Price Commitments, Screening Incentives, and Privacy Protection: A Theoretical and Empirical Analysis

Jin-Hyuk Kim* Liad Wagman†

Abstract

In 1999, Congress enacted the Gramm-Leach-Bliley Act, allowing a variety of financial institutions to sell, trade, share, or give out nonpublic personal information about their customers unless their customers direct that such information not be disclosed. We study a model in which firms offer financial products to individuals, such as loans, post prices for their products, and screen consumers who apply to purchase them. Any information obtained in the screening process may be traded to another firm selling related products. We show that firms' ability to sell consumer information can lead to lower prices, higher screening intensities, higher rejection rates of applicants, and increased social welfare. By exploiting variations in the adoption of local financial-privacy ordinances in five California Bay Area counties, we are able to provide simple estimates of the effects of stricter financial-privacy laws on the denial rates of applications for home-purchase loans and loan refinancing during 2001–2004. Consistent with the model's predictions, we show that denial rates for both purchase loans and loan refinancing decreased in counties where opt-in privacy ordinances were adopted. Moreover, we find that during the financial crisis of 2007–2008, estimated foreclosure rates were higher in counties where the privacy ordinance was adopted.

Keywords: Consumer privacy; financial information; information trade; screening
JEL classification: D18, G14, G21

*Department of Economics, University of Colorado at Boulder. Email: jinhyuk.kim@colorado.edu
†Managerial Economics and Decision Sciences (MEDS), Kellogg School of Management. Email: wagman@gmail.com.

1 Introduction

Many financial institutions routinely collect nonpublic information about their customers to provide financial products or services. When consumers apply for a loan, for instance, they provide, among other details, information about their employment history and various financial statements. A lender may collect additional pieces of information from sources such as credit reports prepared by credit bureaus. Lenders and loan-servicing companies also have records of a customer's current balance, the frequency and timing of payments, and, in some cases, information about insurance policies retained. Such information can be shared with affiliates or sold to outside companies (non-affiliates), including telemarketers, where it will be used to better target consumers who may be interested in related products and services.[1]

The Gramm-Leach-Bliley Act (GLBA) is the primary federal law that specifies privacy provisions.[2] GLBA requires financial institutions to notify customers about how their personal information is collected and used. In particular, financial institutions that share or sell customer data to non-affiliated third parties must give customers a chance to opt out; that is, request that their information not be shared (15 USC §§ 6801-6809). However, there are several exemptions under the GLBA that can permit information sharing despite a consumer's objections.[3] Since the enactment of the GLBA in 1999, there has been much

[1] Banks can sell a variety of consumer information with little scrutiny of the buying party. In 1997, Charter Pacific Bank sold more than three million credit-card numbers to a convicted felon, who used these numbers to run a fraud scheme. The law did not require banks to check for criminal records of its buyers' pasts. See `http://articles.latimes.com/1999/sep/11/news/mn-8867.`

[2] GLBA partially repealed the Glass-Steagall Act of 1933 by allowing banking, insurance, and securities companies to operate under the same entity. Financial holding companies so created can have a variety of non-banking affiliates. Under the GLBA consumers have no right to stop sharing of nonpublic personal information among affiliates.

[3] For instance, a financial institution can share information with an outside marketer in order to jointly offer products. In fact, many of the nation's leading banks use information about their customers' shopping habits to help retailers target offers to customers without actually releasing their data. See `http://money.cnn.com/2011/07/06/pf/banks_sell_shopping_data/index.htm`.

debate about whether GLBA privacy provisions meet increasing public concern surrounding consumer privacy.[4]

In this paper, our first objective is to explore how firms' ability to acquire and sell consumer data influences social welfare. Existing theoretical works on privacy focus primarily on firms' incentives to price discriminate once they acquire consumer information (see, for instance, Taylor 2004; Acquisti and Varian 2005; Calzolari and Pavan 2006; Conitzer et al. 2012). In contrast, the analysis in this paper focuses on firms' incentives to screen applicants when prices are posted upfront (e.g., Burke et al. 2012). In particular, we study a competitive industry where, prior to approving a purchase, sellers may work to acquire information about applicants. Such information is pertinent to sellers because it may affect their cost of servicing an applicant, and it may also be of interest to sellers with overlapping target markets.

We investigate two settings—a confidential regime, in which firms cannot sell consumer information; and a disclosure regime, in which information can be traded. Our paper contributes to the earlier literature by showing that firms' ability to trade in consumer data can lead to lower prices and increased welfare when information acquisition is explicitly taken into account. Importantly, the value of information in prior works is derived from a seller's ability to price discriminate. In contrast, in our model, the value of information comes from reducing the cost of servicing consumers. When the market is competitive, these cost savings are passed on to consumers who overall benefit from significant price cuts. We show that firms' screening standards are more stringent under the disclosure regime, whereby more ap-

[4]The 107th Congress introduced a number of bills that seek to modify the GLBA to require consumer opt-in consent for information transfers. See, for instance, the Financial Institutions Privacy Protection Act of 2001, S. 450, § 3 (2001); Consumer's Right to Financial Privacy Act, H.R. 2720, § 2 (2001).

plicants are rejected. This is because the downstream firms' willingness to pay for consumer information rises as the accuracy of consumer information increases.

Our second objective is to empirically test our model's predictions, that is, whether trade in consumer information can indeed lead to higher screening intensities and therefore to higher rejection rates. One of the main criticisms of GLBA's privacy provisions has been that most consumers do not (and likely will not) take advantage of the *opt-out* option to cease trade in their information.[5] In 2002, three out of five counties in the San Francisco-Oakland-Fremont (SFOF), CA, Metropolitan Statistical Area adopted a local ordinance (effective January 1, 2003) that is more protective than then-current practices by pursuing an *opt-in* approach. Specifically, the local ordinance would require financial institutions to seek a written waiver before sharing consumer information with both affiliates and non-affiliates.

Our empirical results are robust and new to the literature. We analyze loan applications for conventional home purchases as well as loan refinancing in about 800 census tracts in the SFOF Metropolitan Statistical Area (MSA) during 2001–2004 and provide simple difference-in-difference estimates of the effects of the local privacy ordinance. Since unobserved heterogeneities are likely to be less prominent within a single MSA, comparisons of loan denial rates before and after the ordinance offer a unique opportunity for evaluating the effects of strengthening consumer-privacy protection. We find that the opt-in ordinance had a significantly negative effect on loan denial rates (that is, approval rates increased), consistent with our model's predictions. Furthermore, we provide reduced-form estimates of the effects of the privacy ordinance on foreclosure rates during the financial crisis of 2007–2008,

[5]Critics argue that this may be the case because disclosures are often buried in fine print with confusing terminology and because (since consumers need to opt out) the default option is for information trade to take place (see, for instance, Johnson and Goldstein, 2004).

which in fact suggests that foreclosure rates were higher in the counties that adopted the privacy ordinance.

The paper proceeds as follows. Section 2 provides a brief discussion of related literature. Section 3 presents our basic model of consumer screening and information trade. Section 4 derives the positive and normative implications of consumer privacy protection in our model. Section 5 provides some background information on state and local privacy ordinances and describes the data set. Section 6 presents empirical findings, demonstrating the extent to which our model's predictions account for the changes in loan denial rates and foreclosure rates. Section 7 contains concluding remarks.

2 Related Literature

This paper is related to the growing literature on consumer privacy. Taylor (2004) and Villas-Boas (2004) show that in the presence of strategic consumers, a firm may be worse off by targeting prices based on consumers' purchase histories.[6] The reason is that once consumers anticipate future prices, they may choose to forgo purchases today to avoid being identified as past customers tomorrow, and thus have access lower prices targeted at new customers. This strategic waiting by consumers can lower a seller's profit by reducing sales and diminishing the benefit of price discrimination. Our analysis differs from these studies in that our focus is on a competitive market, and because we place the strategic burden of whether or not to move forward with a transaction on firms.

[6]Other theoretical papers that investigate price discrimination based on consumers' purchase histories include Acquisti and Varian (2005), Armstrong (2006), Hermalin and Katz (2006), Chen and Zhang (2009), and Conitzer et al. (2012). See, e.g., Fudenberg and Villas-Boas (2006) and Hui and Png (2006) for surveys of this literature.

The literature has considered the relationship between information revelation to a potential trading partner and the efficiency of outcomes. In a sequential agency model, Calzolari and Pavan (2006) show that committing ex ante to disclose information can sometimes increase social welfare.[7] In their model, a principal can learn an agent's private information through a screening mechanism, and the benefit of disclosure arises from profitably selling information. Although our efficiency results are similar to their findings, the difference is that in their work information acquisition is modeled as a passive process that produces soft information, whereas in this paper screening intensity is endogenously determined and produces hard information.

Recently, researchers have begun to evaluate empirically the effects of privacy regulations. Using variations in state medical privacy laws, Miller and Tucker (2009, 2011) show that privacy regulations restricting a hospital's release of patient information significantly reduced the adoption of electronic medical records, while a 10 percent increase in the adoption of such systems can reduce infant mortality by 16 deaths per 100,000 births. Goldfarb and Tucker (2011) examine the effects of the implementation of the EU Privacy Directive and find some evidence that after the Privacy Directive was passed, advertising effectiveness decreased significantly. We provide new evidence on the effects of privacy regulation with a focus on financial markets.

This paper is also related to the literature on information sharing in credit markets. Pagano and Jappelli (1993, 2002) predict that if banks share information about their customers, they would increase lending to safe borrowers, thereby decreasing default rates.

[7]Their result is reminiscent of the traditional Chicago School argument that privacy protections hinder information flows that would otherwise lead to improved efficiency in the market because individuals may then misrepresent their personal information (see, e.g., Stigler 1980; Posner 1981).

Existing empirical studies mostly focus on the effects of credit bureaus and creditor rights using data from a cross-section of countries (see, e.g., Djankov et al. 2007; Qian and Strahan 2007). However, in a recent paper Doblas-Madrid and Minetti (2013) analyze contract-level data from the equipment finance industry to show the effect of lenders' information sharing in this capital market. Specifically, they show that the entry of lenders into a credit bureau can reduce the incidence of contract delinquencies and defaults.

A number of works have studied the U.S. mortgage default crisis and pointed out the expansion in mortgage credit to subprime borrowers (e.g., Mian and Sufi 2009). This literature, however, has not explicitly considered the effects of privacy protection or privacy legislation. In this paper, we incorporate information acquisition and information trade into a model of consumer screening, and test the model's main predictions using census tract-level data on the disposition of mortgage applications. The results in this paper give rise to the conjecture that part of the role of the GLBA in the subprime mortgage crisis was to increase foreclosure rates by weakening lenders' incentives to screen mortgage applications.

3 Consumer Screening and Information Trade

3.1 The Model

The supply side of the market is composed of two types of symmetric firms, A and B. Firms of type A offer financial products such as conventional home loans (good A), and firms of type B offer related products, such as personal credit lines and insurance policies (good B). All firms are risk neutral and maximize expected profits. The demand side of the market consists of a continuum of ex-ante identical individuals with measure $M > 1$. A fraction of this mass, normalized to 1, is applying to purchase a loan from firm A. Individuals are assumed to be risk neutral and have unit demands (separately) for goods A and B. Let v_A and v_B denote the incremental utilities from consuming one unit of good A and B, respectively.

Suppose that there is uncertainty about cost-relevant consumer characteristics. Specifically, the cost of serving a consumer either turns out to be low (c_L^m) or high (c_H^m), with $c_L^m < c_H^m$ for $m \in \{A, B\}$. For technical simplicity, we assume that consumer types for firm B's products are perfectly correlated with those of firm A's.[8] We assume that $c_L^m < v_m < c_H^m$ for $m \in \{A, B\}$; that is, for both goods, it is efficient to serve only low-cost consumers. However, if screening is imperfect, firms may end up approving some high-cost consumers. It is common knowledge that the proportion of high-cost consumers in the population is $\lambda > 0$. At the onset, information is incomplete and symmetric, and, in particular, consumers (and firms) do not observe the realization of their types.[9]

The game unfolds in several stages. First, each firm j of type A announces a price $p_{j,A} \in \mathbb{R}_+$ at which it will sell the good to a consumer whose application is ultimately approved. Price announcements are made publicly and simultaneously. Next, each consumer applies to purchase the good from a firm A of his choice. Then, each firm of type A acquires information about its applicants and chooses which applicants to qualify. After selling to qualified applicants, each firm of type A makes a take-it-or-leave-it offer to a firm B for purchasing its list of applicants and the information it acquired about them, including whether

[8]Calzolari and Pavan (2006) make an analogous assumption. Taylor (2004) considers imperfect correlation. Incorporating imperfect correlation would not change our qualitative results but it would complicate the exposition without a major gain in intuition.

[9]In practice, consumers belong to different credit categories (e.g., prime, near-prime, etc.), and our analysis under informational symmetry more naturally applies separately for each credit category. When consumers' cost types are privately known by consumers at the beginning of the game, our main results continue to hold as part of a pooling equilibrium, provided that $v_A - v_B$ is sufficiently high, whereby a type c_H consumer would not forego applying to purchase good A in fear of not being offered good B.

they were approved.[10] Finally, firm B decides whether to accept or reject this offer and proceeds to make targeted sale offers to potential customers.

3.2 Information Acquisition

A firm A chooses a sample size, or search intensity $n \geq 0$, which we treat for simplicity as a continuous variable. The cost to the firm of acquiring information about an applicant is kn, where $k > 0$. By choosing a search intensity n, firm A receives n conditionally independent Bernoulli signals, $\{X_1, \ldots, X_n\}$, where

$$\Pr\{X_i = 1 | c^A\} = \begin{cases} 1, & \text{if } c^A = c_L^A, \\ 1 - \alpha, & \text{if } c^A = c_H^A \end{cases}$$

The parameter $\alpha \in (0, 1)$ represents intrinsic signal strength. If $\alpha = 1$, then a signal is fully informative, and if $\alpha = 0$, then signals contain no information. This process is interpreted as follows: A firm A chooses a search report containing $i = 1, \ldots, n$ records, $\{X_1, \ldots, X_n\}$, for each of its applicants, and each record is either positive ($X_i = 1$) or negative ($X_i = 0$).

Since the firm is in effect searching for bad news about its applicants' creditworthiness, it is possible to summarize all the information contained in an applicant's search report with the sufficient statistic $S_n \equiv \min\{X_1, \ldots, X_n\}$. That is, if $S_n = 0$, then at least one of the records was negative, and the applicant (referred to as disqualified) is certainly type c_H^A; whereas if $S_n = 1$, then all records were positive, and the applicant (referred to as qualified)

[10] Selling to a single firm of type B is not restrictive. As we will show below, the value of information is derived from updating beliefs about a consumer's cost type, so that a firm A can indeed sell to multiple firms of type B. An alternative interpretation is that the aggregate measure of B firms is normalized to one.

is type c_H^A with probability

$$\mu(n) = \frac{\lambda(1-\alpha)^n}{\lambda(1-\alpha)^n + (1-\lambda)} < \lambda \tag{1}$$

and type c_L^A with the complementary probability $1 - \mu(n)$.

After acquiring information about a consumer, a firm j of type A decides whether to approve the consumer's application (i.e., sell him the good at its posted price $p_{j,A}$). Approval results in an expected payoff of $p_{j,A} - E[c^A|S_n] - kn$ for the firm. Rejection results in a payoff of zero for the consumer and $-kn$ for the firm. We define a measure of the efficacy of firm A's information-acquisition technology as follows:

$$m \equiv -\frac{k}{\ln(1-\alpha)}.$$

Lower values of m correspond to better technologies involving low sampling costs and/or a high intrinsic signal strength.

3.3 The Value of Information

To focus on the effects of information trade, we assume that firms of type B do not possess the technology to independently acquire information about potential customers. This is consistent with observations that credit card and insurance policies are often approved without pursuing high levels of direct screening; moreover, issuers often make prescreened offers based on summary information (e.g., credit scores) that they purchase from primary sources. Let us suppose that a firm of type B (henceforth, firm B) randomly targets a fraction of the total mass of potential consumers, M. Let ξ denote the probability that a customer in firm B's *initial* target set overlaps with a firm of type A's approved set of applicants.

The benefit to firm B of learning about a *qualified* consumer from A's list is the following: With probability ξ, this consumer is already targeted by firm B, whereby B's benefit is zero. With probability $1-\xi$, this consumer would not have been targeted by B, in which case B's benefit is $E[c^B] - E[c^B|S_n = 1]$ in expected cost savings (where $E[c^B] = \lambda c_H^B + (1-\lambda)c_L^B$ and $E[c^B|S_n = 1] = \mu(n)c_H^B + (1-\mu(n))c_L^B$). Thus, firm B's overall expected benefit from learning about a qualified consumer is $(1-\xi)(E[c^B] - E[c^B|S_n = 1])$. Analogously, firm B's expected benefit from learning about a randomly-selected *disqualified* consumer is given by $\xi(c_H^B - E[c^B])$. Formally, we can derive B's willingness to pay per application contained in firm A's applicant list (given a level of screening intensity n) as follows:[11]

$$\underbrace{[(1-\lambda)+\lambda(1-\alpha)^n](1-\xi)(E[c^B]-E[c^B|S_n=1])}_{\text{Qualified consumer information}} + \underbrace{\left(1-[(1-\lambda)+\lambda(1-\alpha)^n]\right)\xi(c_H^B - E[c^B])}_{\text{Disqualified consumer information}}.$$

Substituting for $\mu(n)$ from (1) and simplifying yields

$$(1-\lambda)\lambda(1-(1-\alpha)^n)(c_H^B - c_L^B). \tag{2}$$

From the expression in (2), it is clear that firm B's expected benefit from information about firm A's applicant list is independent of ξ and is increasing in n. Intuitively, firm B benefits from firm A's information by being able to fine-tune its target customer set. Specif-

[11] We note here that firm B could directly swap out a disqualified consumer and swap in a qualified consumer, which would result in the sum of the two benefits. This would generate the same expected benefit as when firm B swaps out a disqualified consumer with a new consumer (for whom it has no information), and swaps out a consumer for whom it has no information with a qualified consumer from outside its simple.

ically, a higher screening intensity enables B to more effectively avoid high-cost consumers and to better target low-cost consumers. Since the overall benefit is a combination of improving the targeting of individual consumers both in and out of B's initial target set, it is independent of the likelihood that A's information directly overlaps with this initial set.

4 Equilibrium in a Competitive Market

4.1 Main Results

In the proceeding analysis, we assume that the information-acquisition technology for firms of type A is sufficiently effective (i.e., m is sufficiently small) so that an interior solution obtains with positive screening intensity (i.e., $n^* > 0$).[12] The super- and sub-scripts T and NT denote cases with and without information trade, respectively. For notational simplicity, we omit the subscript j when referring to a representative firm of type A. Lemma 1 derives preliminary results that hold in equilibrium both when information trade is permitted and when it is not (all proofs are provided in the Appendix).

Lemma 1 *In a symmetric equilibrium, a consumer's expected utility is decreasing in firm A's price p_A.*

Lemma 1 states that, taking into account a firm A's subsequent choice of screening intensity, an applicant's expected utility decreases monotonically in firm A's posted price. It follows that consumers choose to apply to a firm A that posts the lowest price. As the

[12] A condition sufficient to guarantee this is that $m < \lambda(c_H^A - v_A)$. This follows given that a firm A's expected profit without information trade is $(\lambda(1-\alpha)^n + (1-\lambda))p_A - (1-\lambda)c_L^A - \lambda(1-\alpha)^n c_H^A - kn$ and that its derivative evaluated at $n = 0$ has to be positive. The condition for an interior solution with information trade is weaker. See the proof of Lemma 1.

next result indicates, this choice has an important implication for firms' levels of consumer screening. In particular, the firms that post the lowest prices are also those that will be screening consumers most intensely.

Proposition 1 *Relative to when consumer information cannot be traded, with information trade (i) consumers purchase from a firm A at lower prices, (ii) more information is collected about consumers, and (iii) more consumers are disqualified from purchasing good A.*

Proposition 1 points to an inverse relationship between price and screening intensity. From the standpoint of firms, price competition dissipates profits from selling good A. However, once information trade is permitted, due to profits from selling applicants' information, firms A are able to lower their prices even further. This price reduction is coupled with a stricter screening of applicants. Thus, on the one hand, consumers benefit from lower posted prices; on the other hand, more consumers are disqualified compared to the case where information trade is prohibited. Our next result addresses allocative efficiency.

Proposition 2 *There exists a constant $\bar{c} > 0$, such that for $c_H^B - c_L^B \geq \bar{c}$, allowing for information trade strictly increases ex-ante social welfare.*

That is, when firm B's benefit from applicant information for product A is significant, the price reductions consumers enjoy ex ante for good A offset consumers' disutility from higher rates of ex-post rejections. If $c_H^B - c_H^L$ is sufficiently high, qualified consumers can even be paid to purchase good A; that is, the price for good A when information trade is permitted, $p^*_{A,T}$, could be negative. Intuitively, price commitments, whereby screening is conducted after consumers apply, lead firms selling good A to compete away their downstream profits from selling their applicants' information. In turn, more consumer data is collected and traded, more consumers are rejected from buying good A, but those who are ultimately approved enjoy significant discounts.

While our model purposely focuses on the effects of information trade in the presence of price commitments, our results provide complementary views to those in the literature on price discrimination. In particular, the literature on price discrimination (e.g., Taylor 2004; Villas-Boas 2004; Acquisti and Varian 2005) suggests that when firm B is able to price discriminate based on information that correlates with, for instance, high valuation for its products, consumers may choose to strategically avoid buying good A in an effort not to experience a price hike for good B. Here, we have shown that any such adverse effects due to price discrimination may be offset, and possibly entirely dominated, by consumers' benefit from significant price cuts due to information trade when the market is competitive.

4.2 Disutility from Default Risks

Thus far, we have assumed that consumers differ only in terms of sellers' costs of serving their respective types, where all consumers derive the same marginal utilities from acquiring sellers' products. However, consumers' tastes may differ whereby, for instance, each consumer type derives a different value v_A from purchasing good A. Specifically, net utility from a loan may reflect a consumer's likelihood of default, where high-cost consumers are more likely to default and experience disutility from a potential loss of creditworthiness and foreclosure. That is, letting $v_{A,H}$ and $v_{A,L}$ denote a high and low type's incremental utility, respectively, it may be the case that $v_A = v_{A,H} = v_{A,L}$ no longer holds. Our analysis readily extends to the case in which high-cost consumers experience a higher level of expected disutility from default (and thus place a lower valuation on good A) than low-cost consumers, that is, $v_{A,H} < v_{A,L}$.[13]

Proposition 3 *Suppose high-cost consumers derive a lower utility from good A than low-cost consumers; then allowing for information trade leads to higher ex-ante social welfare.*

Proposition 3 extends our findings in the baseline case by showing that information trade brings the outcome closer to the social planner's solution when taking into account high-cost consumers' higher likelihood of (and hence disutility from) default. In such cases, allowing for information trade is even more important in terms of improving welfare. This is because the social cost of misallocating good A (that is, of mistakenly qualifying high-cost consumers) is now strictly higher. Since firms screen applicants more intensely when information trade is permitted, allocative efficiency improves due to a greater likelihood of avoiding more costly defaults. In other words, when factoring in consumers' disutilities from default, there are greater benefits associated with information trade.

5 Data and Background

5.1 Privacy Legislation

Our theory predicts different levels of screening (and hence approval or rejection rates) depending on whether consumer information can be traded. As previously mentioned, the

[13]Since individuals behave rationally, they anticipate that buying a loan as a high-cost type can result in greater disutility due to a higher likelihood of default. We note that it is possible that high-cost consumers tend to discount future risks of default on a loan even when they may care greatly about such consequences in the future. Although the model can be extended to account for these considerations, the analysis below does not seek to address these behavioral issues.

GLBA's privacy provisions place the burden on individuals to protect their privacy with an opt-out standard. That is, financial institutions can share customers' nonpublic personal information if customers neglect to respond.[14] Since the enactment of the GLBA, there have been considerable legislative activities in state governments in regards to privacy issues, financial in particular, some pertaining to the adoption of an opt-in standard.[15] As a result, there exists significant state-level variation in the protection and trade of consumers' financial information by financial institutions.

However, certain features of state privacy legislations render themselves less suitable for an empirical analysis exploiting state-level policy variation. First, some states, including Alaska, Connecticut, Illinois, North Dakota, and Vermont, have strict laws that currently require an opt-in consent for the sharing of consumer information with non-affiliated third parties while most other states do not. The problem is that these states adopted an opt-in approach in their state banking laws long before the GLBA was enacted in 1999. Therefore, given that our data source (which follows below) begins in 1999, pre-treatment outcomes would not be observed if we were to compare outcomes in these states against others using a difference-in-difference estimation.[16]

[14] To be clear, consumers have no opt-out rights whatsoever against affiliate sharing, and financial institutions can further evade opt-out requirements by exploiting the exceptions in the GLBA (e.g., joint marketing exception).

[15] The GLBA permits states to formulate privacy protections that exceed federal law. The Federal Trade Commission (FTC) was granted sole authority to make determinations as to whether state statutes are inconsistent with (and therefore pre-empted by) the GLBA. In 2001 and 2004, the FTC issued a formal letter in which it determined that affirmative opt-in provisions in Illinois and North Dakota, respectively, were consistent with GLBA.

[16] An interesting case in point is North Dakota. Before GLBA was enacted, North Dakota had a strict opt-in privacy law that required consumer consent to share financial information with unaffiliated third parties. Effective July 1, 2001, the North Dakota State Government passed an emergency legislation that authorized disclosure of customer information by a financial institution to a non-affiliated third party if the disclosure was in compliance with the GLBA. This would have given a unique opportunity to examine the effect of consumer privacy protection; however, in the first state-wide referendum on financial privacy held in June 11, 2002, North Dakota voters repealed the 2001 amendment effective immediately. Given the annual

Second, an increasing number of states have in fact enacted laws that limit the sale of personal information by financial institutions and impose stricter requirements for many third-party uses. Perhaps the best-known example is the California Financial Information Privacy Act (CalFIPA), which in part superseded the opt-out approach of the GLBA.[17] Here, the main issue is that there is substantial heterogeneity in the language (e.g., definitions, scope, and strength) of state privacy laws. For instance, some state privacy laws apply only to banks and credit unions, and others pertain only to account services and electronic-funds transfers. This makes us rather skeptical about the use of state variation due to issues arising from the common trend assumption, which is crucial for difference-in-difference methods.

Thus, we turn our attention toward local-government legislation, where heterogeneity across different areas is more plausibly controlled by the inclusion of observed characteristics. Before the CalFIPA was signed into law in September 2003, the bill had been rejected multiple times by the Assembly. As these bills were being defeated, local governments in the Bay Area led efforts to strengthen consumer financial-privacy protection. In August 2002, they began enacting ordinances that would require financial institutions to obtain consumers' permission before releasing their financial information to a nonaffiliated third party or affiliate.[18] To the authors' knowledge, this event offers a unique opportunity —

frequency of the data we use, the short time period between these two events does not allow for an effective identification.

[17]The opt-out provision of this law regarding affiliate sharing can be preempted by the Fair Credit Reporting Act (see American Bankers Association v. Gould, 412 F.3d 1081). However, in 2008 the Ninth Circuit upheld substantial portions of California's financial privacy law regarding affiliate sharing (see American Bankers Association v. Lockyer, No. 05-17163), and the Supreme Court denied review of the California law in 2009.

[18]Similar to the case of CalFIPA, in response to a lawsuit filed by Bank of America and Wells Fargo, a federal judge struck down the local ordinance with regards to affiliate sharing on July 29, 2003, but the restriction on non-affiliated third-party sharing (unless a consumer gives advance consent) was upheld. See `http://www.abag.ca.gov/privacy`.

we were unable to find any other case where local jurisdictions successfully implemented a consumer financial-privacy ordinance in such a fashion.

According to the Association of Bay Area Governments, the five counties in the San Francisco-Oakland-Fremont Metropolitan Statistical Area (MSA) considered adoption of such opt-in privacy ordinances, and the result was that three of them (Alameda, Contra Costa, and San Mateo) adopted the ordinance, while the other two (Marin and San Francisco) did not (see Figure 1 for a map of this MSA). We note here that, as with many other studies using policy variation, it is difficult to completely rule out the bias that arises from the issue that the policy adoption may be an endogenous outcome. However, we argue that it is reasonable to believe that the adoption would be related to industry lobbying or job-related issues rather than systematically associated with our outcome measures. Further, we control for a relatively rich set of census tract-level characteristics, which mitigates the concern of selection.

5.2 Data Description

The Bay Area local ordinance was adopted in October, 2002, with an effective date of January 1, 2003. However, the CalFIPA, which established state-wide opt-in requirements, went into effect on July 1, 2004, eliminating the policy variation created by the (non)adoption of the local privacy ordinance across counties. Hence, we limit our post-intervention period to 2003 and 2004.[19] We chose 2001 and 2002 for our pre-intervention period because the GLBA

[19]In fact, both the local ordinance and the CalFIPA were in part pre-empted by the affiliate-marketing provisions in Section 214 of the Fair and Accurate Credit Transactions Act of 2003, which amends the Fair Credit Reporting Act. However, the Federal Trade Commission established December 1, 2004, as the effective date for Section 214, so this has no effect during our post-intervention period. See `http://www.gpo.gov/fdsys/pkg/FR-2004-02-11/html/04-2913.htm`.

came into effect on November 13, 2000, which facilitated the sharing of consumer information among affiliates by establishing an opt-out standard and allowing different types of companies to affiliate with each other via a financial-holding company.

As we argue above, geographic proximity and dense population within a single MSA might confer a closer comparison group than areas in different states. In particular, the Association of Bay Area Governments encouraged consistency and uniformity of the local privacy ordinances, which led to a model ordinance.[20] Another advantage is that the data we use comes from government-mandated disclosures that are reported by all financial institutions, including mortgage lenders. Furthermore, the composition of the population would be relatively stable over a four-year period. Although we cannot rule out spurious correlation due to demographic changes, we believe its quantitative effect would be small.

The Home Mortgage Disclosure Act requires financial institutions (including banks, savings associations, credit unions, and other mortgage lending institutions) to annually report disclosures of their lending activities. Using the data submitted by these financial institutions, the Federal Financial Institutions Examination Council releases aggregate lending information on the disposition of mortgage applications at the census tract level.[21] Census tracts are small, relatively permanent statistical subdivisions of a county, and usually consist of 2,500 to 8,000 persons. When first delineated by the government, they are designed to be homogeneous with respect to population characteristics, economic status, and living conditions.

For each year, we observe the disposition of loan applications by property location and

[20] See `http://www.abag.ca.gov/privacy/ordinances.html`.

[21] The aggregate tables are publicly available at `http://www.ffiec.gov/hmda`.

the type of loan. That is, for each census tract, the data shows how many loans were originated, how many were approved but not accepted, and how many applications were denied, withdrawn, or closed for incompleteness. The data also shows the aggregate dollar amounts in each of these five categories. This information is compiled for each of three types of loans: home-purchase loans, loan refinancing, and home-improvement loans. We do not use home-improvement loans in our analysis because, unlike the first two types of loans, a home-improvement loan may not require a lien on a dwelling, and the number of applications is relatively small.

Table 1 shows the means of the variables in our data set. We focus on conventional home-purchase loans and refinancing loans for 1-to-4 family dwellings. In the former case, we exclude government-insured mortgages such as FHA, FSA/RHS, and VA loans because the approval standard for these loans may well be different than that for conventional loans. In constructing the application denial rates, we divide the number (or dollar amount) of applications that were denied by the sum of the number (or dollar amount) of i) loans originated, ii) loans approved but not accepted, and iii) applications denied.[22] We will refer to the denominator as the (total) number of loan applications in the next section. For both home-purchase loans and refinancing loans, mortgage denial rates increased in 2003–2004 compared to the 2001–2002 levels. The increase in denial rates in the treatment group (i.e., Alameda, Contra Costa, and San Mateo counties) appears a little smaller than that in the control group (i.e., Marin, and San Francisco counties). A simple comparison of means is

[22]We do not use information on withdrawn or incomplete applications in our calculation because we do not observe the reason for such actions. That is, it may be the lender or the applicant who decided to withdraw an application. More importantly, withdrawn applicants are likely to apply for a loan elsewhere, and hence by not including them we may avoid double-counting of such applications.

suggestive, but we need to allow for other sources of variation in loan denial rates.

For control variables, we add tract-level economic characteristics (tract median income as % of MSA median; % of population below poverty threshold; an indicator for being inside a central city), population characteristics (% of minority population; % of Asian population; % of Black population; % of Hispanic population), and housing characteristics (median age of housing stock; % of owner-occupied housing units; number-of-households-to-housing-units ratio).[23] Notice that all control variables, except the central city dummy, have comparable means. This confirms the view that the adoption of the local ordinances was mainly influenced by the presence of financial institutions in the central city. However, this dummy variable does not seem to play an important role in our results. On the other hand, in each county there is a large amount of variation in economic, population, and housing characteristics, and they are comparable across counties. Hence, it seems less likely that the composition effect will drive our results. In the next section, we will also use tract-level foreclosure rates provided by the Department of Housing and Urban Development.

6 Empirical Findings

The entries in Tables 2–5 are regression coefficients from the following difference-in-difference specification:

$$Denial\ Rate_{it} = \beta_0 + \beta_1 Treat_{it} + \beta_2 Post_{it} + \beta_3 Treat_{it} \cdot Post_{it} + \gamma X_{it} + \varepsilon_{it}.$$

[23]Minority population is the tract's total population minus its white population. Poverty status is determined by assigning each person or family into various poverty thresholds, which vary by family size and the ages of the family members. See `https://www.census.gov/hhes/www/poverty/data/threshld/index.html` for details.

where for each tract i and year t, $Treat_{it}$ is an indicator variable for those belonging to the treatment group; $Post_{it}$ is an indicator for post-intervention periods; and X_{it} is a set of control variables discussed above. In addition, we employ a full set of county fixed effects to absorb cross-sectional differences, and year fixed effects for nonlinear trends in mortgage denial rates not directly associated with privacy legislation.

In Tables 2–5, columns (i)–(iii) show the ordinary least square (OLS) estimates. In all of the tables, robust standard errors are reported, clustered at the county level, to allow for arbitrary within-county correlation in errors. Because all our dependent variables are in ratio form, the constant variance assumption may bias the OLS results. That is, a mortgage denial rate calculated based on a smaller number of applications may be less reliable than that calculated based on a larger number of applications. To address this concern, we provide alternative estimation results using weighted least squares (WLS) in columns (iv)–(vi), where the weight is the denominator (i.e., number of total applications) used to calculate denial rates.

Specifically, Table 2 reports the estimation results when the outcome variable is mortgage denial rates (in number) for conventional home-purchase loans. The coefficient of interest, $\hat{\beta}_3$, multiplies the usual interaction term, which measures the difference in changes in mortgage denial rates between the treatment and the control groups. The OLS estimates show statistically significant, negative effects of the privacy ordinance on mortgage denial rates, where the magnitude of the effect seems above 1%. Using WLS, we see that the effect of privacy legislation is still statistically significant, but the magnitude of the effect seems to decrease to slightly below 1%. Overall the results seem to support our prediction.

The signs of other coefficients are in line with expectations. Higher-income tracts are associated with lower denial rates, and larger population shares that are below poverty are associated with higher denial rates. Tracts with higher minority population shares have significantly higher loan denial rates. This is consistent with the existing evidence on statistical discrimination against some ethnic groups in the capital market (see, e.g., Cavalluzzo et al. 2002; Blanchflower et al. 2003; and Cavalluzzo and Wolken 2005). Our results largely confirms their findings—tracts with larger African American and Hispanic population shares are associated with significantly higher denial rates. The effects of higher Asian population shares are significant but smaller.

For the tract housing variables, our expectation was that they would mainly affect foreclosure rates (to follow shortly). Here, only the household-to-housing-units ratio seems significant. A higher ratio means that demand for housing is relatively high given the housing stock. This could indicate that there is a higher level of market activity, which may increase underwriters' opportunity costs (e.g., the marginal costs of information acquisition in our theoretical model). On the other hand, there is no clearcut prediction regarding the effect of the median house age because older houses may require further inspections (e.g., to ensure that there are no structural issues and/or liens, etc.), whereas it could be slightly more difficult to appraise newer housing units.

Tables 3–5 replicate the results presented in Table 2 using different measures of loan denial rates. In Table 3, the denial rates are for conventional home-purchase loans as in Table 2, but they are calculated using the dollar amounts rather than number of loans and applications. In Tables 4 and 5, the dependent variables are denial rates calculated using refinancing loans, both in number and amount, respectively. The results turn out to be relatively consistent throughout the tables. That is, the effect of the privacy ordinance is to reduce loan denial rates relative to the control group by slightly more than 1 percentage point (when estimated using OLS) and by slightly less than 1 percentage point (when estimated using WLS).

Finally, our theoretical model implies that if denial rates go down, foreclosure rates should eventually go up, as more unqualified individuals are approved. We examine this hypothesis by using data on foreclosure rates during the financial crisis of 2007–2008. These are tract-level estimates of the percentage of mortgages that began the foreclosure process or were seriously delinquent in these two years (made available by the Department of Housing and Urban Development to implement the Neighborhood Stabilization Program, which was authorized under the Housing and Economic Recovery Act of 2008).[24] Table 6 presents reduced-form estimates of the effect of the privacy ordinance on foreclosure rates.

The coefficient estimates on the treatment dummy in columns (i)–(vi) suggest that the foreclosure rates are higher by 1.7 to 2.8 percentage points in the treatment group than in the control group. Since denial rates fell by approximately 1 percentage point in the treatment group every year during the post-intervention period, these estimates seem to be consistent with a pass-through view suggested by our hypothesis. That is, foreclosure rates rose at roughly the same rate as the fall in loan denial rates after controlling for housing-market conditions. In sum, aside from supporting our model's positive prediction regarding loan denial rates, these results render some support to our model's normative prediction — that in a competitive market with price commitments, social welfare may be higher when trade in consumers' financial information is possible.

[24] The data is publicly available at `http://www.huduser.org/nspgis/nsp_map_by_state.html`. The data is one-shot; to the authors' knowledge there is no other reliable data set that estimates foreclosure rates at the census tract level for other years.

7 Conclusions

The economic impact of trade in consumer information depends largely on how this information is used. In this paper, consumer types determined firms' costs of serving them as customers, and information obtained in the screening process was traded to another firm with correlated costs. Firms' ability to sell consumer information led to lower prices, higher screening intensities, higher rejection rates, and, perhaps more importantly, increased ex-ante social welfare. We believe that this paper provides a framework for studying certain aspects of privacy protection in markets such as financial markets where prices to qualified consumers are posted upfront. In particular, our model's welfare implication should be given consideration alongside models of price discrimination.

We closely scrutinized state and local privacy legislations, seeking ordinances that required financial institutions to obtain explicit opt-in consents prior to sharing or selling consumer information. We found statistically-reliable evidence that in areas affected by opt-in policy legislation, loan denial rates decreased by about one percentage point. This decrease was robust to the use of alternative loan denial rates as the dependent variable. Our empirical analysis also suggested that the decrease in loan approval rates may have led to a relative increase in foreclosure rates in the treatment counties. Taken along with our model's social welfare implication, these findings offer some caution against proposals to amend the privacy provisions of the GLBA by requiring an opt-in standard.

References

[1] Acquisti, A., and H. Varian. 2005. Conditioning prices on purchase history. *Marketing Science* 24(3): 367–381.

[2] Armstrong, M. 2006. Recent developments in the economics of price discrimination. Newey, Blundell, Persson, eds., *Advances in Economics and Econometrics: Theory and Applications*. Cambridge University Press.

[3] Blanchflower, D., P. Levine, and D. Zimmerman. 2003. Discrimination in the small-business credit market. *Review of Economics and Statistics* 85(4): 930–943.

[4] Burke, J., C. Taylor, and L. Wagman. 2012. Information acquisition in competitive markets: An application to the US mortgage market. *American Economic Journal: Microeconomics* 4(4): 65–106.

[5] Calzolari, G., and A. Pavan. 2006. On the optimality of privacy in sequential contracting. *Journal of Economic Theory* 130(1): 168–204.

[6] Cavalluzzo, K., L. Cavalluzzo, and J. Wolken. 2002. Competition, small business financing, and discrimination: Evidence from a new survey. *Journal of Business* 75(4): 641–679.

[7] Cavalluzzo, K., and J. Wolken. 2005. Small business loan turndowns, personal wealth, and discrimination. *Journal of Business* 78(6): 2153–2178.

[8] Chen, Y. and Z. Zhang. 2009. Dynamic targeted pricing with strategic consumers. *International Journal of Industrial Organization* 27(1): 43–50.

[9] Conitzer, V., C. Taylor, and L. Wagman. 2012. Hide and seek: Costly consumer privacy in a market with repeat purchases. *Marketing Science* 31(2): 277–292.

[10] Djankov, S., C. McLiesh, and A. Shleifer. 2007. Private credit in 129 countries. *Journal of Financial Economics* 84(2): 299–329.

[11] Doblas-Madrid, A., and R. Minetti. 2013. Sharing information in the credit market: Contract-level evidence from U.S. firms. *Journal of Financial Economics*, `http://dx.doi.org/10.1016/j.jfineco.2013.02.007`.

[12] Fudenberg, D., and J. Villas-Boas. 2006. Behavior-based price discrimination and customer recognition. T. Hendershott, ed., *Handbook on Economics and Information Systems*. North-Holland.

[13] Goldfarb, A., and C. Tucker. 2011. Privacy regulation and online advertising. *Management Science* 57(1): 57–71.

[14] Hermalin, B., and M. Katz. 2006. Privacy, property rights and efficiency: The economics of privacy as secrecy. *Quantitative Marketing and Economics* 4(3): 209–239.

[15] Hui, K.-L., and I. Png. 2006. Economics of privacy. T. Hendershott, ed., *Handbooks in Information Systems*, Vol. 1. Elsevier.

[16] Johnson, E., and D. Goldstein. 2004. Defaults and donation decisions. *Transplantation*, 78(12): 1713–1716.

[17] Mian, A., and A. Sufi. 2009. The consequences of mortgage credit expansion: Evidence from the U.S. mortgage default crisis. *Quarterly Journal of Economics* 124(4): 1449–1496.

[18] Miller, A., and C. Tucker. 2009. Privacy protection and technology diffusion: The case of electronic medical records. *Management Science* 55(7): 1077–1093.

[19] ———, and ———. 2011. Can health care information technology save babies? *Journal of Political Economy* 119(2): 289–324.

[20] Pagano, M., and T. Jappelli. 1993. Information sharing in credit markets. *Journal of Finance* 48(5): 1693–1718.

[21] ———, and ———. 2002. Information sharing, lending and defaults: Cross-country evidence. *Journal of Banking and Finance* 26(10): 2017–2045.

[22] Posner, R. 1981. The economics of privacy. *American Economic Review* 71(2): 405–409.

[23] Qian, J., and P. Strahan. 2007. How laws and institutions shape financial contracts: The case of bank loans. *Journal of Finance* 62(6): 2803–2834.

[24] Stigler, G. 1980. An introduction to privacy in economics and politics. *Journal of Legal Studies* 9(4): 623–644.

[25] Taylor, C. 2004. Consumer privacy and the market for customer information. *RAND Journal of Economics* 35(4): 631–650.

[26] Villas-Boas, J. 2004. Price cycles in markets with customer recognition. *RAND Journal of Economics* 35(3): 486–501.

Appendix

Proof of Lemma 1. Whether or not trade is permitted, a consumer's expected utility from applying to purchase the good at a price p_A is

$$U(p_A, n) = (\lambda(1-\alpha)^n + (1-\lambda))(v_A - p_A).$$

Without information trade, firm A's expected profit from each of its applicants is given by

$$\Pi^{NT}(p_A, n) = (\lambda(1-\alpha)^n + (1-\lambda))p_A - (1-\lambda)c_L^A - \lambda(1-\alpha)^n c_H^A - kn.$$

The first-order condition with respect to n (given p_A) is

$$\frac{\partial \Pi^{NT}}{\partial n} = \lambda \ln(1-\alpha)(p_A - c_H^A)(1-\alpha)^n - k \leq 0.$$

Since $\frac{\partial^2 \Pi^{NT}}{\partial n^2} = \lambda(\ln(1-\alpha))^2(p_A - c_H^A)(1-\alpha)^n < 0$, an interior solution obtains if $\frac{\partial \Pi^{NT}}{\partial n} > 0$ at $n = 0$, or $m < \lambda(c_H^A - v_A)$.

Thus, the expected utility assuming interior solutions can be rewritten as

$$U^{NT}(p_A, n) = [\frac{m}{c_H^A - p_A} + (1-\lambda)](v_A - p_A).$$

Differentiating with respect to p_A yields

$$\frac{dU^{NT}}{dp_A} = \frac{m(v_A - c_H^A)}{(c_H^A - p_A)^2} - (1-\lambda) < 0.$$

Similarly, in the case with trade, firm A's profit per application is given by

$$\Pi^T = (\lambda(1-\alpha)^n + (1-\lambda))p_A - (1-\lambda)c_L^A - \lambda(1-\alpha)^n c_H^A - kn +$$
$$(1-\lambda)\lambda(1-(1-\alpha)^n)(c_H^B - c_L^B).$$

The first-order condition with respect to n (given p_A) is

$$\frac{\partial \Pi^T}{\partial n} = \lambda \ln(1-\alpha)(1-\alpha)^n[p_A - c_H^A - (1-\lambda)(c_H^B - c_L^B)] - k \leq 0.$$

Since $\frac{\partial^2 \Pi^T}{\partial n^2} = \lambda[\ln(1-\alpha)]^2(1-\alpha)^n[p_A - c_H^A - (1-\lambda)(c_H^B - c_L^B)] < 0$, an interior solution obtains if $\frac{\partial \Pi^T}{\partial n} > 0$ at $n = 0$, or $m < \lambda[c_H^A - v_A + (1-\lambda)(c_H^B - c_L^B)]$.

Expected utility when information trade is permitted is thus given by

$$U^T(p_A, n) = [\frac{m}{c_H^A - p_A + (1-\lambda)(c_H^B - c_L^B)} + (1-\lambda)](v_A - p_A).$$

Differentiating with respect p_A yields

$$\frac{dU^T(p_A, n)}{dp_A} = \frac{m[v_A - c_H^A - (1-\lambda)(c_H^A - c_L^A)]}{[c_H^A - p_A + (1-\lambda)(c_H^B - c_L^B)]^2} - (1-\lambda) < 0.$$

Proof of Proposition 1. The proof proceeds in two steps.

Step 1: We first show that for a given p_A, under information trade, more information is collected about consumers, and the amount of information collected is decreasing in p_A.

Without information trade, given a price p_A, firm A's optimal search intensity $n_{NT}^*(p_A)$ is given by

$$n_{NT}^*(p_A) = \frac{\ln m - \ln \lambda - \ln(c_H^A - p_A)}{\ln(1-\alpha)}.$$

With trade, given a price p_A, its optimal search intensity $n_T^*(p_A)$ is

$$n_T^*(p_A) = \frac{\ln m - \ln \lambda - \ln[c_H^A - p_A + (1-\lambda)(c_H^B - c_L^B)]}{\ln(1-\alpha)}.$$

We need only compare the last term in the numerators. Since $\ln(c_H^A - p_A) < \ln[c_H^A - p_A + (1-\lambda)(c_H^B - c_L^B)]$ and the denominator $\ln(1-\alpha)$ is negative for $\alpha \in (0,1)$, we have $\frac{-\ln(c_H^A - p_A)}{\ln(1-\alpha)} < \frac{-\ln[c_H^A - p_A + (1-\lambda)(c_H^B - c_L^B)]}{\ln(1-\alpha)}$. Hence, it follows that $n_{NT}^*(p_A) < n_T^*(p_A)$. Note that if $p_{A,T} < p_{A,NT}$, then the gap between $c_H^A - p_{A,NT}$ and $c_H^A - p_{A,T} + (1-\lambda)(c_H^B - c_L^B)$ expands.

From the first-order condition of the firm's expected profit without trade,

$$\lambda(1-\alpha)^{n_{NT}(p_{A,NT})}(c_H^A - p_{A,NT}) = m,$$

it follows that $n_{NT}^*(p_{A,NT})$ is decreasing in $p_{A,NT}$.

Similarly, from the first-order condition of the firm's expected profit with trade,

$$\lambda(1-\alpha)^{n_T(p_{A,T})}[c_H^A - p_{A,T} + (1-\lambda)(c_H^B - c_L^B)] = m,$$

it follows that $n_T^*(p_{A,T})$ is decreasing in $p_{A,T}$.

Step 2: Next, we show that under information trade, consumers purchase from firms of type A at a lower price.

Since firms are competing in price, equilibrium profits will be driven down to zero. Therefore, $p^*_{A,NT}$ and $p^*_{A,T}$, along with $n^*_{NT} = n_{NT}(p_{A,NT})$ and $n^*_T = n_T(p_{A,T})$, satisfy

$$p^*_{A,NT} = \frac{(1-\lambda)c_L^A + \lambda(1-\alpha)^{n^*_{NT}}c_H^A + kn^*_{NT}}{\lambda(1-\alpha)^{n^*_{NT}} + (1-\lambda)}$$

and

$$p^*_{A,T} = \frac{(1-\lambda)c_L^A + \lambda(1-\alpha)^{n^*_T}c_H^A + kn^*_T - (1-\lambda)\lambda[1-(1-\alpha)^{n^*_T}](c_H^B - c_L^B)}{\lambda(1-\alpha)^{n^*_T} + (1-\lambda)}.$$

Since $\Pi_A^{NT}(p^*_{A,NT}, n^*_{NT}) = 0$, once information trade is permitted, profits are positive at the price and search intensity $(p^*_{A,NT}, n^*_{NT})$; that is, $\Pi_A^T(p^*_{A,NT}, n^*_{NT}) > 0$. By Lemma 1, consumers apply to the lowest priced firm. It follows that $p^*_{A,T} < p^*_{A,NT}$ must be satisfied in equilibrium. Combined with Step 1, we further have that $n^*_T > n^*_{NT}$ in a symmetric equilibrium.

Proof of Proposition 2. The proof proceeds in two steps.

Step 1: We first show that there exists a lower bound $b(c_L^B, c_H^B)$ on the difference in the equilibrium prices with and without trade, $p^*_{A,NT} - p^*_{A,T}$.

Notice that for a given search intensity n, the gain from information trade is $(1-\lambda)\lambda(1-(1-\alpha)^n)(c_H^B - c_L^B)$. Since firm A's expected profits are zero in equilibrium without trade, the following relationship must hold when information trade is permitted:

$$p^*_{A,NT} - p^*_{A,T} \geq \frac{(1-\lambda)\lambda(1-(1-\alpha)^{n^*_{NT}})(c_H^B - c_L^B)}{\lambda(1-\alpha)^{n^*_{NT}} + 1 - \lambda},$$

else a firm A possesses a profitable deviation by undercutting its competitors and attracting all consumers, while benefiting from a positive expected profit.

Step 2: Next, we compare social welfare with and without information trade.

Welfare without trade, W_{NT}, is given by

$$W_{NT} = (\lambda(1-\alpha)^{n_{NT}} + (1-\lambda))v_A - (1-\lambda)c_L^A - \lambda(1-\alpha)^{n_{NT}}c_H^A - kn_{NT}.$$

Welfare with trade, W_T, is given by

$$W_T = (\lambda(1-\alpha)^{n_T} + (1-\lambda))v_A - (1-\lambda)c_L^A - \lambda(1-\alpha)^{n_T}c_H^A - kn_T + (1-\lambda)\lambda(1-(1-\alpha)^{n_T})(c_H^B - c_L^B).$$

From firm A's first-order conditions, we have $\lambda(1-\alpha)^{n^*_{NT}}(c_H^A - p^*_{A,NT}) = m$ and $\lambda(1-\alpha)^{n^*_T}[c_H^A - p^*_{A,T} + (1-\lambda)(c_H^B - c_L^B)] = m$ for each case. Substituting n^*_{NT} and n^*_T, welfare with information trade is greater than without trade when

$$\frac{m(v_A - p^*_{A,T})}{c^A_H - p^*_{A,T} + (1-\lambda)(c^B_H - c^B_L)} + (1-\lambda)(v_A - p^*_{A,T}) \geq \frac{m(v_A - p^*_{A,NT})}{c^A_H - p^*_{A,NT}} + (1-\lambda)(v_A - p^*_{A,NT}).$$

Since $m < \lambda(c_H - v_A)$ by assumption (interior solutions), a sufficient condition for the above inequality to hold is

$$\frac{1-\lambda}{\lambda(c^A_H - v_A)}(p^*_{A,NT} - p^*_{A,T}) \geq \frac{v_A - p^*_{A,NT}}{c^A_H - p^*_{A,NT}} - \frac{v_A - p^*_{A,T}}{c^A_H - p^*_{A,T} + (1-\lambda)(c^B_H - c^B_L)}.$$

This inequality can be rewritten as

$$(1-\lambda)\frac{c^A_H - p^*_{A,NT}}{c^A_H - v_A}(p^*_{A,NT} - p^*_{A,T})(c^A_H - p^*_{A,T} + (1-\lambda)(c^B_H - c^B_L)) + \lambda(p^*_{A,NT} - p^*_{A,T})(c^A_H - v_A)$$

$$\geq \lambda(v_A - p^*_{A,NT})(1-\lambda)(c^B_H - c^B_L).$$

Since $p^*_{A,NT} > p^*_{A,T}$ and $0 < c_H - v_A < c_H - p^*_{A,NT}$, whereby $\frac{c_H - P_{A,NT}}{c_H - v_A} > 1$, a sufficient condition for the above is

$$p^*_{A,NT} - p^*_{A,T} \geq \lambda(v_A - p^*_{A,NT}).$$

Using $p^*_{A,NT} - p^*_{A,T} \geq \frac{(1-\lambda)\lambda(1-(1-\alpha)^{n^*_{NT}})(c^B_H - c^B_L)}{\lambda(1-\alpha)^{n^*_{NT}} + 1 - \lambda}$ from Step 1 and simplifying, a sufficient condition is given by

$$c^B_H - c^B_L \geq \frac{(1-\lambda+\lambda(1-\alpha)^{n^*_{NT}})(v_A - p^*_{A,NT})}{(1-\lambda)(1-(1-\alpha)^{n^*_{NT}})} \equiv \bar{c}.$$

The right-hand side of the inequality does not depend on $c^B_H - c^B_L$, giving a lower bound $\bar{c}$.

Proof of Proposition 3. The proof proceeds in two steps.

Step 1: We first show that consumers' expected utility from applying to purchase good A is decreasing in the price P_A. To save notation, we let $v_A = v_{A,L}$.

Whether or not information trade is permitted, a consumer's expected utility from applying to a firm A is now given by

$$U(p_A, n) = (\lambda(1-\alpha)^n + (1-\lambda))(v_A - p_A) - \lambda(1-\alpha)^n(v_A - v_{A,H}),$$

and for a given p_A a firm A chooses $n(p_A)$ from its first-order condition, $\lambda(1-\alpha)^n(c^A_H - p_A) = m$ (without trade) and $\lambda(1-\alpha)^n[c^A_H - p_A + (1-\lambda)(c^B_H - c^B_L)] = m$ (with trade). Thus, a consumer's expected utility is

$$U(p_A, n) = \left(\frac{m}{c^A_H - p_A + \mathbf{1}\{\text{trade}\}(1-\lambda)(c^B_H - c^B_L)} + (1-\lambda)\right)(v_{A,H} - p_A) + (1-\lambda)(v_A - v_{A,H}),$$

where $\mathbf{1}\{\text{trade}\}$ is the indicator function for information trade. Differentiation with respect to p_A yields

$$\frac{dU}{dp_A} = -\frac{m(c_H^A - v_{A,H} + \mathbf{1}\{\text{trade}\}(1-\lambda)(c_H^B - c_L^B))}{(c_H^A - p_A + \mathbf{1}\{\text{trade}\}(1-\lambda)(c_H^B - c_L^B))^2} - (1-\lambda),$$

which is negative since $c_H > v_{A,H}$ and $c_H > p_A$.

Step 2: Next, we show that a symmetric equilibrium consists of an analogous characterization to the base model. Notice that the planner's problem is equivalent to the monopolist's problem.

Without information trade, the highest price that a monopolist can charge has to satisfy consumers' participation constraints, $(1-\lambda)(v_A - p_A) + \lambda(1-\alpha)^n(v_{A,H} - p_A) \geq 0$. In equilibrium,

$$p_A^* = \frac{(1-\lambda)v_A + \lambda(1-\alpha)^n v_{A,H}}{1-\lambda+\lambda(1-\alpha)^n}.$$

The monopolist's profit is given by $(1-\lambda+\lambda(1-\alpha)^n)p_A - (1-\lambda)c_L^A - \lambda(1-\alpha)^n c_H^A - kn$. Substituting for p_A^*, the monopolist's profit can be rewritten as $\lambda(1-\alpha)^n(v_{A,H} - c_H^A) + (1-\lambda)(v_A - c_L^A) - kn$. The first term is negative and represents the cost of selling to high-cost consumers. Differentiating to obtain the first-order condition gives

$$\lambda(1-\alpha)^n(c_H^A - v_{A,H}) = m$$

A simple comparison between the optimal search condition employed by the monopolist above and the optimal search condition employed by a firm A in equilibrium, given by $\lambda(1-\alpha)^n(c_H^A - p_{A,NT}) = m$, reveals that since $p_A > v_{A,H}$, in equilibrium applicants are not sufficiently screened relative to the efficient level.

It is straightforward to show that the problem is analogous in the case of information trade. That is, comparing the optimal search condition employed by a monopolist, $\lambda(1-\alpha)^n(c_H^A - v_{A,H} + (1-\lambda)(c_H^B - c_L^B)) = m$, to the optimal search condition employed by a firm A, given by $\lambda(1-\alpha)^n(c_H^A - p_A + (1-\lambda)(c_H^B - c_L^B)) = m$, reveals the following: Since $p_{A,T} < p_{A,NT}$ is satisfied by Proposition 1, allowing for information trade would move the outcome closer to the social optimum.

Figure 1: Map of the five counties in the San Francisco-Oakland-Fremont MSA.

	Treatment Group		Control Group	
Variable	Mean	Std. Dev.	Mean	Std. Dev.
1. Pre-intervention years (2001, 2002)				
Loan denial rate (in number)	.1388	(.0883)	.1077	(.0940)
Loan denial rate (in $000's)	.1343	(.0891)	.1043	(.1026)
Refinance denial rate (in number)	.1473	(.0762)	.1389	(.1095)
Refinance denial rate (in $000's)	.1482	(.0720)	.1403	(.1080)
Median income as % of MSA median	103.6	(42.22)	97.45	(42.83)
% of population below Poverty Line	9.040	(9.060)	11.45	(10.88)
Inside central city?	.2545	(.4358)	.7487	(.4343)
% of Minority population	.4051	(.2654)	.4054	(.2638)
% of Asian population	.1290	(.1102)	.2112	(.1969)
% of Black population	.1383	(.2176)	.0889	(.1502)
% of Hispanic population	.1314	(.1128)	.1002	(.1166)
Median age of housing stock	30.36	(13.16)	40.85	(13.03)
% of owner-occupied units	.5731	(.2411)	.4017	(.2577)
Households to housing units ratio	.9566	(.0861)	.9341	(.0651)
Number of observations	1218		398	
2. Post-intervention years (2003, 2004)				
Loan denial rate (in number)	.1649	(.0826)	.1406	(.0945)
Loan denial rate (in $000's)	.1615	(.0856)	.1354	(.0919)
Refinance denial rate (in number)	.1797	(.0930)	.1785	(.1288)
Refinance denial rate (in $000's)	.1915	(.0951)	.1855	(.1256)
Median income as % of MSA median	105.6	(45.73)	98.46	(44.59)
% of population below Poverty Line	9.400	(9.255)	10.51	(7.328)
Inside central city?	.3758	(.4845)	.7974	(.4024)
% of minority population	.5135	(.2558)	.4730	(.2683)
% of Asian population	.1745	(.1426)	.2370	(.2010)
% of Black population	.1155	(.1675)	.0727	(.1246)
% of Hispanic population	.1799	(.1502)	.1294	(.1350)
Median age of housing stock	37.80	(13.58)	50.10	(13.33)
% of owner-occupied units	.5920	(.2423)	.4204	(.2556)
Households to housing units ratio	.9722	(.0852)	.9510	(.0687)
Number of observations	1288		454	

Table 1: Summary statistics by treatment status and period.

	Ordinary Least Squares			Weighted Least Squares		
Independent variable	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Treat	.0100***	.0421***	.0290***	.0175***	.0430***	.0340***
	(.0002)	(.0022)	(.0062)	(.0050)	(.0046)	(.0047)
Post	.0536***	.0490***	.0577***	.0515***	.0463***	.0525***
	(.0017)	(.0019)	(.0024)	(.0062)	(.0047)	(.0046)
Treat*Post	−.0065***	−.0143***	−.0129**	.0051	−.0090**	−.0091**
	(.0004)	(.0016)	(.0031)	(.0060)	(.0046)	(.0045)
Median income %		−.0002**	−.0001		−.0002***	−.0001***
		(.0000)	(.0001)		(.0000)	(.0000)
Below poverty %		.0018***	.0009*		.0019***	.0010***
		(.0001)	(.0004)		(.0002)	(.0002)
Inside city?		−.0039	−.0033		−.0059**	−.0063**
		(.0030)	(.0035)		(.0028)	(.0030)
Minority %		.1235***			.1197***	
		(.0237)			(.0051)	
Asian %			.0606**			.0725***
			(.0154)			(.0072)
Black %			.1695***			.1573***
			(.0220)			(.0077)
Hispanic %			.1731***			.1710***
			(.0247)			(.0092)
Median house age			−.0004*			.0000
			(.0002)			(.0001)
Owner-occupier %			−.0105			.0077
			(.0102)			(.0058)
HH to housing unit			−.1066*			−.0922***
			(.0468)			(.0227)
County dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	3326	3314	3314	3326	3314	3314
R^2	.0740	.3304	.3589	.1152	.4920	.5108

Table 2: Difference-in-Difference Models

The dependent variable is loan denial rates, measured in *number*, for home-purchase loans. For OLS estimates, standard errors are clustered by county and reported in the parentheses. For WLS estimates, the weight is the number of total home-purchase loan applications. *** Significant at 1 percent, ** Significant at 5 percent, * Significant at 10 percent.

	Ordinary Least Squares			Weighted Least Squares		
Independent variable	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Treat	.0063***	.0395***	.0270***	.0166***	.0434***	.0345***
	(.0011)	(.0016)	(.0051)	(.0050)	(.0047)	(.0048)
Post	.0522***	.0480***	.0558***	.0541***	.0484***	.0547***
	(.0045)	(.0031)	(.0035)	(.0062)	(.0049)	(.0048)
Treat*Post	−.0034	−.0115***	−.0100***	.0052	−.0086*	−.0088*
	(.0021)	(.0013)	(.0013)	(.0060)	(.0047)	(.0047)
Median income %		−.0001	.0000		−.0001***	−.0001**
		(.0001)	(.0001)		(.0000)	(.0000)
Below poverty %		.0023***	.0015*		.0021***	.0011***
		(.0002)	(.0006)		(.0002)	(.0002)
Inside city?		−.0034	−.0030		−.0037	−.0039
		(.0047)	(.0037)		(.0029)	(.0031)
Minority %		.1173**			.1162***	
		(.0260)			(.0053)	
Asian %			.0577**			.0688***
			(.0148)			(.0074)
Black %			.1588***			.1543***
			(.0197)			(.0080)
Hispanic %			.1638***			.1691***
			(.0283)			(.0095)
Median house age			−.0003			.0000
			(.0002)			(.0001)
Owner-occupier %			−.0087			.0071
			(.0087)			(.0060)
HH to housing unit			−.0888*			−.1013***
			(.0346)			(.0235)
County dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	3326	3314	3314	3326	3314	3314
R^2	.0753	.3189	.3411	.1208	.4664	.4858

Table 3: Difference-in-Difference Models

The dependent variable is loan denial rates, measured in *amount*, for home-purchase loans. For OLS estimates, standard errors are clustered by county and reported in the parentheses. For WLS estimates, the weight is the number of total home-purchase loan applications. *** Significant at 1 percent, ** Significant at 5 percent, * Significant at 10 percent.

	Ordinary Least Squares			Weighted Least Squares		
Independent variable	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Treat	−.0234***	.0166**	.0081	−.0069	.0322***	.0137***
	(.0028)	(.0037)	(.0088)	(.0050)	(.0035)	(.0034)
Post	.0879***	.0887***	.1001***	.0869***	.0755***	.0830***
	(.0060)	(.0104)	(.0131)	(.0054)	(.0033)	(.0030)
Treat*Post	−.0056	−.0153*	−.0157	.0045	−.0079**	−.0079***
	(.0052)	(.0066)	(.0094)	(.0054)	(.0032)	(.0030)
Median income %		−.0004**	−.0004***		−.0004***	−.0003***
		(.0001)	(.0001)		(.0000)	(.0000)
Below poverty %		.0022	.0012		.0023***	.0009***
		(.0011)	(.0007)		(.0001)	(.0002)
Inside city?		−.0065	−.0034		−.0005	−.0008
		(.0045)	(.0034)		(.0021)	(.0021)
Minority %		.1052**			.1139***	
		(.0262)			(.0037)	
Asian %			.0514**			.0400***
			(.0113)			(.0047)
Black %			.1455***			.1795***
			(.0191)			(.0057)
Hispanic %			.1565***			.1905***
			(.0237)			(.0064)
Median house age			−.0006*			−.0002***
			(.0002)			(.0001)
Owner-occupier %			−.0145			.0217***
			(.0180)			(.0040)
HH to housing unit			−.2212**			−.1134***
			(.0647)			(.0172)
County dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	3339	3324	3324	3339	3324	3324
R^2	.1346	.4639	.5053	.2469	.7337	.7773

Table 4: Difference-in-Difference Models
The dependent variable is loan denial rates, measured in *number*, for refinancing loans.
For OLS estimates, standard errors are clustered by county and reported in the parentheses.
For WLS estimates, the weight is the number of total refinancing loan applications.
*** Significant at 1 percent, ** Significant at 5 percent, * Significant at 10 percent.

	Ordinary Least Squares			Weighted Least Squares		
Independent variable	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Treat	−.0207***	.0174***	.0084	−.0055	.0335***	.0158***
	(.0021)	(.0037)	(.0075)	(.0050)	(.0038)	(.0037)
Post	.0909***	.0906***	.1010***	.0940***	.0821***	.0888***
	(.0042)	(.0076)	(.0104)	(.0055)	(.0035)	(.0033)
Treat*Post	−.0004	−.0109	−.0108	.0047	−.0081**	−.0080**
	(.0041)	(.0055)	(.0081)	(.0055)	(.0035)	(.0032)
Median income %		−.0004**	−.0003**		−.0003***	−.0003***
		(.0001)	(.0001)		(.0000)	(.0000)
Below poverty %		.0018	.0008		.0021***	.0009***
		(.0008)	(.0005)		(.0001)	(.0002)
Inside city?		−.0059	−.0038		.0002	−.0009
		(.0033)	(.0029)		(.0022)	(.0023)
Minority %		.1192***			.1206***	
		(.0182)			(.0040)	
Asian %			.0658**			.0505***
			(.0168)			(.0051)
Black %			.1582***			.1829***
			(.0121)			(.0062)
Hispanic %			.1703***			.1958***
			(.0216)			(.0070)
Median house age			−.0005			−.0001**
			(.0002)			(.0001)
Owner-occupier %			−.0120			.0225***
			(.0137)			(.0044)
HH to housing unit			−.1967**			−.0974***
			(.0469)			(.0188)
County dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	3338	3323	3323	3338	3323	3323
R^2	.1470	.4557	.4903	.2561	.7071	.7447

Table 5: Difference-in-Difference Models
The dependent variable is loan denial rates, measured in *amount*, for refinancing loans.
For OLS estimates, standard errors are clustered by county and reported in the parentheses.
For WLS estimates, the weight is the number of total refinancing loan applications.
*** Significant at 1 percent, ** Significant at 5 percent, * Significant at 10 percent.

	Ordinary Least Squares			Weighted Least Squares		
Independent variable	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Treat	.0164***	.0286***	.0176***	.0176***	.0287***	.0173***
	(.0034)	(.0026)	(.0027)	(.0034)	(.0024)	(.0024)
Median income %		−.0001***	−.0001***		−.0001***	−.0001***
		(.0000)	(.0000)		(.0000)	(.0000)
Below poverty %		.0005***	.0003**		.0011***	.0005***
		(.0001)	(.0001)		(.0001)	(.0002)
Inside city?		−.0013	−.0027		−.0044***	−.0045***
		(.0020)	(.0018)		(.0015)	(.0014)
Minority %		.0699***			.0564***	
		(.0042)			(.0037)	
Asian %			.0300***			.0236***
			(.0052)			(.0044)
Black %			.0886***			.0828***
			(.0061)			(.0064)
Hispanic %			.0959***			.1020***
			(.0058)			(.0061)
Median house age			−.0002***			−.0002***
			(.0001)			(.0000)
Owner-occupier %			.0134***			.0158***
			(.0041)			(.0039)
HH to housing unit			−.0272**			−.0731***
			(.0129)			(.0207)
County dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	871	868	868	863	861	861
R^2	.5955	.8113	.8410	.6376	.8526	.8788

Table 6: Reduced-Form Models

The dependent variable is foreclosure-rate estimates, measured in number, during 2007-2008. For OLS estimates, standard errors are clustered by county and reported in the parentheses. For WLS estimates, the weight is the number of owner-occupied housing units. *** Significant at 1 percent, ** Significant at 5 percent, * Significant at 10 percent.