COLUMBIA UNIVERSITY
Center for Pricing and Revenue Management

Price-Driven Adverse Selection in Consumer Lending

Robert Phillips,
Columbia University and Nomis Solutions

Robin Raffard,
Nomis Solutions

2011

http://www.cprm.columbia.edu

Price-Driven Adverse Selection in Consumer Lending

Robert Phillips

Columbia University and Nomis Solutions

Robin Raffard

Nomis Solutions

October 4, 2011

Abstract

This paper shows how price-driven adverse selection in consumer lending can be explained in terms of differential price-sensitivity between lenders who will default ("bads") and those who will not ("goods"). We show that price-driven adverse selection will characterize any market in which "goods" are more price-elastic than "bads". In such a situation, it is possible that there are some segments to whom it is unprofitable to lend at any price. This provides an explanation of the common practice of credit rationing in commercial credit markets.

1   Introduction

Adverse selection is an important characteristic of credit markets. While adverse selection can take many forms, we consider *price-driven adverse selection* – the phenomenon that, all else being equal, *raising the price of credit charged to a market segment will result in deteriorating average loss behavior for that segment.* The obverse is also true: *lowering the price charged to a segment will result in improved average loss behavior from that segment.* The change in loss behavior may manifest itself in changing probability of default, changing loss-given-default, or both. Price-driven adverse selection is a widely acknowledged phenomenon at commercial lenders, yet, in our experience, many lenders are unable to quantify its effect. An improved ability to quantify price-driven adverse selection would support better credit pricing and underwriting decisions. Ideally, this will enable improved prediction of default rate for consumer lending portfolios in which price-driven adverse selection can be significant.

We posit that price-driven adverse selection can be explained by *differential price sensitivity* – that is, customers with a lower probability of default (better credit quality) are, on the average, more price sensitive than customers with a higher probability of default. As the price is raised, customers with a low probability of default tend to seek out alternatives more quickly than those with higher probability of default – and vice versa. This effect has been understood, at least in principle, since Adam Smith (1776) who wrote of lending rates, "If the legal rate . . . was fixed so high . . ., the greater part of the money which was to be lent, would be lent to prodigals and

projectors who alone would be willing to give this higher interest. Sober people, ... would not venture into the competition ..."[1]. One consequence of price-driven adverse selection is that there are some customer segments that it is not profitable to serve at any price. As a result, in the words of Stiglitz and Weiss (1981), credit is "rationed" – that is, there are some borrowers who will not be extended credit at any price. In practical terms, lenders typically choose an "underwriting threshold" or cutoff risk and will not accept any customers whose risk exceeds the cutoff.

Empirically, many lenders have noticed that offering a less-desirable (e.g. higher price) product to the same group of prospective borrowers (a *segment*) leads to higher default rates[2]. The typical explanation for this phenomenon is *adverse selection* which can traced to either or both of two underlying causes:

1. The classic explanation for adverse selection as identified by Akerlof (1970) is *asymmetric information*. In terms of consumer lending, this means that the borrower possesses private "adverse information" not possessed by the lender – information that would increase the lender's *ex ante* probability that the borrower would default. The borrower's private information is "above and beyond" the information that the lender extracted from the borrower either through its own efforts (e.g. via a loan application) or from third parties such as the credit bureaus. An example of private adverse information might be a series of bad job reviews that has led the borrower to believe that his job is in jeopardy. This information, which is unknown to the lender, would tend to both increase the probability that this borrower will take the loan at a higher price (he is desperate to get the loan before he loses his job) and that the borrower will default relative to other customers who appear *ex ante* identical.

2. Another possibility is that borrowers who appear otherwise identical might differ in their ability to understand and manage their financial commitments. "Financially savvy" borrowers might simply be better at calculating their expected future ability to repay given the prospects that they face and therefore make better borrowing decisions. Less-savvy borrowers might be less good at estimating their future ability to repay and therefore do not make good borrowing decisions. The less-savvy would thus be more likely to accept higher rates and more likely to default due to exogenous shocks or future financial mis-management even though they appear *ex ante* identical to the lender and do not possess any private adverse information. The role of financial numeracy and cognitive skills on financial decision making is a very active topic of current research, see, for example, Bergstresser and Beshears (2009); Burks, et.al. (2009); and Gerardi, et. al. (2009). However, there appears to be no research specifically on its role in adverse selection.

We refer to these two causes collectively as "adverse selection". However, there is another explanation for default rates being affected by loan prices that is not related to adverse selection

---

[1] *The Wealth of Nations*, pg. 388. "Projectors" in this context means schemers.

[2] We note that *default rate* (DR) refers to the fraction of loans in a portfolio that default during their life. *Loss rate* (LR) is the fraction of the original balance that defaults. *Loss given default* (LGD) is the average balance lost per default. If average loan size is normalized to 1, then $LR = LGD \times DR$

– the so-called *capacity effect* (sometimes called *affordability*. For most borrowers, repaying loan obligations is a lower priority than paying for food, rent, and taxes. The difference between a borrower's monthly disposable income and her expenditure on food, rent, and other necessities is called the borrower's *capacity*. If the monthly payment for a prospective loan is a large fraction of a borrower's capacity, then an increase in the monthly payment due to a higher APR could increase the probability of default due to financial shocks such as temporary loss of employment or a large medical bill. Thus, a higher APR would tend to lead to higher default rates. For this reason, some, but by no means all, lenders estimate a borrower's capacity as part of the underwriting process (Wilkinson and Tingay [2004]). Our basic model does not incorporate the effect of affordability on default – we discuss how this can be done in the final section.

In this paper, we show that a simple model based on differential price-sensitivity can provide considerable insight into adverse selection. In particular, our model predicts that adverse selection increases with risk. As a consequence, the optimal policy for a lender who scores potential borrowers on the basis of predicted risk is to set a "cutoff score" and not to lend to borrowers with score below the cutoff – that is high risk borrowers. This is true even in the absence of a usury limit that would cap the rate that the lender can offer. We can also show that the optimal price in the presence of adverse selection is consistently lower than the optimal price ignoring adverse selection. We derive the magnitude of "indirect adverse selection" – that is, the change in risk for borrowers with the same score – based on the magnitude of "direct adverse selection" – the shift in borrower scores as rates change. This is important because direct adverse selection is immediately observable, while indirect adverse selection can take months or years to manifest.

2   Previous Research

Adverse selection has been the topic of many academic papers since its initial identification in the classic paper of Akerlof (1970). Much of this rich literature has focused on modeling and detecting the existence of adverse selection and moral hazard in insurance markets and markets for various "unique goods" such as thoroughbred racehorses (Wimmer and Chezum [2006]) and used cars (Genesove [1993]). While this research has provided considerable insight, many of the results do not directly translate from insurance and unique goods to consumer credit markets.

There have been a handful of studies that have focused on adverse selection in consumer credit markets. Edelberg (2004) used data from the Triennial Survey of Consumer Finances conducted by the US Department of the Treasury to determine the existence of adverse selection and moral hazard in mortgages and automobile loans. Her conclusions were that there is strong evidence for adverse selection in both types of credit, with weaker evidence of moral hazard. A number of researchers have found strong evidence for price-driven adverse selection in credit-card markets. Ausubel (1999) observed that customers choosing an inferior credit card product *ex ante* exhibited a higher default rate *ex post*. Using response and risk data from a credit card company, Agarwal et. al. (n.d.) found that "consumers who respond to inferior offer types (e.g., higher APR) exhibit worse credit risk characteristic than those responding to superior offer types" (pg. 3). Karlan and Zinman (2005) worked with a major South African lender to

design a randomized experiment in which 58,000 direct mail offers for credit were randomized in terms of interest rate. Rather surprisingly, they found strong evidence of adverse selection among women and moral hazard among men but not the other way around.

Models that incorporate price-driven adverse selection into consumer lending have not been widely studied. One exception is Thomas (2009), who considers the effect of a linear relationship between price and probability of default on optimal prices and underwriting decisions.

3   Adverse Selection as Differential Price Sensitivity

Consider a segment of "risky" borrowers – that is, a set of borrowers with an expected probability of default greater than zero. For any given price, p, let the demand from "goods" (those customers who will not default) be $d_g(p)$, the demand from "bads" (those customers who will default) be $d_b(p)$ and denote total demand by $d(p) = d_b(p) + d_g(p)$. We write the price-response functions for the two types of customer as $d_i(p) = d_i \bar{F}_i(p)$ where $d_i > 0$ is the demand at zero price, $\bar{F}_i(0) = 1$, and $\bar{F}_i'(p) \leq 0$ for $i = g, b$. Note that $\bar{F}_i(p)$ can be interpreted as the complementary cumulative distribution function of a distribution with density $f_i(p) = -\bar{F}_i'(p)$ – specifically, $F_i(p)$ is the fraction of the segment of size d_i whose willingness-to-pay is greater than or equal to p.

The default rate (fraction of loans that will default) at a price $p > 0$ is given by:

$$DR(p) = \frac{d_b(p)}{d(p)} = \frac{d_b \bar{F}_b(p)}{d_b \bar{F}_b(p) + d_g \bar{F}_g(p)}. \tag{1}$$

By convention, we set $DR(p) = 1$ if $\bar{F}_b(p) = \bar{F}_g(p) = 0$. We say that a segment demonstrates adverse selection at a price p if $DR'(p) > 0$.

A common measure of the risk associated with a loan or loan portfolio is the *odds of default* or, simply, *odds*[3] which is defined as the probability of default divided by the probability of non-default:

$$o(p) = DR(p)/(1 - DR(p)) = d_b(p)/d_g(p).$$

The *hazard rate*[4] (sometimes called the *failure rate*) associated with a normalized price-response function $\bar{F}$ is the density divided by the c.c.d.f. that is, $r(p) = f(p)/\bar{F}(p)$ where $r(p)$ is the hazard rate at p. An important result of this paper is that the adverse selection behavior of a segment – that is, the behavior of $DR(p)$ and $o(p)$ – is closely tied to the hazard rates of the normalized price-response curves for goods and bads.

For two distributions, f_i *is strictly smaller than* f_j *in the hazard rate order*[5] on P if $r_i(p) > r_j(p)$ for all $p \in P$. We write $f_i >_{hr} f_j$ to denote that f_i is strictly larger than f_j in the hazard

---

[3]The odds of default needs to be distinguished from the so-called *good:bad odds* which is its inverse. See Thomas (2009), page 19.

[4]The hazard rate should not be confused with the loss rate.

[5]This involves a slight abuse of notation – it is actually the random variable X_i with density function f_i that is smaller in hazard rate order than the random variable X_j. Note that, if a random variable X_i is smaller than X_j in hazard rate order then X_i will also be smaller than X_j in the usual stochastic order (Shaked and Shantikumar [2007]).

rate order. We can now characterize the existence of adverse selection in terms of the underlying price-response for "goods" and "bads".

Proposition 1 *The following statements are equivalent:*

1. *A segment demonstrates price-driven adverse selection at a price $p > 0$.*

2. $f_b >_{hr} f_g$.

3. $\epsilon_g(p) > \epsilon_b(p)$.

Proof.

Differentiating equation 1 and simplifying gives:

$$DR'(p) = \frac{d_b d_g [f_g(p)\bar{F}_b(p) - f_b(p)\bar{F}_g(p)]}{[d_b\bar{F}_b(p) + d_g\bar{F}_g(p)]^2} \tag{2}$$

For $DR'(p) > 0$, it is necessary and sufficient that the numerator be greater than 0. That is, $f_g(p)\bar{F}_b(p) > f_b(p)\bar{F}_g(p)$, or, equivalently $f_g(p)/\bar{F}_g(p) > f_b(p)/\bar{F}_b(p)$, which is equivalent to statement 2. Multiplying both sides of Inequality (2) by p gives Inequality (3). ∎

Note that substituting Equation 1 into Equation 2 and rewriting gives:

$$
\begin{aligned}
DR'(p) &= \frac{[f_g(p)\bar{F}_b(p) - f_b(p)\bar{F}_g(p)]}{\bar{F}_b(p)\bar{F}_g(p)} DR(p)(1 - DR(p)) \\
&= [\frac{f_g(p)}{\bar{F}_g(p)} - \frac{f_b(p)}{\bar{F}_b(p)}] DR(p)(1 - DR(p) \\
&= (r_g(p) - r_b(p)) DR(p)(1 - DR(p)). \tag{3}
\end{aligned}
$$

One implication of equation 3 is that price-driven adverse selection should be more severe in sub-prime markets than in prime or near-prime segments. Define the *hazard rate differential* by $z(p) = r_g(p) - r_b(p)$. Figure 1 plots the derivative of the default rate against the default rate using the relationship in Equation 3 when the hazard differential $z(p) = k$ is constant. If a segment consists entirely of "goods" ($DR(p) = 0$) or entirely of "bads" ($DR(p) = 1$), there is no adverse selection. $DR'(p)$ is maximized when $DR(p) = .5$. Since the expected bad rate is typically well below 50% even for highly sub-prime segments, adverse selection will increase with the expected risk of a portfolio – specifically, adverse selection is a stronger effect in sub-prime portfolios than in prime portfolios.

Equation 3 can be used as the starting point to estimate the relationship between the price-response functions for goods and bads.

Proposition 2 *Any segment in which $DR(0) > 0$ and $z(p) > 0$ will exhibit adverse selection. Furthermore, if there exists an $\epsilon > 0$ such that $z(p) > \epsilon$ for all $p \in P$, then $DR(p)$ will be strictly increasing in p and $\lim_{p\to\infty} DR(p) = 1$.*

Proof. We can rewrite Equation 3 as:

$$DR'(p) = z(p)DR(p)(1 - DR(p)). \tag{4}$$

If $z(p) > 0$ and $DR(0) > 0$, then $DR'(p) > 0$ for all $p \geq 0$, which establishes adverse selection. It is clear by the definition that $DR(p) \leq 1$. Now, assume that $\lim_{p\to\infty} DR(p) = \Delta$ for some $\Delta < 1$.
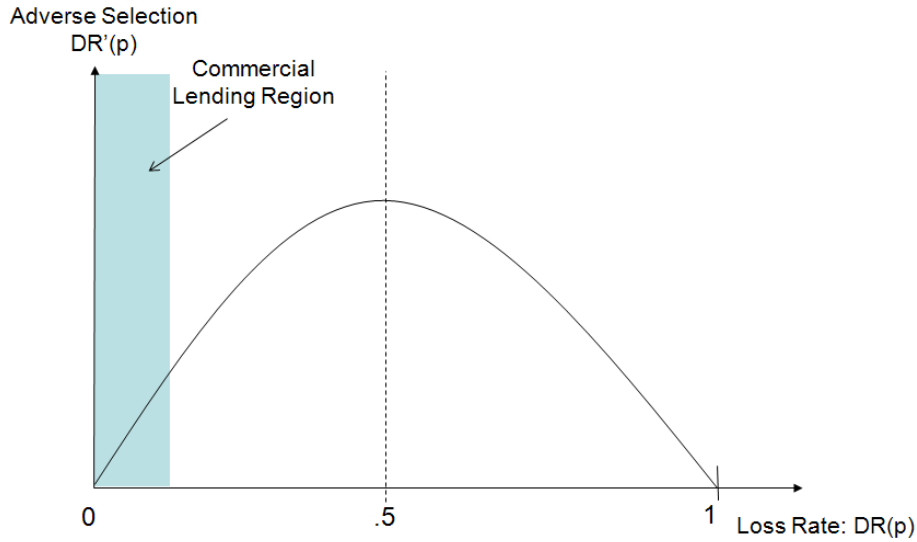
5

Figure 1: Adverse selection ($DR'(p)$) as a function of default rate $DR(p)$ when the hazard rate differential is constant.

Then, by Equation 4, $DR'(p) \geq \epsilon DR(p)(1 - DR(p)) \geq \min[\epsilon DR(0)(1 - DR(0)), \epsilon\Delta(1 - \Delta)] \equiv \delta > 0$, which implies that $DR(p) \geq DR(0) + \delta p$, which is inconsistent with a limit of Δ. ∎

We can use Equation 4 to test the implications of different assumptions on the hazard rate differential. To do this, we use the following Lemma.

Lemma 1 *For* $z(p) \geq 0$,

$$DR(p) = \frac{A(p)DR(0)}{1 - DR(0) + DR(0)A(p)} \quad where \quad A(p) = e^{\int_0^p z(p)dp}. \tag{5}$$

Proof. Rewrite Equation 4 as:

$$\frac{DR'(p)}{DR(p)(1 - DR(p))} = z(p).$$

Therefore,

$$\int_0^p \frac{DR'(p)}{DR(p)(1 - DR(p))}dp = \int_0^p z(p)dp,$$

which implies,

$$\ln\left[\frac{DR(p)}{(1 - DR(p))}\right]\Big|_0^p = \int_0^p z(x)dx,$$

which, when solved for $DR(p)$, gives the desired result. ∎

We note that one of the implications of Equation 5 is that the log odds at a price p can be computed by:

$$\ln[o(p)] = \ln[o_0] + \int_0^p z(p)dp, \quad where \quad o_0 = \frac{d_b}{d_g} \tag{6}$$

6

Equation 5 can be used to derive the default rate function for any hazard rate differential. For example, consider a constant hazard rate differential: $z(p) = k > 0$. Then, Equation 5 immediately gives:

$$DR(p) = \frac{DR(0)e^{kp}}{1 + DR(0)(e^{kp} - 1)}$$

And, from Equation 6:

$$\ln[o(p)] = kp + \ln(o_0)$$

In words, a constant hazard rate differential implies that the log odds is a linear function of price. Note that this result is independent of the actual form of the price-response functions for goods and bads.

We can also analyze the implications of a constant difference in elasticities. In this case, $z(p) = k/p$ and the solution to Equation 5 is:

$$DR(p) = \frac{p^k}{C + p^k} \quad \text{where} \quad C = \frac{1 - DR(1)}{DR(1)} = \frac{1}{o(1)},$$

which implies that log odds is linear in the logarithm of price.

$$\ln(o(p)) = k\ln(p) + \ln(o(1))$$

Finally, we note that:

$$o'(p) = \frac{DR'(p)(1 + o(p))}{(1 - DR(p))}.$$

Substituting from Equation 4 gives:

$$
\begin{aligned}
o'(p) &= z(p)DR(p)(1 + o(p)) \\
&= z(p)DR(p)/(1 - DR(p)) \\
&= z(p)o(p),
\end{aligned}
$$

which implies that $z(p) = o'(p)/o(p)$ – a relationship that is useful in estimating the price-response functions for goods and bads.

4   Pricing and Adverse Selection

We define a *customer segment* (or simply *segment*) as being characterized by the parameters d_g and d_b and the corresponding price-response functions $\bar{F}_g$ and $\bar{F}_b$. Consider a risk neutral lender whose loans are uniformly of size 1. We assume a constant *loss-given-default* ℓ with $0 < \ell \leq 1$ which specifies the fraction of the original loan balance that is lost upon default. The unit profit margin realized by the lender for each funded loan if he offers a price p is:

$$m(p) = (p - c)(1 - DR(p)) - \ell DR(p), \tag{7}$$

and his total profit is given by:

$$\Pi(p) = [(p - c)(1 - DR(p)) - \ell DR(p)]d(p), \tag{8}$$

where c is the unit cost of capital possibly including a charge for required reserves[6].

4.1 Approaches to Pricing

We consider four different approaches that a lender might take to setting price, p:

1. A *zero-risk price* is one that optimizes profitability assuming no risk, i.e. assuming that $DR(p) = 0$ for all p. We denote the zero-risk price by $\bar{p}$.

2. A *constant margin price* is equal to cost plus a factor that compensates for risk at that price plus a pre-defined margin $\mu > 0$. We denote a constant margin price with margin μ by $\tilde{p}_\mu$.

3. A *price optimized without adverse-selection* is the price that optimizes total profitability *assuming that there is no adverse selection*, that is assuming $DR'(p) = 0$ for all p. We denote the price optimized without adverse selection by $\hat{p}$.

4. The *optimal price* is the price that maximizes total profitability as defined in Equation (8) including the effect of adverse selection. We denote the optimal price by p^*

We show that for each approach except the zero-risk price, there are segments in which a price meeting the specified condition does not exist. For these segments, the magnitude of adverse selection is so great that it precludes profitability and the best policy for a profit-maximizing lender would be not to lend to the segment. We consider each of the policies in turn.

4.2 Zero-Risk Pricing

A lender who does not consider the probability of default (or believes that it is 0) would set the price $\bar{p}$ that maximizes $(p - c)d(p)$. It is easy to determine from the first order conditions that:

$$\bar{p} = c + \frac{1}{r(\bar{p})}. \tag{9}$$

where $r(p)$ is the hazard rate for total demand defined by:

$$r(p) = (1 - DR(p))r_g(p) + DR(p)r_b(p) = r_g(p) - z(p)r_b(p).$$

It is well-known (Lariviere [2006], van den Berg [2007]), that a sufficient condition for a unique $\hat{p}$ satisfying Equation 9 to exist is that $pr(p)$ be increasing in p: this is known as the *Increasing Generalized Failure Rate* (IGFR) property. IGFR distributions include the gamma, Weibull, exponential, and normal with $\mu > 0$ (Barlow and Proschan, 1965) as well as the logistic (Balakrishnan, 1992).[7]

We note that:

$$m(\bar{p}) = (1 - DR(\bar{p}))[\frac{1}{r(\bar{p})} - o(\bar{p})\ell].$$

---

[6]We note that c could also be interpreted as a "risk-free return" in an alternative investment, in which case this model is equivalent to the "rate of return model" for a single-period loan described in Thomas (2009), pp. 54-55.

[7]Note that $f_b(p)$ and $f_g(p)$ both IGFR is not sufficient to assure that $f(p) = f_b(p) + f_g(p)$ is IGFR.

It is easy to generate cases in which a zero-risk price exists, but generates a negative margin. For example, let $r(p) = k$, a constant. Then $o(p) = o_0 e^{kp}$ and the term $[1/r(\bar{p}) - o(\hat{p})\ell]$ will be negative for sufficiently high $\bar{p}$. However, from Equation 9, $\bar{p} > c$ so $\bar{p}$ can be driven arbitrarily high by increasing c.

4.3   Constant Margin Pricing

Under constant margin pricing, a lender determines a margin μ that he wishes to achieve above his cost of capital and his losses and sets the price that achieves that margin (if one exists). The margin μ could include both profit and a capital charge as imposed by regulation. Constant margin pricing of loans is commonly referred to as *risk-based pricing* and is the most common approach to pricing used by American lenders (Edelberg, 2006). In theory, a lender could choose to lend at a negative margin if, for example, he was seeking to gain short-run market share at the expense of profitability. However, we only consider the case in which $\mu \geq 0$, in which case, the constant margin price $\tilde{p}_\mu$ must satisfy:

$$\tilde{p}_\mu = c + \ell o(\tilde{p}(\mu)) + \mu/(1 - DR(\tilde{p})) \tag{10}$$

Note that $m(\tilde{p}_\mu) = \mu$ by construction, so a constant margin price – if it exists – will be profitable. However, it is not the case that $\tilde{p}_\mu$ will necessarily exist. If adverse selection is strong, losses may increase so quickly that margin begins to decline as a function of price before it reaches μ.

Proposition 3 *Define $k_\mu > 0$ as the unique root of*

$$k_\mu = \frac{e^{-[(c+\mu)k_\mu + 1]}}{o_0(\ell + \mu)} \tag{11}$$

If $z(p) > k_\mu$ for all $p > c$, then there is no price which can achieve a margin of μ and Equation (10) has no solution. If $z(p) \leq k_\mu$ for all $p > c$, then Equation (10) has a solution with corresponding margin μ.

Proof in the Appendix.

The implication of proposition 3 is that a lender cannot expect to achieve an arbitrary target margin from all segments: adverse selection creates an upper bound to the margin. A lender faced with this situation has two alternatives – he can either reduce his required margin or he can choose not to lend. However, for some segments, there may be no price at which it is profitable to lend.

Corollary 1 *Define $k_0 > 0$ as the unique root of*

$$k_0 = e^{-(ck_0 + 1)}/\ell o_0. \tag{12}$$

If $z(p) > k_0$ for all $p > c$, then there is no profitable price.

The importance of Corollary 1 is the implication that under sufficiently strong adverse selection, there is no price under which it is profitable to lend to a segment.

4.4 Optimal Pricing without Adverse Selection

We now consider a lender who optimizes his price without incorporating the effect of adverse selection. That is, the lender is aware of risk but assumes that $DR'(p) = 0$ in setting his price. The first order condition for profit maximization is:

$$\Pi'(p) = d'(p)m(p) + d(p)[1 - DR(p) - DR'(p)(p - c - \ell)] = 0 \tag{13}$$

Setting $DR'(p) = 0$ and some algebra gives us the formula for the optimal price without adverse selection, which we denote as $\hat{p}$:

$$\hat{p} = H(\hat{p}) \equiv c + \ell o(\hat{p}) + \frac{1}{r(\hat{p})}. \tag{14}$$

We note that $m(\hat{p}) = (1 - DR(\hat{p}))/r(\hat{p})$. Thus, if $\hat{p}$ exists, it will generate positive profit.

An alternative derivation of $\hat{p}$ can be obtained by assuming that the lender has initially set a price p_0 and has observed $DR(p_0)$ and $r(p_0)$. Then, not considering the dependence of DR and r on p, he seeks to set a price p_1 that satisfies the Equation 13 assuming $DR'(p) = 0$. In this case, he would update his price to p_1 according to:

$$\frac{1 - DR(p_0)}{(p_1 - c)(1 - DR(p_0)) - \ell DR(p_0)} = r(p_0)$$

or, expressed as a general recursion:

$$p_{k+1} = c + \ell o(p_k) + \frac{1}{r(p_k)}. \tag{15}$$

If the sequence defined by Equation 15 converges, it will converge to $\hat{p}$: thus $\hat{p}$ would be the price charged by a lender who continually observes default rates and adjusts his price to maximize profitability given his current default rate.

4.5 Optimal Pricing with Adverse Selection

Finally, we consider the case of a lender who fully maximizes profit. The first-order conditions for Equation 8 give:

$$1 - z(p^*)o(p^*)\ell = r_g(p^*)(p^* - c - \ell o(p^*)), \tag{16}$$

where p^* is the profit maximizing price. An implication of Corollary 1 is that, for k_0 defined in Equation 12, if $z(p) > k_0$ for $p \geq c$, then no p^* satisfying Equation 16 can exist. Lemma 2 shows that the converse is also true.

Lemma 2 *If there exists an $\epsilon > 0$ such that $\epsilon \leq z(p) < k_0$ as defined in Corollary 1 for all $p > c$, then there exists a profit-maximizing price p^* satisfying Equation 16.*

Proof. The conditions for p^* to exist are a direct consequence of proposition 3. Because $z(p) \geq \epsilon$, we must have:

$$o(p) \geq o_0 e^{\epsilon p}$$

10

from Equation 6. Thus, there exists some p^+ such that $m(p) < 0$ for $p > p^+$. Thus, there must exist a finite profit-maximizing price and it must satisfy 16. ■

Figure 4.5 shows the dependence of total profit on k in the case when the hazard rate $z(p)$ is equal to a constant k. This figure shows profit, $\Pi(p)$ as a function of p for $k = 0, .5, 1, 2$. For this example, $\ell = .9, c = 1$, and $o_0 = .25$. Lemma 2 shows that there is a profitable price for $k < k_0$ and no profitable price for $k > k_0$ where $k_0 \approx .7626$ solves Equation 12. As shown in the figure, there is a profitable region of prices for $k =$ and $k = .5$ but not for $k = 1$ and $k = 2$.
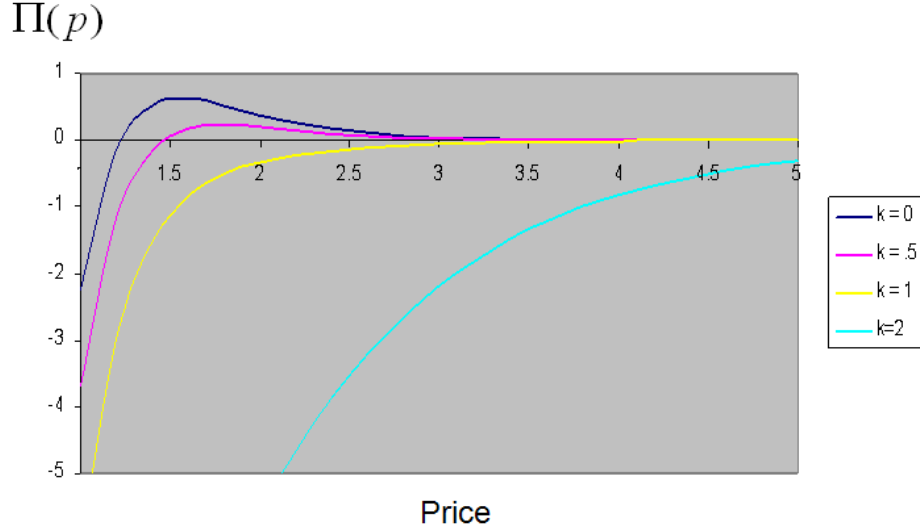

Figure 2: Dependence of total profit $\Pi(p)$ on p for various values of the hazard rate difference k when $z(p) = k$ is constant. For this example, $\ell = .9, c = 1$, and $o_0 = .25$. The normalized price responses for goods and bads are both exponential c.c.d.f.'s so that $d_i(p) = d_i e^{-\lambda p}$ for $i = g, b$. For this example $\lambda_g = 3$ and $\lambda_b = \lambda_g - k$ for $k = 0, .5, 1, 2$.

Finally, we derive the first order condition for the profit-maximizing price when it does exist.

Proposition 4 *If a profit-maximizing price exists in a segment displaying adverse selection, it satisfies:*

$$p^* = c + \frac{1}{r_g(p^*)} + \ell o(p^*)\frac{r_b(p^*)}{r_g(p^*)} \tag{17}$$

Proof. We note first that profit can be written as:

$$\Pi(p) = (p - c)d_g\bar{F}_g(p) - \ell d_b\bar{F}_b(p) \tag{18}$$

Taking the derivative and setting to 0, we must have:

$$0 = d_g\bar{F}(p^*) - (p^* - c)d_g f_g(p^*) + \ell d_b f_b(p^*)$$

11

Therefore:

$$
\begin{aligned}
p^* &= c + \bar{F}_g(p^*)/f_g(p^*) + \frac{\ell d_b f_b(p^*)}{d_g f_g(p^*)} \\
&= c + \frac{1}{r_g(p^*)} + \frac{\ell d_b \bar{F}_b(p^*) r_b(p^*)}{d_g \bar{F}_g(p^*) r_g(p^*)} \\
&= c + \frac{1}{r_g(p^*)} + \frac{\ell d_b(p^*) r_b(p^*)}{d_g(p^*) r_g(p^*)}
\end{aligned}
$$

which, noting that $d_b(p)/d_g(p) = o(p)$, gives the desired result. Here, the first step uses the fact that

$$
\frac{f_b(p)}{f_g(p)} = \frac{\bar{F}_b(p) r_b(p)}{\bar{F}_g(p) r_g(p)}.
$$

∎

To calculate the corresponding margin, we substitute from equation 17 into 7 to obtain:

$$
\begin{aligned}
m(p^*) &= \frac{1}{r_g(p^*)}\Big[1 + \ell r_b(p^*)\frac{d_b(p^*)}{d_g(p^*)}\Big]\frac{d_g(p^*)}{d(p^*)} - \ell\frac{d_b(p^*)}{d(p^*)} \\
&= \frac{1}{d(p^*)r_g(p^*)}\big[d_g(p^*) + \ell d_b(p^*)r_b(p^*) - \ell r_g(p^*)d_b(p^*)\big] \\
&= \frac{1}{d(p^*)r_g(p^*)}\big[d_g(p^*) + \ell d_b(p^*)z(p^*)\big] \\
&= \frac{1}{r_g(p^*)}\big[1 - DR(p^*) + \ell z(p^*)DR(p^*)\big],
\end{aligned}
$$

where we have used the fact that $DR(p) = d_b(p)/d(p)$ and $1 - DR(p) = d_g(p)/d(p)$.

4.6   Comparing Pricing Approaches

Table 1 shows the first-order conditions for all four pricing approaches and the corresponding total profits. We note that the zero-risk price will exist and be unique if $r(p)$ is IGFR. Existence of prices satisfying the first-order conditions for the other three approaches is only guaranteed if the hazard rate differential $z(p)$ is sufficiently low.

We can now compare results under three of the pricing strategies. Specifically, we define the following three strategies:

1. The lender sets the price $\bar{p}$ that would maximize profitability if the default rate is 0, that is, $DR(p) = 0$ for all $p \in P$.

2. The lender sets the price $\hat{p}$ that maximizes profitability assuming risk, but no adverse selection, so that $\hat{p}$ satisfies 14.

3. The lender sets the optimal price p^* that satisfies Equation 17.

Proposition 5 *Assume that $\bar{p}, \hat{p}$, and p^* as defined in Table 1 all exist. Then, $\hat{p} \geq p^*$ and $\hat{p} \geq \bar{p}$.*

12

Pricing Approach	Formula	Profit (Π)
Zero-Risk	$\bar{p} = c + \frac{1}{r(\bar{p})}$	$d_g(\bar{p})[\frac{1}{r(\bar{p})} - o(\bar{p})\ell]$
Constant Margin	$\tilde{p}_\mu = c + \ell o(\tilde{p}_\mu) + \frac{\mu d(\tilde{p}_\mu)}{1 - DR(\tilde{p}_\mu)}$	$\mu d(\tilde{p}_\mu)$
Optimal wo. Adv. Sel	$\hat{p} = c + \ell o(\hat{p}) + \frac{1}{r(\hat{p})}$	$d_g(\hat{p})/r(\hat{p})$
Optimal with Adv. Sel	$p^* = c + \frac{1}{r_g(p^*)} + \ell o(p^*)\frac{r_b(p^*)}{r_g(p^*)}$	$\frac{1}{r_g(p^*)}[d_g(p^*) + \ell z(p^*)d_b(p^*)]$

Table 1: Formulae and margins for four pricing approaches.

Proof. The conditions for the three prices are:

$$
\begin{aligned}
\bar{p} &= c + \frac{1}{r(\bar{p})} \\
\hat{p} &= c + \ell o(\hat{p}) + \frac{1}{r(\hat{p})} \\
p^* &= c + \frac{1}{r_g(p^*)} + \ell o(p^*)\frac{r_b(p^*)}{r_g(p^*)}
\end{aligned}
$$

The first price inequality is immediate because $\ell o(p) > 0$. The second price inequality requires,

$$
\frac{1}{r_g(p)} + \ell o(p)\frac{r_b(p)}{r_g(p)} \le \ell o(p) + \frac{1}{r(p)},
$$

which follows from the fact that $r_b(p) \le r(p) \le r_g(p)$ by hazard rate dominance. $\blacksquare$

It is instructive to compare optimal prices with and without adverse selection. Specifically, consider the case of a segment characterized by some $o_0, z(p)$, and ℓ such that a profitable, profit-maximizing price p^* exists with corresponding default rate $DR(p^*)$. Now, consider a segment of equal size with the same value of ℓ and c but with $z(p) = 0$ for all p and with constant default rate $DR = DR(p^*)$. In other words, the second segment does not display adverse selection but has a constant default rate equal to the optimal default rate of the first segment. By proposition x, a profit-maximizing price p^{**} must exist. Then, the following proposition describes the relationship between the optimal prices and margins in the two segments.

5 Adverse Selection and Scoring

The discussion in the previous section focused on the problem of setting a single price to a segment of prospective lenders. We assumed that *ex ante*, the lender knows the number of "goods" and "bads" in the underlying segment as well as how the "goods" and "bads" will respond to the price that he offers. However, the lender has no available mechanism for distinguishing between the "goods" and the "bads". In this situation, the lender's decisions are whether or not to offer loans to the segment and, if so, what rate to charge.

In practice, lenders do not usually face a single homogeneous segment of prospective lenders to which they will offer a single price. Rather, they seek to estimate the risk associated with lending to a particular customer. Based on the lender's estimate of the risk associated with a customer, a lender may choose whether or not to approve an application (the *underwriting* decision) and, if the lender approves the application, what price to charge. The underwriting and pricing approach followed by many lenders can be summarized in four steps:

1. A prospective borrower applies for a loan.

2. The lender estimates the risk associated with making the loan to that customer.

3. Based on the estimated risk, the lender decides whether or not to offer the loan. If the loan is offered, the lender must also decide on the price.

4. Assuming that the lender has decided to offer a loan, the borrower decides whether or not to accept the loan at the offered price.

This is a highly stylized version of the process, which can vary for different markets and lenders. For mortgages, a lender may offer a variety of different products – 15-year fixed, 30-year fixed, 5-year ARM, etc. – at different rates. The underwriting decision may not be a strict "yes" or "no" but may take the form of restrictions or additional requirements on the borrower – for example, mortgage insurance or a higher down-payment.

The most common approach in North America and Western Europe to quantify the risk associated with a prospective borrower is some form of *credit scoring*. Let $\theta_i \in \Re^N$ refer to a vector of observable customer characteristics associated with customer i. Typically, these characteristics are contained in the customer's credit file and relate to various aspects of the customer's current financial commitments. For example, an Equifax credit file in Canada contains 132 elements associated with a borrower such as number of accounts more than 30 days past due, number of defaults in last 120 days, etc. A *credit score* is a scalar function $s(\theta)$ such that *ceteris paribus* a higher value of the credit score is associated with a lower probability of default. Typically, credit scores are positive integer values. For example, the FICO score used in the United States spans a range from 300 to 800 with 300 the worst possible score and 800 a "perfect" score (Poon [2009]). Scores are commercially available through credit bureaus such as TransUnion and Equifax. In addition, lenders often generate their own scores using a combination of commercially available credit data and internal customer data. There is a substantial literature describing different approaches to computing and using credit scores; see, for example, Thomas et. al. (2002) and de Servigny and Renault (2004).

We consider the case of a lender who is offering a single price to a segment such that each customer within the segment has a credit score s with $s \in [s_L, s_U] \subset Z^+$. We assume that credit score is the only criterion used by the lender in underwriting so that all customers with $s \in [s_L, s_U]$ are approved. This guarantees that there is no selection, either adverse or otherwise, occurring across any customer characteristic not incorporated in the score. Let $d(s)$ be the number of approvals with score s and let $D = \sum_s d(s)$ be the total number of approvals. Define $g(s) = d(s)/D$ as the density function of scores across approved applications. Define $d_b(s)$ and $d_g(s)$ as the number of "bads" and "goods" respectively with score s that have been approved so that $d_b(s) + d_g(s) = d(s)$ for all s. Let $DR_A(s)$ be the expected default rate for

all loans with score s. Here and it what follow, the subscript A indicates that the quantity over *approved* loans. Then, from the point of view of the lender, $d_b(s)$ is a binomial random variable with mean $DR_A(s)d(s)$. The expected default rate across the entire set of approvals is:

$$DR_A = \sum_{s=s_L}^{s_U} DR_A(s)g(s). \tag{19}$$

We are interested in deriving the expected default rate for the *booked* loans, given $DR_A(s), g(s)$, and p. We assume in what follows that the price-sensitivity of "goods" and "bads" is independent of score – in other words $f_b(p)$ and $f_g(p)$ do not depend on score. An important implication of this assumption is that the scores provide a consistent ordering of the underlying segment in terms of default rate. That is, if the same price is offered to lenders with different scores, the segment with the lower score will have a higher default rate, no matter what price is offered. The consistent ordering property implies that $o_0(s_i) \geq o_0(s_j)$ for all $s_i, s_j, \in [s_L, s_U]$ whenever $i < j$ where $o_o(s) \equiv d_b(s)/d_g(s)$.

The expected default rate from booked loans with score s given price p is:

$$DR_B(s|p) = \frac{DR_A(s)\bar{F}_b(p)}{DR_A(s)\bar{F}_b(p) + [1 - DR_A(s)]\bar{F}_g(p)}, \tag{20}$$

where the subscript B indicates that the quantity is calculated for the portfolio of *booked* loans. We can also derive the density function of scores for booked loans given a price of p as:

$$g_B(s|p) = \frac{[DR_A(s)\bar{F}_b(p) + (1 - DR_A(s))\bar{F}_g(p)]g(s)}{DR_A\bar{F}_b(p) + (1 - DR_A)\bar{F}_g(p)}; \quad s \in [s_L, s_U] \tag{21}$$

The expected default rate across all booked loans is:

$$DR_B(p) = \sum_{s=s_L}^{s_U} DR_B(s|p)g_B(s|p). \tag{22}$$

We define *total adverse selection* $TAS(p)$ as the derivative of the expected default rate for booked loans with respect to price, that is,

$$
\begin{aligned}
TAS(p) &= \frac{dDR_B(p)}{dp} \\
&= \sum_{s=s_L}^{s_U} \frac{\partial DR_B(s|p)}{\partial p}g_B(s|p) + \sum_{s=s_L}^{s_U} DR_B(s|p)\frac{\partial g_B(s|p)}{\partial p} \tag{23} \\
&\equiv IAS(p) + DAS(p). \tag{24}
\end{aligned}
$$

We term the first term on the right hand side of Equation 23 *indirect adverse selection* and denote it by $IAS(p)$. Indirect adverse selection is the change in the default rate as price changes due to the change in default rate for borrowers with the same score. The second term on the right hand side of Equation 23 measures *direct adverse selection* $DAS(p)$ – the change in the expected default rate as price changes due to the shift in credit scores of booked loans. We note that direct adverse selection only occurs when the same price is offered to a range a credit scores. If there is only a single credit score in the segment, then there is no direct adverse selection. Note that total adverse selection is the sum of indirect adverse selection and direct selection.

15

Lemma 3 *The expected default rate for a portfolio of bookings is:*

$$DR_B(p) = \frac{DR_A \bar{F}_b(p)}{DR_A \bar{F}_b(p) + (1 - DR_A) \bar{F}_g(p)} \tag{25}$$

Proof. Substituting Equations 21 and 20 into Equation 22 and reducing gives:

$$
\begin{aligned}
DR_B(p) &= \sum_{s=s_L}^{s_U} \frac{DR_A(s) \bar{F}_b(p) g(s)}{DR_A \bar{F}_b(p) + (1 - DR_A) \bar{F}_g(s)} \\
&= \left[\frac{\bar{F}_b(p)}{DR_A \bar{F}_b(p) + (1 - DR_A) \bar{F}_g(p)} \right] \sum_{s=s_L}^{s_U} DR_A(s) g(s),
\end{aligned}
$$

substituting from 19 gives the desired result. ∎

Proposition 6 *The elements of adverse selection are:*

$$
\begin{aligned}
TAS(p) &= z(p) DR_B(p)(1 - DR_B(p)) &&(26) \\
IAS(p) &= z(p) E[DR_B(s|p)(1 - DR_B(s|p))] &&(27) \\
DAS(p) &= z(p)[E[DR_B^2(s|p)] - DR_B^2(p)] = z(p) var[DR_B(s|p)] &&(28)
\end{aligned}
$$

where the expectation and the variance are taken over booked loans.

Proof in Appendix.

The resemblance of Equation 26 to Equation 4 is clear and has a similar interpretation – namely, the magnitude of adverse selection in a segment with diverse credit scores is low when the expected default rate is close to 0 or close to 1. If $z(p)$ is constant, then the expected total average selection for a segment will follow the same relationship to expected default rate as that shown in Figure 1.

Note that Direct Adverse Selection can be calculated immediately after pricing from the distribution of scores among customers who have accepted the loan. However, Indirect Adverse Selection cannot be observed until a statistically significant number of loans default – which can be six months or more after the loans fund. It is therefore useful to have a bound on the fraction of total adverse selection that is due to indirect versus direct causes. Define $\delta_U = DR(s_U)$ and $\delta_L = DR(s_L)$ – that is δ_U is the lowest expected default rate and δ_L is the highest among scores in the segment. Also, let $\zeta(\delta) = \sqrt{\delta(1 - \delta)}$ for $\delta \in [0, 1]$.

Proposition 7 *For $DR(s_L) > 0$, the ratio of Indirect Adverse Selection to Total Adverse Selection satisfies:*

$$IAS(p)/TAS(p) \geq 1 - \delta_L - \delta_U + 2\delta_U \delta_L + 2\zeta(\delta_L)\zeta(\delta_U) \geq 1 - \delta_L \tag{29}$$

Furthermore, the first bound is tight.

Proof in Appendix.

Table 2 shows the lower bound on $IAS(p)/TAS(p)$ for various values of δ_L and δ_U. An implication of proposition 7 is that, in most cases, we would expect Indirect Adverse Selection to be higher in magnitude than Direct Adverse selection in most lending situations. Most commercial lenders do not lend to segments whose expected default rate is greater than 10% or so – and many lenders will not lend to segments whose expected default rate is greater than 2%.

| | $\ln(\delta_L)$ |
$\ln(\delta_U)$	-9	-8	-7	-6	-5	-4	-3	-2	-1	0
$-\infty$	.999877	.999665	.999088	.997521	.993262	.981684	.950213	.864655	.632121	0
-10	.999981	.999866	.999450	.998146	.994320	.983448	.953103	.869241	.638607	4.5 E-05
-9	1	.999948	.999636	.998503	.994958	.984545	.954934	.872175	.642801	.000123
-8		1	.999859	.999009	.995927	.986272	.957877	.876949	.649694	.000635
-7			1	.999616	.997301	.988901	.962522	.884650	.660990	.000912
-6				1	.998953	.992632	.969612	.896877	.679423	.002479
-5					1	.997133	.979733	.915271	.702940	.006738
-4						1	.992052	.943046	.756605	.018316
-3							1	.977162	.828739	.049787
-2								1	.926282	.135335
-1									1	.367879

Table 2: Lower bound on $IAS(p)/TAS(p)$ as a function of the highest and lowest expected default rates. The first row represents the case in which $\delta_U = 0$.

5.1 Underwriting and Pricing with Scoring

A lender with access to credit scores faces two decisions: which prospective lenders to accept and what price(s) to charge to those that are accepted. We assume as before that the scores incorporate all of the information about prospective borrowers relevant to risk. In order to characterize the underwriting and pricing decisions, we need to recognize that the marginal cost of loans may depend upon the score of the borrower. Basel II and related regulations require that banks maintain capital reserves who size in part depends on the riskiness of their portfolios. Because of this, a lender may need to set aside more reserves when lending to a risky customer (one with a lower score) than a less risky customer. The cost of the additional reserves is typically calculated as the lost risk-free return on the reserve for the term of the loan. While the detailed calculation of the capital cost that should be assigned to a particular loan can be complex, the cost will be an increasing function of risk and thus a decreasing function of score, that is $c(s_i) \geq c(s_j)$ for $i < j; s_i, s_j \in [s_L, s_U]$.

We consider first a lender who has full autonomy to determine which scores to accept and which to reject and has total freedom to charge customers with different scores differently.

Proposition 8 *A profit-maximizing lender with full pricing flexibility and a segment with scores $s \in [s_L, s_U]$ would adopt one of two policies:*

1. *Not lend to any prospective borrower.*

2. *Set a minimum $\hat{s} \in [s_L, s_U]$ and lend to all borrowers with $s \in [\hat{s}, s_U]$, charging a price to each approved borrower of $p^*(s) = c(s_i) + \frac{1}{r_g(p^*(s))} + \ell o(p^*|s)\frac{r_b(p^*(s))}{r_g(p^*(s))}$,*

where $o(p^*|s)$ are the odds at price p^* given score s.

Proof. Clearly, a profit-maximizing lender with full pricing flexibility would choose to lend to every score for which there is a profitable price and would do so at the profit-maximizing price

17

as defined in Equation 17. It remains to be shown that, if it is unprofitable to lend to score s_i, then it is unprofitable to lend to score s_{i-1}. By the consistency of the scores $o_0(s_{i-1}) \geq o_0(s_i)$. Assume that it is unprofitable to lend to score s_i. Then:

$$m(p|s_i) = (p - c(s_i))(1 - DR(p|s_i)) - \ell DR(p|s_i) \leq 0 \quad \text{for} \quad p > c(s_i),$$

where $m(p|s)$ is the margin for a loan at price p to customers of score s. But, $o_0(s_{i-1}) \geq o_0(s_i) \longrightarrow DR(p|s_{i-1}) \geq DR(p|s_i)$ for all $p > c(s_{i-1})$. Thus, $m(p|s_{i-1}) \leq m(p|s_i) \leq 0$ for all $p > c(s_{i-1}) > c(s_i)$ and it is unprofitable to lend to borrowers with score s_{i-1}. ∎

We note that the prices $p^*(s)$ will be decreasing in score.

On the other end of the spectrum, a lender might offer only a single price, independent of credit score. We can show that a cutoff policy is also optimal for such a lender.

Proposition 9 *A profit-maximizing lender who can only charge a single price would adopt one of two policies:*

1. *Not lend to any prospective borrower.*

2. *Set a minimum $\tilde{s} \in [s_L, s_U]$ and lend to all borrowers with $s \in [\tilde{s}, s_U]$.*

Proof. We observe that the profit margin $m(p|s)$ is increasing in s for all p because both $DR(p|s)$ and $c(s)$ are decreasing in s. Thus, for any p, if it is optimal to lend to customers with score s_{i-1}, it is also profitable to lend to customers with score s_i. ∎

We note the following relationship between the optimal cutoff score for the fully flexible pricing policy $\hat{s}$ and the cutoff score for the single-price policy $\tilde{s}$.

Corollary 2 *Given any distribution of scores on $[s_L, s_U]$, $\hat{s} \leq \tilde{s}$.*

Proof. Let p^* be the optimal single price. Then, it must be that $(p^* - c)(1 - DR(p|\tilde{s})) - \ell DR(p|\tilde{s}) \geq 0$, otherwise the lender could increase profitability by setting $\tilde{s} = \tilde{s} + 1$ or by not lending at all if $\tilde{s} = s_U$. Therefore, a fully-flexible lender would choose to lend to customers with score $\tilde{s}$. ∎

In reality, most lenders do not follow either of the two pure policies. A common practice is to divide scores into a smaller number of mutually exclusive *risk bands*: $[s_0, s_1), [s_1, s_n), \ldots, [s_n, s_U]$. Applicants with scores in $[s_0, s_1)$ will be rejected while applicants in the other risk bands will be accepted and charged a rate p_i where i is the index of the risk band $[s_i, s_{i+1})$. Reasons often given for charging fewer rates than credit scores include limitations in the software used to set and/or communicate prices as well as a desire to keep pricing policies simple so they can be easily understood and communicated. In any case, the results of propositions 8 and 9 and Corollary 2 hold for risk bands as for individual risk scores.

6   Conclusions and Future Research

This paper showed that a relatively simple model of differential price-sensitivity can be used to derive a model of price-driven adverse selection that explains a number of real-world phenomena. In our model, adverse selection is strongly driven by the difference in the price-response hazard

rates displayed by goods and bads. In particular, a higher hazard-rate differential leads to a higher rate of price-driven adverse selection. We show that, *ceteris paribus*, a sufficiently high hazard rate differential will lead to a situation in which it is unprofitable to lend to a market segment at any price.

One area for further investigation is the role of *affordability* or *capacity* effects in default. Our model has assumed that the decision of an individual borrower to default is independent of the rate, that is, borrowers can *ex ante* be sorted into "goods" and bads. However, it is possible that a higher monthly payment could increase stress in a borrower that would lead to an increased likelihood of default. In this case, there would not be fixed values of d_g and d_b – instead, the number of "bads" would increase with the price. Our model could be modified to incorporate such an effect by including an increasing function $h(p)$ such that $D_b(p) = d_b + h(p)d_g$ where d_b and d_g can now be interpreted as the number of "bads" and "goods" in the population assuming a price of 0, $D(p)$ is the total number of bads in the population at price p and $0 \leq h(p) \leq 1$ specifies the rate at which goods become bads as a function of price.

7 Acknowledgment

8 Appendix

Proof of proposition 3

First we note that Equation 11 has a unique root k_μ with $0 < k_\mu < e^{-1}/\ell(o_0 + \mu)$ because the right hand side of Equation 11 decreases from $e^{-1}/[o_0(\ell+\mu)]$ to 0 as k_μ increases to infinity. Substituting from Equation 5 into Equation 10 means that $\tilde{p}_\mu$ must solve:

$$
\begin{aligned}
p &= c + \ell A(p)o_0 + \mu\Big[\frac{1 - DR(0) + DR(0)A(p)}{1 - DR(0)}\Big] \\
&= c + \ell A(p)o_0 + \mu(1 + o_0 A(p))
\end{aligned}
$$

which implies $p - c - \mu = o_0 A(p)(\ell + \mu)$, where $A(p)$ is defined in Equation 5. Let $z(p) = k$ and take logarithms of both sides to obtain;

$$
\ln(p - c - \mu) = \ln(\ell + \mu) + \ln(o_0) + kp. \tag{30}
$$

The left side of Equation (30) is concave in p and strictly increases from $-\infty$ at $p = c + \mu$ to ∞. The derivative of $\ln(p-c-\mu)$ w.r.t. p decreases from ∞ at $p = c+\mu$ continuously toward 0. Taken together, these mean that, for any value of $k > 0$ there is a unique $\hat{p}(k) > c$ such that there exists a line with slope k that is tangent to $\ln(p - c - \mu)$ at $\hat{p}(k)$. Denote the intercept of this line by $(0, a(k))$. Since the slope at $\hat{p}(k)$ is k, we must have $k = [\partial \ln(p - c - \mu)/\partial p]|_{\hat{p}(k)} = 1/(\hat{p}(k)-c-\mu)$ which implies that $\hat{p}(k) = c+\mu+1/k$ and $\ln(\hat{p}(k)-c-\mu) = -\ln(k)$. It can be easily be calculated that $a(k) = -\ln(k)-k(c+\mu)-1$. As shown in Figure 3, if $\ln(\ell+\mu)+\ln(o_0) > -\ln(k)-k(c+\mu)-1$, then the line $\ln(\ell+\mu)+\ln(o_0)+kp$ will always be above the curve $\ln(p-c-\mu)$ so that $m(p) < \mu$

for all p. If, on the other hand, $\ln(\ell + \mu) + \ln(o_0) < -\ln(k) - k(c + \mu) - 1$, then the curve and the line will intersect and Equation 10 will have at least one solution. We note that:

$$\ln(\ell + \mu) + \ln(o_0) > -\ln(k) - k(c + \mu) - 1 \Rightarrow (\ell + \mu)o_0 > e^{-(k(c+\mu)+1)}/k,$$

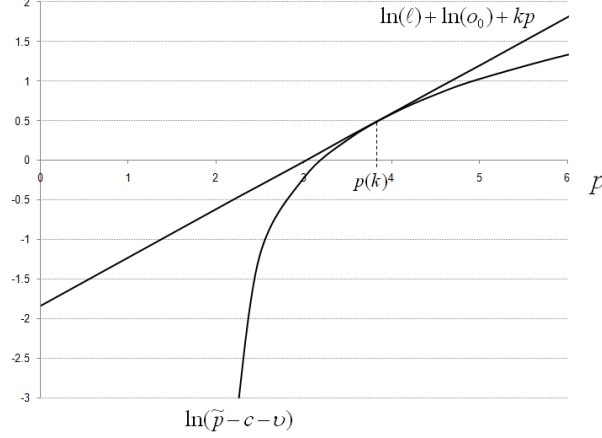which, when rearranged, gives the required conditions.

Figure 3: The line $ln(\ell) + \ln(o_0) + k^*p$ and the curve $ln(p - c - \mu)$ are shown for the case of $\ell = .8, o_0 = .2, c = 2$ and $\mu = .2$. In this case, $k^* = 6.68$.

Proof of proposition 6

We wish to prove the inequality in Equation 29. We proceed in three stages. First, we prove Lemma 4 to show that, for all discrete distributions over $[s_L, s_U]$, $r(p)$ is minimized at a two-point distribution such that $g(s) = 0$ for all $s \neq L, U$. We then calculate a two-point distribution that minimizes $r(p)$ and show that it implies the first inequality in 29. Finally, we show that the first inequality in 29 is increasing in δ_L which enables us to derive the second inequality in 29. Throughout, we assume without loss of generality that the default rate δ_i is strictly decreasing in $i = L, L + 1, \ldots U$, otherwise risk classes with the same default rates could be combined. We also assume that there are at least three risk classes – otherwise the result is trivial. Finally, for simplicity, we write $\delta_i = \delta(s_i)$ and $\mu = DR_A$ and recall that $\zeta(\delta) = \sqrt{\delta(1 - \delta)}$.

Lemma 4 *For any value of $\mu \in [\delta_U, \delta_U]$, there is a two point distribution minimizing $r(p)$ such that g_L and g_U only are non-zero.*

Proof. We note that, for any $\mu \in [\delta_U, \delta_L]$, minimizing $r(p)$ is equivalent to finding $g = (g_L, g_{L+1}, \ldots, g_U)$ that maximizes $E[DR_A^2(s)]$ over all possible distributions with mean μ. This

20

can be written as a linear program:

$$\max_{g} V(g) \equiv \sum_{i=L}^{U} g_i \delta_i^2 \tag{31}$$

$$\text{s.t.} \sum_{i=L}^{U} g_i \delta_i = \mu$$

$$\sum_{i=L}^{U} g_i = 1$$

$$g \geq 0$$

From complementary slackness, it must be that there is a g^* that solves the linear program 31 such that at most two elements of g^* are greater than 0 with all other elements equal to zero. Let $g_i^*, g_j^* \in [g_L, g_U]$ be those two values with $g_i^* < g_j^*$ and let δ_i^* and δ_j^* the corresponding default rates. Note that $\delta_j^* < \mu < \delta_i^*$, so, we must have $g_i^* \delta_i^* + (1 - g_i^*) \delta_j^* = \mu$, which implies that:

$$g_i^* = (\mu - \delta_j^*)/(\delta_i^* - \delta_j^*) \quad \text{and} \quad g_j^* = (\delta_i^* - \mu)/(\delta_i^* - \delta_j^*).$$

Substituting into the objective function in Equation 31 gives:

$$V(g^*) = \frac{(\mu - \delta_j^*)(\delta_i^*)^2 + (\delta_L^* - \mu)(\delta_j^*)^2}{\delta_i^* - \delta_j^*}$$

$$= \mu(\delta_i^* + \delta_j^*) - \delta_i^* \delta_j^*.$$

Because $\delta_j^* < \mu < \delta_i^*$, $V(g^*)$ is increasing in δ_i^* and decreasing in δ_j^*. Thus, we must have that $\delta_j^* = \delta_U$ and $\delta_i^* = \delta_L$.

For any value of μ, $r(p)$ is minimized by a two point distribution such that $g_L = (\mu - \delta_U)/(\delta_L - \delta_U)$ and $g_U = (\delta_L - \mu)/(\delta_L - \delta_U)$ and $g_i = 0$ for all $i \neq L, U$. The corresponding value of $E[DR_A^2(s)] = \mu(\delta_L + \delta_U) - \delta_L \delta_U$. This means that the lower bound on $r(p)$ can be found by solving the optimization problem:

$$\min_{\mu} \psi(\mu) \equiv \frac{\mu(1 - \delta_L - \delta_U) + \delta_L \delta_U}{\mu(1 - \mu)} \tag{32}$$

subject to $\delta_U \leq \mu \leq \delta_L$. Taking the derivative of $\psi(\mu)$ and setting equal to zero gives the first order condition:

$$0 = (1 - \delta_L - \delta_U)(\mu^*)^2 + 2\delta_L \delta_U \mu^* - \delta_L \delta_U$$

The only positive root of this equation is:

$$\mu^* = \frac{\zeta(\delta_L)\zeta(\delta_U) - \delta_L \delta_U}{1 - \delta_L - \delta_U} \tag{33}$$

The first bound in 29 can be obtained by substituting the expression for μ^* in Equation 33 into the objective function in Equation 32 and performing some algebra.

To obtain g_L and g_U, we substitute Equation 33 back into the formula for g_L to obtain:

$$
\begin{aligned}
g_L &= \frac{\zeta(\delta_L)\zeta(\delta_U) - \delta_L\delta_U - \delta_U(1 - \delta_L - \delta_U)}{(\delta_L - \delta_U)(1 - \delta_L - \delta_U)} \\
&= \frac{\zeta(\delta_L)\zeta(\delta_U) - \zeta^2(\delta_U)}{\zeta^2(\delta_L) - \zeta^2(\delta_U)} \\
&= \frac{\zeta(\delta_U)}{\zeta(\delta_U) + \zeta(\delta_L)},
\end{aligned}
$$

which means that $g_U = \zeta(\delta_L)/[\zeta(\delta_U) + \zeta(\delta_L)]$.

We now need to show the second inequality in 29. Define $\rho(\delta_U, \delta_L) = r(p)$, that is ρ is the lower bound parameterized with δ_U and δ_L. We show that $\partial\rho(\delta_U, \delta_L)/\partial\delta_U \geq 0$. Note that $\rho(\delta_U, \delta_L)$ can be written as $\rho(\delta_U, \delta_L) = [\sqrt{\delta_L\delta_U} + \sqrt{(1 - \delta_L)(1 - \delta_U)}]^2$. Since both terms that are being squared are greater than or equal to zero, we only need to consider the derivative of the terms inside the bracket. That is, we can calculate:

$$
\frac{\partial[\sqrt{\delta_U\delta_L}]}{\partial\delta_U} + \frac{\partial[\sqrt{(1 - \delta_U)(1 - \delta_L)}]}{\partial\delta_U} = \frac{\sqrt{\delta_L}}{2\sqrt{\delta_U}} - \frac{\sqrt{1 - \delta_L}}{2\sqrt{1 - \delta_U}}.
$$

This partial derivative will be greater than or equal to zero if $\sqrt{\delta_L(1 - \delta_U)} \geq \sqrt{\delta_U(1 - \delta_L)}$. But this inequality is guaranteed to hold since $\delta_L \geq \delta_U$. ∎

References

[1] Agarwal, S., Chomsisengphet, S., and Liu, C. n.d. "Adverse Selection in the Credit Card Market: Evidence from a Natural Experiment" Working Paper, The Federal Reserve Bank of Chicago, available at http://www.geocities.com/ushakri/CreditCardAdvSelection.pdf+agarwal+%22natural+experiment%22.

[2] Akerlof, G. A. 1970. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism". *Quarterly Journal of Economics* **84** 3: 488500.

[3] Altman, D., Cutler, D. M., and Zeckhauser, R. J. 1998. "Adverse Selection and Adverse Retention." *AEA Papers and Proceedings* **88** 2.

[4] Ausubel, L. 1999. "Adverse Selection in the Credit Card Market." Working Paper. University of Maryland.

[5] Balakrishnan, N. (ed.) 1992. *Handbook of the Logistic Distribution,* Marcel Dekker, Inc. New York.

[6] Barlow, R. E. and Proschan, F. 1965. *Mathematical Theory of Reliability.* John Wiley and Sons, New York.

[7] Bergstresser, D. and J. Bashears. 2009. "Who Selected Adjustable-Rate Mortgages? Evidence from the 1989-2007 Surveys of Consumer Finance," Working Paper. National Bureau of Economic Research.

[8] Burks, S. V., Carpenter, J. V., Goette, L., and Rustichini, A. 2009. "Cognitive Skills Affect Economic Preferences, Strategic Behavior, and Job Attachment," *Proceedings of the National Academy of Sciences*. **106** 19: 7745-7750.

[9] de Servigny, A. and Renault, O. 2004. *Measuring and Managing Credit Risk* McGraw Hill: New York, NY.

[10] Engelmann, B. and Rauhmeier, R. (eds.) 2006 *The Basel II Risk Parameters: Estimation, Validation, and Stress Testing.* Springer Verlag: Heidelberg, Germany.

[11] Genesove, D. 1993. "Adverse Selection in the Wholesale Used Car Market," *Journal of Political Economy* **101**: 644-645.

[12] Edelberg, W. 2004. "Testing for Adverse Selection and Moral Hazard in Consumer Loan Markets", Finance and Economics Discussion Series 2004-9, The US Federal Reserve: Washington, DC.

[13] Edelberg, W. 2006. "Risk-based Pricing of Interest Rates for Consumer Loans," *Journal of Monetary Economics* **53** 8: 2283-2298.

[14] Gerardi, K., Goette, L., and Meier, S. 2009. "Sustainable Homeownership and Numerical Ability," Working Paper. Federal Reserve Bank of Atlanta.

[15] Karlan, D. S. and Zinman, J. 2005. "Observing Unobservables: Identifying Information Asymmetries with a Consumer Credit Field Experiment," *Yale University Economic Growth Center Discussion Paper No. 911.* Available at http://ssrn.com/abstract=725563.

[16] Lariviere, M. A. 2006. "A Note on Probability Distributions with Increasing Generalized Failure Rates" *Operations Research* **54** 3: 602-604.

[17] Lariviere, M. A. and Porteus, E. 2001. "Selling to a Newsvendor: An Analysis of Price-Only Contracts". *Manufacturing Service OPerations Management.* **3** 4: 293-305.

[18] Phillips, R.L. 2005. *Pricing and Revenue Optimization.* Stanford University Press: Stanford, CA.

[19] Phillips, R. L. 2009. "An Introduction to Price Optimization for Consumer Credit," forthcoming in *Credit Control.*

[20] Poon, M. 2007. "Scorecards as Devices for Consumer Credit: The case of Fair, Isaac & Company Incorporated" in *Market Devices.* Eds. M. Callon, F. Muniesa and Y. Millo. Blackwell Publishing, London.

[21] Shaked and Shantikumar. 2007. *Stochastic Orders* Springer: New York, NY.

[22] A. Smith. 1776 *The Wealth of Nations.* Reprinted 2000. Modern Library: New York, NY.

[23] Stiglitz, J. E. and Weiss, A., 1981. "Credit Rationing in Markets with Imperfect Information," *American Economic Review*, **71** 3: 393 - 410.

[24] Thomas, L. C., 2009. *Consumer Credit Models: Pricing, Profit, and Portfolios* Oxford University Press, Oxford, UK.

[25] Thomas, L.C., Edelman, D. B., and Crook, J. N. 2002. *Credit Scoring and Its Applications* SIAM. Philadelphia, PA.

[26] van den Berg, G. J., 2007. "On the Uniqueness of Optimal Prices Set by Monopolistic Sellers," *Journal of Econometrics*. **141** 482-491.

[27] Wilkinson, G. and Tingay, J. 2004. "The use of Affordability Data – Does it add Real Value," in *Readings in Credit Scoring*, eds. Thomas, L. C.; Edelman, D. B.; Crook, J. N. Oxford University Press. Oxford, UK.

[28] Wimmer, B. S. and Chezum, B., 2006. "Adverse Selection, Seller Effort, and Selection Bias," *Southern Economic Journal*, **73** 1: 201-218.

[29] Ziya, S., Ayhan, H.. and Foley, R. D., 2004. "Relationships Among Three Assumptions in Revenue Management," *Operations Research*. **52** 5: 804-809.