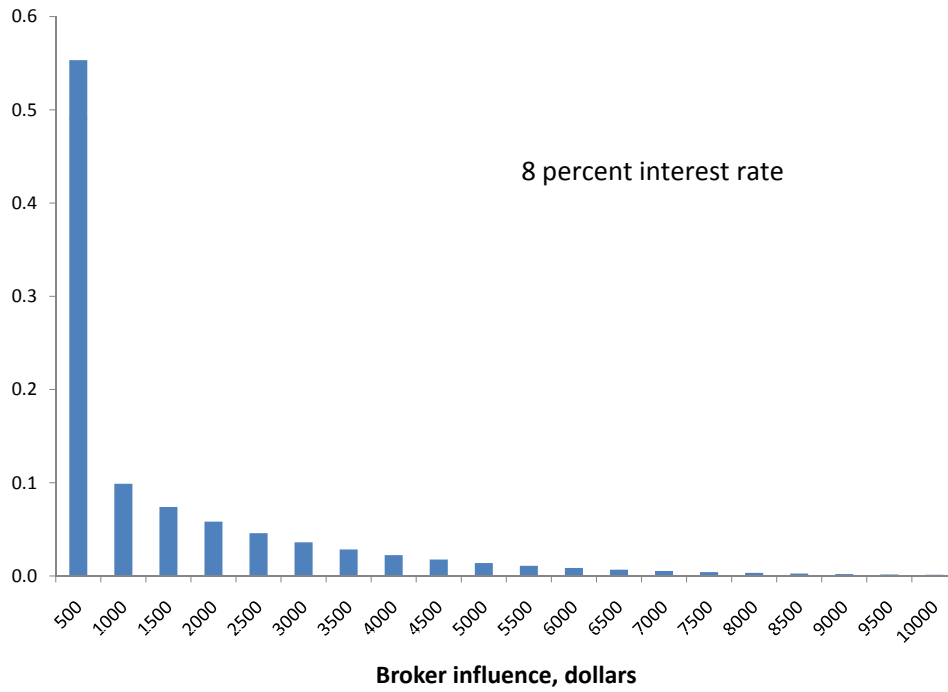


Figure 13: Estimated Distribution of Broker Influence on Closing Charge, High Interest Rate

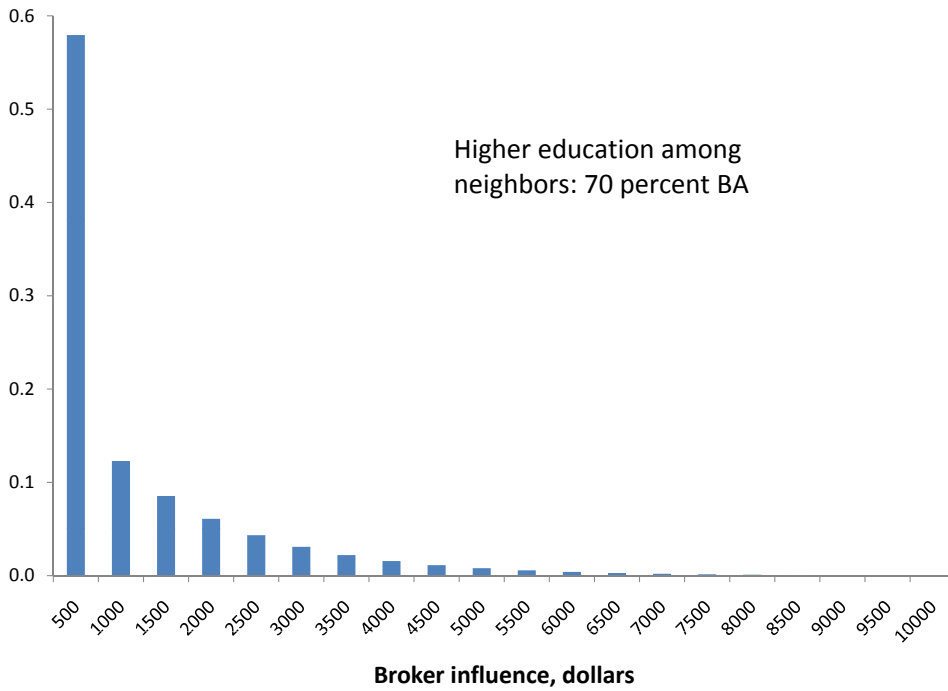


Figure 14: Estimated Distribution of Broker Influence on Closing Charge, Highly Educated Neighbors

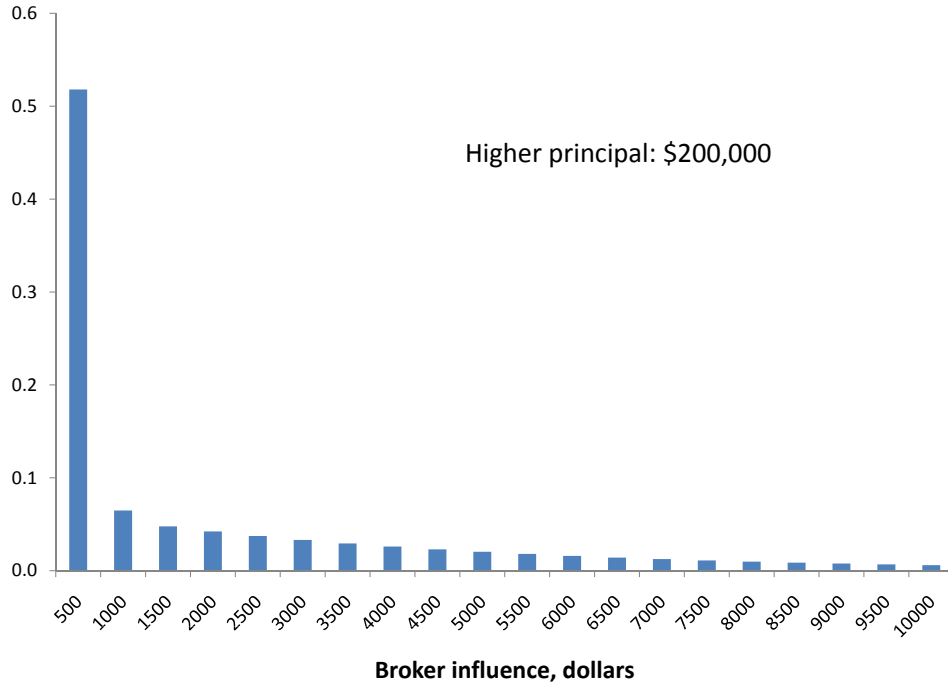


Figure 15: Estimated Distribution of Broker Influence on Closing Charge, High Principal

generate a stationary equilibrium. The rest of the distribution spreads over all the possible levels—the upper tail of very large payments to brokers above the brokers’ cost requires a small probability of high influence.

Figure 17 shows an alternative distribution of the broker’s influence, shifted somewhat toward lower influence. We think of this distribution as illustrative of the shift that would occur under a program of more effective disclosure of mortgage terms and counseling of borrowers about the role of brokers.

Figure 18 shows the distribution of payments in excess of cost, $L - \bar{L}$, corresponding to the alternative distribution in Figure 17 of the broker’s influence, for the case where borrowers consult with two friends as well as listening to the broker. Consulting substantially amplifies the effect of the lower broker influence, because the friends report lower transaction prices, under less influence upward from their brokers.

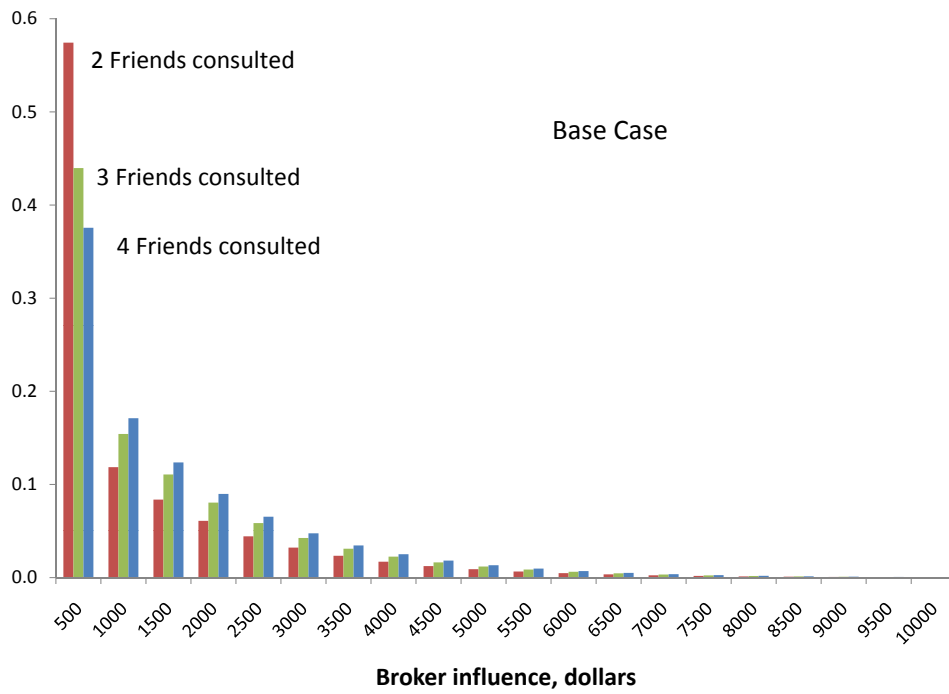


Figure 16: Estimated Distribution of Broker Influence on Closing Charge, Base Case with 2, 3, and 4 Friends Consulted

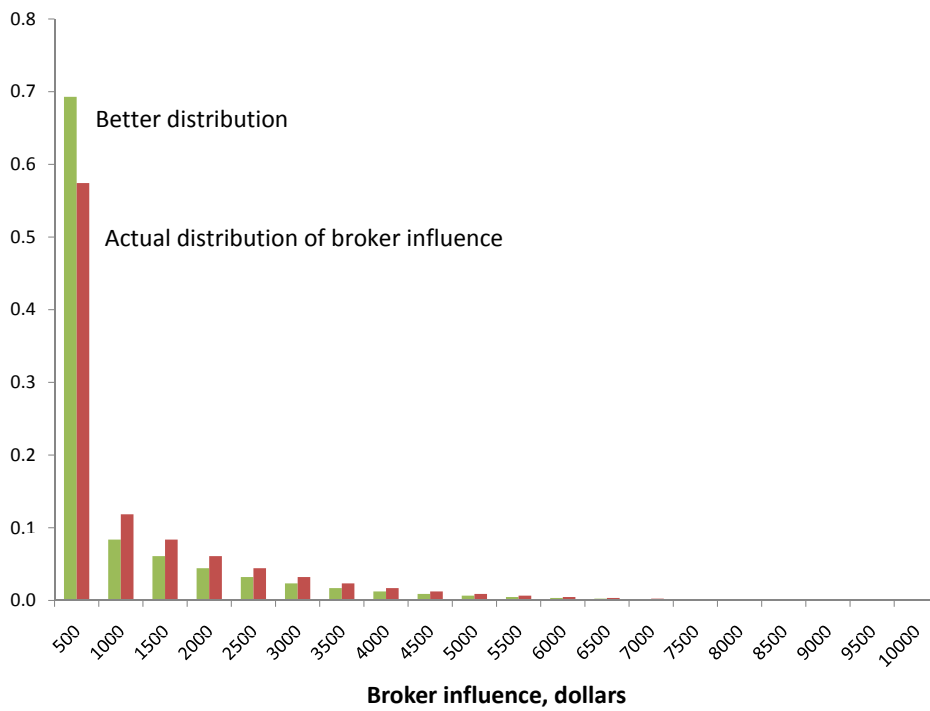


Figure 17: Original and Better Distributions of the Broker's Influence, Base Case

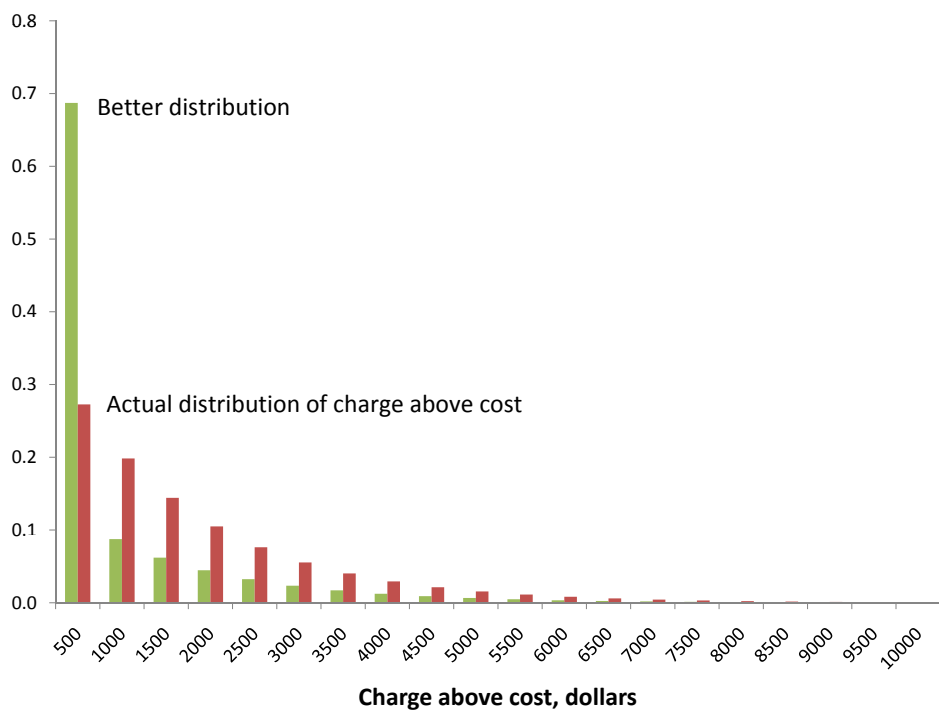


Figure 18: Alternative Distributions of the Closing Payment in Excess of Cost

Appendix

Lemma 1 *The polynomial $p(x) = x + \gamma(1-x)^N - \alpha$ with $N > 1$ and $\gamma < \alpha \leq 1$ has exactly one root in $[0, 1]$.*

Proof: Because $p(0)$ is negative and $p(1)$ is non-negative, and p is continuous, there is at least one root in $(0, 1]$. We now rule out the possibility of more than one root. Suppose $\gamma N > 1$. Then $p'(0)$ is negative and $p'(x) = 0$ has the single root in the interval $[0, 1]$,

$$\bar{x} = 1 - (\gamma N)^{-\frac{1}{N-1}}. \quad (39)$$

For $0 \leq x \leq \bar{x}$, $p(x)$ is decreasing and so cannot have a root. For $\bar{x} < x \leq 1$, $p(x)$ is increasing and can have only one root, because $p''(x) = \gamma N(N-1)x^{N-2} > 0$. If $\gamma N \leq 1$, $p(x)$ is increasing in $(0, 1]$ and by the argument above can have only one root. \square

Proposition 1 *With experts ($\chi > 0$), the equilibrium distribution is discrete; it has positive probability on the grid $\bar{L} + n\delta$ and zero density elsewhere.*

Proof: Let $x = F(L)$ for an L in $[\bar{L}, \bar{L} + \delta)$. From equation (34),

$$x = \chi + (1 - \chi)\pi_0[1 - (1 - x)^N]. \quad (40)$$

This equation satisfies the assumptions of Lemma 1 with $\gamma = (1 - \chi)\pi_0$ and $\alpha = \chi + (1 - \chi)\pi_0$, so it has the same unique solution for all L values in the interval. Thus $F(L)$ is constant over the interval. \square

Proposition 2 *The equilibrium distribution of closing payments L is unique provided there are experts: $\chi > 0$*

Proof: Consider equation (34) for the minimum value, $L = \bar{L}$:

$$F(\bar{L}) = \chi I(\bar{L}) + (1 - \chi)\pi_0[1 - (1 - F(\bar{L}))^N]. \quad (41)$$

We let $x_0 = F(\bar{L})$ and write this equation as

$$p(x_0) = x_0 + (1 - \chi)\pi_0(1 - x_0)^N - \chi - (1 - \chi)\pi_0 = 0, \quad (42)$$

which puts it into the form of Lemma 1, with $\gamma = (1 - \chi)\pi_0$ and $\alpha = \chi + (1 - \chi)\pi_0$. Note that α exceeds γ by the positive amount χ . Thus the lemma establishes that x_0 is unique. Now we let $x_i = F(\bar{L} + \delta i)$ be the values of F on the grid of values and define the polynomial for determining x_i given earlier values:

$$p(x_i) = x_i + (1 - \chi)\pi_0(1 - x_i)^N - \chi - (1 - \chi)\pi - (1 - \chi) \sum_{n < i} \pi_n [1 - (1 - x_{i-n})^N]. \quad (43)$$

This polynomial in x_i , with the lower-numbered values of x already determined, puts it into the form of Lemma 1, so x_i is uniquely determined as the single root of the polynomial. \square

Proposition 3 *With no experts, if F is an equilibrium distribution, so is another distribution in which all values of L are higher by δ : the distribution \hat{F} with $\hat{F}(\bar{L}) = 0$ and $\hat{F}(\bar{L} + n\delta) = F(\bar{L} + (n - 1)\delta)$ is also an equilibrium distribution.*

Proof: With no experts, $\chi = 0$, the equilibrium condition for $x_0 = \hat{F}(\bar{L})$ is

$$x_0 - \pi_0[1 - (1 - x_0)^N] = 0, \quad (44)$$

which has a root at zero and one positive root. Unless the second root is in $(0, 1]$, there cannot be an equilibrium at all, because the unique solution to equation (34) would be $F(L) = 0$ for all L , which is not a distribution. Thus at the lowest $x_n = F(\bar{L} + n\delta)$ that is positive, we can substitute a zero and use the positive root for the next higher point on the grid. Equation (43) will have the same solutions for the succeeding values of \hat{F} as F has one point earlier on the grid.

Proposition 4

$$\lim_{\chi \rightarrow 0} \lim_{N \rightarrow \infty} F(L) = \Pi(L), \quad (45)$$

where Π is the cumulative distribution of the broker influence, π .

Proof: The value of $x_0 = F(\bar{L})$ satisfies

$$x_0 = \chi + (1 - \chi) + (1 - \chi)\pi_0(1 - x)^N. \quad (46)$$

We need to show that x_0 is bounded uniformly away from zero as N increases. For $N = 0$, $x_0 = \chi$. Let Δx_0 be the difference between x_0 at $N - 1$ and N . We have

$$\Delta x_0 = (1 - \chi)\pi_0(1 - x)^{N-1}x_0 \geq 0. \quad (47)$$

so x_0 is increasing in N and thus $x_0 \geq \chi$ for all N . Taking the limit with N yields

$$\lim_{N \rightarrow \infty} x_0 = \chi + (1 - \chi)\pi_0. \quad (48)$$

For x_1 , a similar argument bounds it uniformly above χ and the limit is

$$\lim_{N \rightarrow \infty} x_1 = \chi + (1 - \chi)(\pi_0 + \pi_1). \quad (49)$$

The argument proceeds in the same way for the rest of the x_i . \square

Proposition 5 *If $b = 0$, $N > 1$, and $\chi > 0$, the equilibrium distribution of L places all probability at the single value \bar{L} : $F(\bar{L}) = 1$.*

Proof: x_0 be $F(\bar{L})$. Then

$$x_0 = \chi + (1 - \chi)[1 - (1 - x_0)^N], \quad (50)$$

which satisfies the assumptions of Lemma 1 with $\gamma = 1 - \chi$ and $\alpha = 1$. The single root is $x_0 = 1$. \square .

Proposition 6 *With no experts, equilibrium is impossible unless $\pi_0 \geq 1/N$.*

Proof: The polynomial for x_0 is

$$x_0 = \pi_0[1 - (1 - x_0)^N]. \quad (51)$$

Thus

$$\pi_0 = \frac{x_0}{1 - (1 - x_0)^N}. \quad (52)$$

This function is increasing in the interval $[0, 1]$ and its minimum value, at $x_0 = 0$, is $1/N$, by Bernoulli's (l'Hôpital's) Rule. When π_0 is less than $1/N$, the positive root is greater than one. Although x_0 could be chosen to be zero, the same equation would then govern x_1 , so if $\pi < 1/N$, the only solution is $x_i = 0$ for all i , but that is not a distribution, so equilibrium is impossible. \square