A MODEL OF ADD-ON PRICING*

GLENN ELLISON

This paper examines competitive price discrimination with horizontal and vertical taste differences. Consumers with higher valuations for quality are assumed to have stronger brand preferences. Two models are considered: a standard competitive price discrimination model in which consumers observe all prices; and an "add-on pricing" game in which add-on prices are naturally unobserved and firms may advertise a base good at a low price in hopes of selling add-ons at high unadvertised prices. In the standard game price discrimination is self-reinforcing: the model sometimes has both equilibria in which the firms practice price discrimination and equilibria in which they do not. The analysis of the add-on pricing game focuses on the Chicago-school argument that profits earned on add-ons will be competed away via lower prices for advertised goods. A conclusion is that add-on practices can raise equilibrium profits by creating an adverse selection problem that makes price-cutting unappealing. Although profitable when jointly adopted, using add-on pricing is not individually rational in a simple extension with endogenous advertising practices and costless advertising. Several models that could account for add-on pricing are discussed.

I. INTRODUCTION

In many businesses it is customary to advertise a base price for a product and to try to sell additional "add-ons" at high prices at the point of sale. The quoted price for a hotel room typically does not include phone calls, in-room movies, minibar items, dry cleaning, or meals in the hotel restaurant. Personal computers advertised in weekly sales circulars typically have little memory, a low-capacity hard disk, and no separate video card. Appliance stores push extended warranties. Car rental agencies push insurance and prepaid gasoline. Manufacturers of new homes offer a plethora of upgrades and options that can add tens or hundreds of thousands of dollars to a home's price. In some cases, add-ons can be thought of as a classic price discrimination strategy: the base good and the base good plus the add-on are two different quality levels. In many of the above applications, however, there is a noteworthy feature absent from the classic price discrimination model: add-on prices are not advertised and would be costly or difficult to learn before one arrives at the point of sale.

* email: gellison@mit.edu. I would like to thank the National Science Foundation (SES-0219205) and the Institute for Advanced Study for financial support. This paper grew out of joint work with Sara Fisher Ellison. Joshua Fischman provided excellent research assistance. Jonathan Levin and Steven Tadelis provided helpful feedback.

The Quarterly Journal of Economics, May 2005

In this paper I address two questions: why do firms offer high-priced add-ons; and what difference does it make? More attention is given to the latter question. I focus on add-ons with unadvertised/unobservable prices, but also discuss the more traditional model with all prices observable. Add-ons are clearly a major source of revenue for many firms, and some consumer groups complain bitterly about them.[1] Whether we should care much about add-ons is not obvious, however. The above examples all involve fairly competitive industries. The classic Chicago-school argument would be that profits earned on add-ons will be competed away in the form of lower prices for the base good.[2]

The analysis focuses on a simple competitive price discrimination model. Two firms are located at the opposite ends of a Hotelling line. Each firm has two products for sale: a base good and an add-on. The add-on provides additional utility if consumed with the base good. The consumer population has both vertical and horizontal taste heterogeneity. There are two continuums of consumers: "high types" with a low marginal utility of income; and "low types" or "cheapskates" with a high marginal utility of income. Within each subpopulation, consumers have unit demands for the base good with the standard uniformly distributed idiosyncratic preference for buying from firm 1 or firm 2.

A crucial assumption inherent in this specification is that "high type" consumers are both more likely to buy high-priced add-ons and less likely to switch between firms to take advantage of a small price difference. This is intended to fit two types of applications. The traditional application would be discrimination between wealthy versus poor consumers (or businessmen versus tourists). Both assumptions about behavior would be natural consequences of wealthy consumers' having a lower marginal utility of income. A second "behavioral" application would be to sophisticated versus unsophisticated consumers, with unsophis-

1. Credit card companies, for example, were reported to have received \$7 billion in late payment fees in 2001. See http://money.cnn.com/2002/05/21/pf/banking/cardfees/.

2. Two formalizations of this argument can be found in the literature. Lal and Matutes [1994] develop a model of loss-leader pricing in which the Chicago view is true to an extreme—every consumer purchases the same bundle at the same price regardless of whether the prices of add-ons are or are not advertised. Verboven [1999] analyzes a model of add-on pricing with different assumptions about preferences in which add-on pricing again has no effect on profits.

ticated consumers as the high types. Unsophisticated consumers may be less sensitive to price differences because they are worse at comparison shopping.[3] They may also be more likely to intentionally or unintentionally buy overpriced add-ons; e.g., they may incur late-payment fees on their credit cards or be talked into unnecessary rental car insurance. There are, of course, other markets where willingness to pay for add-ons and sensitivity to interfirm price differences are better modeled as independent or having the reverse correlation. My analysis will not apply to these.[4]

The analysis consists primarily of contrasting the outcome of two games.[5] In the "standard pricing game" the firms publicly announce both a price for the base good and a price for a bundle containing the base good and the add-on. This game is relevant for applications in which add-on prices are as readily observable as base-good prices. For example, the price of movie-theater popcorn is just as observable to consumers as ticket prices: neither price tends to be advertised, and both are posted at the theater and known to repeat customers.

The second game, the "add-on pricing game," is intended to apply to situations in which add-on prices are not as readily observable. For example, it is typically much harder to learn a hotel's long-distance telephone charges or the quality-adjusted price of its restaurant than it is to learn the hotel's room rate. In this game, firms are assumed to only announce base-good prices, and consumers must incur a (possibly small) sunk cost to learn a firm's add-on price.[6] Consumers, of course, have rational expectations and will correctly infer the unobserved prices in any pure strategy equilibrium, but may not learn about deviations from the

3. Hausman and Sidak [2004] present evidence that less-educated and lower-income customers pay more for long distance service.

4. Most of the existing literature on competitive second-degree price discrimination with horizontal and vertical differentiation has examined the independent case. See Stole [2004].

5. In this regard the paper is similar to Verboven [1999], which also analyzes these two games in an environment with horizontal and vertical differentiation. Verboven's paper, however, is more like those of Holton [1957], Lal and Matutes [1994], and Gabaix and Laibson [2004] in that it focuses on the fact that add-ons are sold at high prices. It does not explicitly discuss whether profits earned on add-ons are competed away, and does not identify the effect highlighted in this paper. Indeed, the competition-softening effect I highlight is not present in Verboven's model due to a difference in the structure of the vertical preferences.

6. This approach follows Lal and Matutes [1994] and Verboven [1999].

equilibrium.[7] I will refer to firms playing this game as "practicing add-on pricing."

Section III of the paper shows that whether firms play the standard pricing game or practice add-on pricing is irrelevant when the preferences of the high and low types are not too different. It is not hard to construct models in which practicing add-on pricing has no effect: the simplest would be a price competition game where firms announce a price and then are allowed to charge all consumers exactly \$17 more than the price they announced. The mechanism behind my irrelevance result is similar. Because types are not very different, firms sell add-ons to everyone rather than using them to price discriminate. Hence, the unadvertised add-on is just like an extra \$17 fee that everyone pays. Every consumer buys the same good that they would buy if the add-on was simply bundled in, and the firms' profits are also unaffected. The result is a fairly easy extension of Lal and Matutes' [1994] result. I include it primarily to contrast with later results about cases where add-on pricing is important.

Section IV analyzes the more interesting case. The preferences of the high and low types are assumed to be more different. One consequence is that the "standard pricing game" becomes a model of competitive second-degree price discrimination. Proposition 2 shows that it has an equilibrium where the firms offer the base good at a low price and the base good plus the add-on at a higher price, and consumers self-select with the low types buying the base good and the high types also buying the add-on. One reason why most work on price discrimination examines monopolies is that competitive second-degree price discrimination models can be complicated; e.g., Borenstein [1985] relies heavily on numerical simulations. The model of this paper illustrates that they can also be simple: the incentive compatibility constraints all turn out to be nonbinding, so one can (almost) just analyze competition for the low and high types separately. An interesting feature of the standard pricing game is that it sometimes has multiple equilibria: there can be a second equilibrium in which the firms do not discriminate. The multiplicity reflects that the benefit from price discriminating is larger when one's rival is

7. For example, most readers of this paper are aware that items in hotel minibars are expensive, but most people choosing between hotels in Cambridge would not be aware of the change if one hotel cut the price of a minibar Coke from \$3 to \$2.

discriminating. Firm profits are higher in the discriminatory equilibrium.

Section IV also contains the main analysis of add-ons with unobserved prices. Perhaps the most important contribution of this paper is that it identifies a reason why the joint adoption of add-on pricing can raise equilibrium profits. This comes out in Proposition 3, where the add-on pricing game is shown to have an equilibrium with higher profits than any of the equilibria of the standard pricing game. I find it striking that in many of the above examples, e.g., hotels, car rental agencies, and retail stores, firms are minimally differentiated and prices must be well above marginal cost to cover fixed costs. The effects of add-on pricing may be important to understanding how firms survive in these industries.

The mechanism behind the result on unobservable add-on pricing is fairly intuitive. It is obvious that the add-ons will be very expensive—as in Diamond's [1971] original search model (and Lal and Matutes [1994]) the fact that firms will otherwise have an incentive to make the unadvertised prices ϵ higher than consumers expect leads to the add-ons being sold at the monopoly price. The more subtle question is whether the rents earned selling add-ons are fully competed away. One way to think about why they are not in the situations analyzed in Section IV is to think of the firms as intentionally creating an adverse selection problem in order to soften competition. It is standard to argue that adverse selection limits the completeness of insurance policies: if a firm offered a more complete policy, then it would attract a customer pool with a disproportionate share of sick people. When customers are heterogeneous in their marginal utility of income, there is a similar selection effect in any business: a firm that undercuts on price will attract a customer pool that contains a disproportionate share of cheapskates. If each firm sells a single good, this is a selection effect, but not an adverse selection effect: a cheapskate's money is as good as anyone else's. When firms offer multiple goods and add-on prices are unobserved, the selection becomes adverse: the unobservability results in low- and high-quality prices' being far apart, and firms do not want to attract a disproportionate share of cheapskates who only buy the low-priced base good. The incentive to cut price is reduced, and equilibrium profits go up.

Some of the welfare results are exactly as one would expect. Comparing the equilibrium of the add-on pricing game with the equilibrium of the standard pricing game (in which add-ons are cheaper), I note that high-type consumers are made worse off and low-type consumers are made better off by the practice of add-on pricing. A more interesting welfare result concerns what would happen if the government could mandate that the add-on must be provided free of charge, e.g., via laws like those mandating that landlords in Massachusetts cannot charge tenants for water and that rental car companies in California cannot charge for a spouse as an additional driver. In contrast to what one normally finds in monopoly price discrimination models (and in contrast to basic intuition about restricting consumer choice being bad), such a policy would make all consumers better off. High types gain because they pay lower prices. Low types are better off despite paying more because they get a higher quality good.

Section V turns to the question of how one can account for the prevalence of unobserved add-on pricing. The first observation is that with an endogenous choice of what to advertise and costless advertising, leaving add-on prices high and unadvertised is not individually rational. Deviating from the equilibrium described in Proposition 3 and advertising a slightly lower price for the add-on would be an effective way to steal profitable high types away from one's rival. I then informally discuss a variety of ways in which one could write down models in which add-on pricing is individually rational. The most obvious is to simply assume that advertising add-on prices is impossible or prohibitively expensive. In many applications, this seems reasonable. For example, it would be difficult to convey to potential guests that a hotel's lobby restaurant is slightly less overpriced on a quality-adjusted basis than the restaurants in competing hotels. I also discuss a potential behavioral explanation: in a population with rational and irrational consumers, the additional profits a firm could extract from rational consumers by advertising prices for add-ons might be outweighed by losses incurred when the advertisements inform irrational consumers. I find this plausible for many applications and think it provides a nice example of the potential for small changes in assumptions about the rationality of consumers to have a large effect on the outcome of a model. Gabaix and Laibson [2004] develop a similar argument formally and argue generally for the consideration of behavioral models in industrial organization.

Section VI examines a variant of the model in which only a small fraction of the population are cheapskates. In this model adopting add-on pricing is a classic example of a competitive strategy that turns lemons into lemonade. It does not just mitigate the damage that cheapskates do to equilibrium profits; it creates an environment where firms benefit from the presence of cheapskates.

Section VII relates the paper to the literatures on loss leaders, competitive price discrimination, switching costs, and other topics. Section VIII concludes.

II. Model

I consider a variant of the standard competition-on-a-line model with vertical as well as horizontal differentiation. There are two firms indexed by $i \in \{1,2\}$. Each firm sells vertically differentiated goods L and H at prices p_{iL} and p_{iH}. The firms can produce either L or H at a constant marginal cost of c.[8] Consumers differ in two dimensions. First, they differ in their marginal utility α of income. There are a unit mass of consumers with $\alpha = \alpha_h$ and a unit mass of consumers with $\alpha = \alpha_l$. Assume that $\alpha_h < \alpha_l$. Alluding to willingness to pay, I will refer to group h as the "high" types and to group l as the "low" or "cheapskate" types. Within each group customers are differentiated by a parameter $\theta \sim U[0,1]$ that reflects how well the two firms' products match their tastes.[9] Each consumer wishes to purchase at most one unit of one of the two products: he or she receives zero utility if he or she does not make a purchase and if a type (α,θ) purchases exactly one unit, his or her utility is

$$u(q_{1L},q_{1H},q_{2L},q_{2H};\alpha,\theta) = \begin{cases} v - \theta - \alpha p_{1H} & \text{if } q_{1H} = 1 \\ v - (1-\theta) - \alpha p_{2H} & \text{if } q_{2H} = 1 \\ v - w - \theta - \alpha p_{1L} & \text{if } q_{1L} = 1 \\ v - w - (1-\theta) - \alpha p_{2L} & \text{if } q_{2L} = 1. \end{cases}$$

Note the assumption of a lower marginal utility of income implies that the high types have a higher incremental valuation for high quality in money terms and are less sensitive to price differences between the firms. One could apply the model to any situation where this association makes sense even if it has nothing to do

8. Good L can be thought of as a "damaged good" as in Deneckere and McAfee [1996].

9. Note that I have set the range of the idiosyncratic taste parameter to one. To capture markets with only a small amount of horizontal differentiation, one would assume that α_l and α_h are both large.

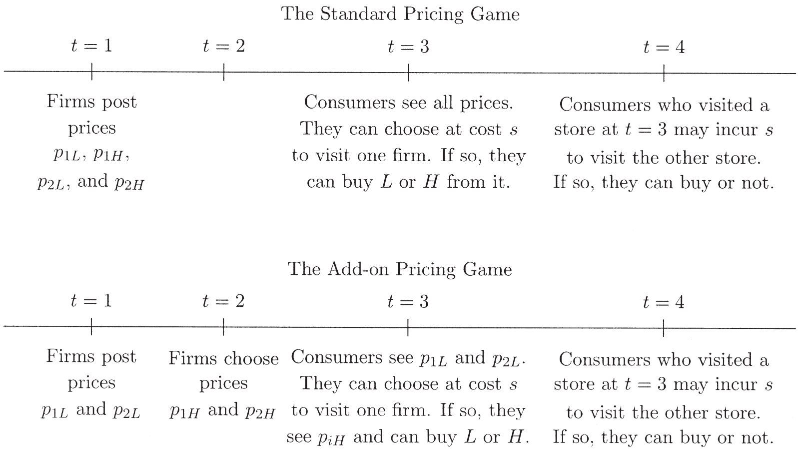

FIGURE I
Timelines for the Standard Pricing and Add-on Pricing Games

with differences in the marginal utility of wealth. For example, in the credit card market the low types could be wealthier, more sophisticated consumers who compare annual fees and interest rates more carefully when choosing between offers and who also are less likely to incur late payment fees.

Sections III and IV will contrast the outcomes of two games: a standard price competition game in which the firms simultaneously post prices for both products; and an add-on pricing game where the firms post prices for good L and reveal their prices for good H only when consumers visit the firm. Consumers will, of course, have rational expectations about the nonposted prices. To model what happens if (out of equilibrium) these expectations turn out to be incorrect, I adopt a version of Diamond's search model where consumers incur a small sunk cost of s utils in visiting a firm. This cost must be incurred to purchase from a store or to learn its price for good H. Timelines for the standard pricing game and the add-on pricing game are shown in Figure I.[10] The standard pricing game is similar, but with each firm choosing both prices at $t = 1$ and with consumers observing all prices.

10. The slightly odd-looking assumption that consumers cannot visit a store at $t = 4$ if they have not visited a store at $t = 3$ is a device to rule out equilibria in which all consumers wait until $t = 4$ to shop and thereby lose the opportunity to switch stores if prices are not as they expect.

In analyzing the model, I will look at sequential equilibria. If the model were specified as a game between the firms with consumer behavior represented by demand functions, then it would be a complete information game in which one would require subgame perfection. With consumers as players in the game, however, one must deal with consumers' beliefs about the nonposted prices. The key restriction that sequential equilibrium places on these beliefs is that if a consumer visits firm 1 at $t = 3$ and learns that it has deviated from its equilibrium strategy, then the consumer continues to believe that firm 2's nonposted price is given by firm 2's equilibrium strategy. In the standard pricing game the sequential and subgame perfect equilibria coincide.

In the model all consumers will purchase either L or H in equilibrium if v is sufficiently large. Rather than letting this paper get cluttered with statements about how large v must be at various points, I will just make the blanket assumption here that v is sufficiently large so that all consumers are served in the relevant cases and not mention it again.

III. The Lal-Matutes Benchmark: Add-ons Sold to Everyone Have No Effect

Although the Lal and Matutes [1994] article is best-known for its conclusion that multiproduct retailers may advertise a single good as a loss leader to save on per-product advertising expenditures, it also contains an irrelevance result about loss-leader pricing—it shows that the bundle of goods each consumer purchases and the total amount each consumer pays are exactly the same with loss-leader pricing as they are when all prices are advertised.[11] With no advertising costs, profits are unaffected as well. When $\alpha_h = \alpha_l$, the add-on pricing game of this paper is essentially the same as that of Lal and Matutes. In this section I verify that the irrelevance result also carries over when α_h and α_l are a bit different.

Intuitively, the result should not be surprising. When α_l and α_h are not too different, customers can forecast that they will be

11. The exact irrelevance result obviously requires special assumptions. Most notably, demands are assumed to be inelastic up to a cutoff point. I have chosen to make the same assumptions here both because it makes the model tractable and because it creates the contrast that highlights the competition-softening effect discussed in the next section.

held up for the low type's valuation for the add-on once they visit the firm. Hence, it is little different from a game where instead of announcing their prices, firms announce a number that is exactly \$17 below their price. The argument is virtually identical to that of Lal and Matutes (and tedious), so I will not try to prove it under the weakest possible assumptions and will only sketch the argument in the text leaving the details to the Appendix.

PROPOSITION 1. Suppose that $\alpha_l/\alpha_h \leq 1.6$. Write $\bar{\alpha}$ for $(\alpha_l + \alpha_h)/2$. Then, for v sufficiently large, the standard- and add-on-pricing games have pure strategy equilibria and
(a) In any symmetric pure-strategy sequential equilibrium of the standard pricing game, all consumers buy the high-quality good from the closest firm at a price of $c + 1/\bar{\alpha}$.
(b) In any symmetric pure-strategy sequential equilibrium of the add-on pricing game, all consumers buy the high-quality good from the closest firm at a price of $c + 1/\bar{\alpha}$.

Sketch of Proof. (a) In the standard pricing game, if all consumers buy H at a price of p_H^*, then if firm 1 deviates to a price p_{1H} in a neighborhood of p_H^*, its profits are

$$\pi_1(p_{1H}) = \left(1 + \frac{\alpha_l + \alpha_h}{2}(p_H^* - p_{1H})\right)(p_{1H} - c).$$

A necessary condition for Nash equilibrium is that the derivative of this expression be zero at $p_{1H} = p_H^*$. This gives $p_H^* = \frac{1}{2}(c + 1/\bar{\alpha} + p_H^*)$, which implies that any equilibrium of this form has $p_H^* = c + 1/\bar{\alpha}$.

The proof in the Appendix verifies that the various possible nonlocal deviations also do not increase a firm's profits and hence that any profile where each firm's prices satisfy $p_{iH} = c + 1/\bar{\alpha}$ and $p_{iL} \geq c + 1/\bar{\alpha} - w/\alpha_l$ does yield an equilibrium.

The one alternative form of equilibrium that is not implausible is that the firms might sell good L to the low types and good H to the high types as part of a "damaged good" second-degree price discrimination strategy as in Deneckere and McAfee [1996]. Damaged goods, however, are not always useful in price discrimination models. Good L is less valuable, but no less costly to produce. To get the low types to buy L instead of H, it must be offered at a substantially lower markup. The Appendix shows that for the parameter values considered here (with α_l and α_h not too different) this makes the damaged good strategy nonviable.

(b) In the add-on pricing model, we can think of the firm i as advertising a price p_{iL} for good L at $t = 1$ and then choosing a nonposted price $p_{iU} \equiv p_{iH} - p_{iL}$ for an upgrade from L to H at $t = 2$. As in Diamond [1971], the fact that consumers' search costs are sunk when they arrive at the firm ensures that the firms will set the monopoly price for the upgrade in equilibrium. When p_{1L} and p_{2L} are not too different and α_l and α_h are sufficiently close together, a monopolist would choose to sell the upgrade to everyone at a price of w/α_l. When p_{1L} is in a neighborhood of the symmetric equilibrium price p_L^*, consumers will correctly anticipate that if they visit firm j they will end up buying H at a price of $p_{jL} + w/\alpha_l$. Firm 1's profits are thus

$$\pi_1(p_{1L}) = \left(1 + \frac{\alpha_l + \alpha_h}{2}(p_L^* - p_{1L})\right)\left(p_{1L} + \frac{w}{\alpha_l} - c\right).$$

The FOC gives that the only possible equilibrium price is $p_L^* = c + 1/\bar{\alpha} - w/\alpha_l$.

The proof in the Appendix again verifies that there is an equilibrium in which firms charge this price for the low-quality good and that there are no other symmetric pure-strategy equilibria. QED

Note that although everyone buys good H at a price of $c + 1/\bar{\alpha}$, the price of good L is $c + 1/\bar{\alpha} - w/\alpha_l$. The proposition contains no restrictions on w, so this price can be below cost. Lal and Matutes [1994] describe their model as a model of loss leaders for this reason.

III.A. Other Specifications

In Verboven's [1999] model consumers are horizontally and vertically differentiated with the taste for quality being independent of the logit preferences for horizontal characteristics. The complete irrelevance result of Lal and Matutes does not hold. Low types pay less in the add-on pricing game than in the standard pricing game, and high types pay more. The profits part of the irrelevance result nonetheless carries over. The higher price paid by one group exactly offsets the lower prices paid by the other, and the firms' profits are identical in the two games.[12]

The model of this paper would be similar to Verboven's if the low and high types were instead assumed to differ only in a

12. I thank Frank Verboven for this observation.

marginal-utility-of-quality parameter. With such an assumption, the firms' profits would again be the same in the standard and add-on pricing games. In the standard pricing game all consumers would buy the high-quality good at a price of $c + 1/\alpha$.[13] In the add-on pricing game the profits earned selling the unadvertised add-on will be exactly offset by reductions in the price of the base good. The profit-neutrality result is much stronger in such a model than in my model—it holds also for the parameter values discussed in the next section where I show that profits are higher in the add-on pricing game.

IV. Discriminatory Add-on Pricing Softens Competition

This section analyzes a more interesting case: the preferences of the high and low types are more different so that there is a greater incentive to price discriminate. There are two main observations. First, the standard pricing game becomes a model of competitive price discrimination with multiple equilibria. Second, the adoption of add-on pricing can soften competition. The observations are brought out by comparing the outcomes of the add-on and standard pricing games for a common set of parameters.

Proposition 2 contains results on the standard pricing game. The equilibrium described in part (a) illustrates that the standard pricing game becomes a tractable competitive price discrimination model. The prices at which the two products are sold are those that would prevail if the firms were competing in two entirely separate Hotelling markets: one in which good L is sold to a population of low types; and one in which good H is sold to a population of high types. The incentive compatibility constraints that play such a crucial role in monopoly price discrimination models are nonbinding for the range of parameters under consideration.

Part (b) gives the equilibrium multiplicity result: for a subset of the parameter values the model also has an equilibrium where the firms sell the add-on to everyone and do not price discriminate. This illustrates a complementarity in practicing price discrimination: it is optimal for the firm to discriminate when its rival is discriminating and optimal for it not to discriminate when its rival is not discriminating. Intuitively, the reason why this

13. See Stole [2004] for more on this model.

occurs is that it is better to sell the high-quality good to everyone (because it is more highly valued and no more costly to produce) unless the optimal prices for the high-quality good in the two populations are very different. When a firm's rival is charging the same price in both populations, the firm's unconstrained best-response prices will be similar in the two populations, so it is optimal to choose an in-between price and sell the high-quality good to everyone. When a firm's rival discriminates and charges more to high types, the firm's unconstrained best-response prices are farther apart, and it is optimal to discriminate.

PROPOSITION 2. Suppose that $\alpha_l/\alpha_h \in [3.2,10]$. Let $\underline{w} = \alpha_h(1/\alpha_h - 1/\alpha_l)$. Let $\bar{w} = 4(\bar{\alpha}/\sqrt{\alpha_l\alpha_h} - 1)$. Then $\bar{w} > \underline{w}$ and for $w \in (\underline{w},\bar{w})$,

(a) The standard pricing game has a "discriminatory" sequential equilibrium in which the low types buy good L from the closest firm at a price of $c + 1/\alpha_l$ and the high types buy good H from the closest firm at a price of $c + 1/\alpha_h$.

(b) If $\alpha_l/\alpha_h > 6.4$ (and for some other parameter values), the standard pricing game also has a sequential equilibrium in which all consumers buy good H from the closest firm at a price of $c + 1/\bar{\alpha}$. There are no other symmetric pure-strategy equilibria.

Sketch of Proof. (a) When the firms choose $p_{iL} = c + 1/\alpha_l$ and $p_{iH} = c + 1/\alpha_h$ in the standard pricing game, high types will buy good H rather than good L because $\alpha_h(p_{iH} - p_{iL}) = \alpha_h(1/\alpha_h - 1/\alpha_l) = \underline{w} < w$.

After some algebra one can also see that the $w < \bar{w}$ condition is sufficient to ensure that low types prefer L to H. For small deviations in price it is as if the firms were playing two separate competition-on-a-line games: one involving selling good L to low types and one involving selling good H to high types. The standard calculations for these games show that a small change in p_{1L} or p_{1H} will not increase firm 1's profits.

Completing the proof that this is an equilibrium requires showing that firm 1 also cannot increase its profits by selling H to members of both populations. When w is large enough, such a deviation is profitable—good L is sufficiently damaged so as to make the benefits from selling the low types a better product outweigh the price discrimination benefits of selling L. The upper bound $\bar{w}$ was chosen to ensure that a deviation that involves selling only H is not profitable. The Appendix contains this calculation along with other details of the argument above.

(b) Any strategy profile with $p_{1H} = p_{2H} = c + 1/\bar{\alpha}$ and $p_{iL} > c + 1/\bar{\alpha} - w/\alpha_l$ satisfies the first-order conditions for profit maximization just as it did in Proposition 1. To show that this is indeed an equilibrium, it remains only to show that it is not profitable to make various nonlocal deviations. The most natural of these is raising the price of good H and selling good L at a lower price to the low types. The Appendix shows that no nonlocal deviations are profitable when α_l/α_h is above the bound given in the statement of the proposition. Uniform pricing equilibria also exist when α_l/α_h is smaller provided that w is sufficiently large. For some parameter values covered in part (a), however, there is no pure strategy equilibrium with good H sold to everyone. The Appendix also contains a verification that there are no other pure strategy equilibria. QED

Remark

1. I have not tried to state the propositions of this section for the broadest possible sets of parameter values. The set covered here is sufficient to illustrate the observations I want to bring out and simplifies the algebra. The lower bound $w > \underline{w}$ makes the high type's incentive compatibility constraint nonbinding. The upper bound $w < \bar{w}$ makes the low type's incentive compatibility constraint nonbinding and is also used to ensure that the firms are not tempted to sell good H to everyone.

Proposition 3 characterizes behavior in the add-on pricing game for the same set of parameter values. Because α_l is more than twice as large as α_h, it is more profitable to sell the add-on to high types at a price of w/α_h than to sell it to everyone at a price of w/α_l. Part (a) describes the equilibrium that seems most reasonable. Part (b) notes another possibility that one could imagine might also arise in some industries—an expectations trap in which consumer beliefs that add-ons will be sold at low prices make it impossible for firms to charge high prices.

PROPOSITION 3. Suppose that $\alpha_l/\alpha_h \in [3.2,10]$ and $w \in (\underline{w},\bar{w})$. Then,

(a) The add-on pricing game has a sequential equilibrium in which the firms set $p_{iL} = c + 1/\bar{\alpha} - w/2\bar{\alpha}$, low types buy good L from the closest firm, and high types pay w/α_h more to upgrade to good H.

(b) This is the only symmetric pure strategy equilibrium in which the equilibrium played at $t = 2$ is always that which is optimal for the firms. The game has other equilibria for some of the parameter values, including one in which firms sell good H to everyone at a price of $c + 1/\bar{\alpha}$.

Sketch of Proof. (a) In the add-on pricing game the lower bound on α_l/α_h ensures that when p_{1L} and p_{2L} are close together, the best equilibrium for the firms has both firms pricing the add-on at $p_{iU} = w/\alpha_h$ at $t = 2$. Firm 1's profit function (for small deviations) is thus

$$\pi_1(p_{1L},p_{2L}) = \left(\frac{1}{2} + \frac{\alpha_l}{2}(p_{2L} - p_{1L})\right)(p_{1L} - c) + \left(\frac{1}{2} + \frac{\alpha_h}{2}(p_{2L} - p_{1L})\right)\left(p_{1L} + \frac{w}{\alpha_h} - c\right).$$

Considering the first-order conditions for firm 1's profit maximization shows that $p_{iL} = c + 1/\bar{\alpha} - w/2\bar{\alpha}$ is the only possible first-period price in a symmetric pure-strategy equilibrium. This profit function is concave, so no price p_{1L} for which the profit function applies can increase firm 1's profits. It remains only to show that firm 1 cannot increase its profits via a larger deviation, for example, with a larger reduction in price that will let it sell to all of the low types (which yields a higher profit than the above formula gives when p_{1L} is below cost). The assumption that $\alpha_l/\alpha_h < 10$ in the proposition is a convenient way to ensure that the profile is indeed an equilibrium. (Weaker conditions could be given.) Details are in the Appendix.

(b) The uniqueness claim is immediate from the uniqueness of the solution to the first-order condition corresponding to the profit function above.

To see that the nondiscriminatory profile is an equilibrium for some of the parameter values covered under Proposition 3, note that if consumers' beliefs are that the firms set $p_{iL} = c + 1/\bar{\alpha} - w/\alpha_l$ at $t = 1$ and then set $p_{iU} = w/\alpha_l$ on the equilibrium path and after nearby deviations, then if firm 1 raises its upgrade price at all at $t = 2$, all low types who visit will refuse to buy the upgrade, and some high types will decide to purchase nothing and visit firm 2 at $t = 4$. When the search cost s is small, firm 1's profits will be approximately equal to

$$\pi_1(p_{1L},p_{1H}) = \left(\frac{1}{2} + \frac{\alpha_l}{2}\left(c + \frac{1}{\bar{\alpha}} - \frac{w}{\alpha_l} - p_{1L}\right)\right)(p_{1L} - c) + \left(\frac{1}{2} + \frac{\alpha_h}{2}\left(c + \frac{1}{\bar{\alpha}} - p_{1H}\right)\right)(p_{1H} - c).$$

This is precisely the expression I considered when assessing whether in the standard pricing game there was any profitable deviation from a profile which sold good H to all consumers at a price of $c + 1/\bar{\alpha}$. The fact that that deviation is not profitable implies that the deviation under consideration here is not profitable. QED

Remarks

1. Good L can easily be sold at a loss in the add-on pricing model. Its price, $c + 1/\bar{\alpha} - w/2\bar{\alpha}$, is less than c whenever $w > 2$. The upper bound $\bar{w}$ on w in the proposition is greater than 2 when $\alpha_l/\alpha_h > (7 + \sqrt{45})/2 \approx 6.85$.
2. The equilibrium multiplicity noted in part (b) is a consequence of the fact that Diamond's result about monopoly pricing being the unique equilibrium of the search game needs a concavity assumption on the profit function. The discrete set of types in my model yields a nonconcave profit function. I think that the idea that firms may sometimes be unable to set high add-on prices because consumers expect not to be held up is intriguing, although the nonconcave profit function it requires clearly will not be reasonable for many applications.
3. Something I did not discuss in detail in the proposition is that the add-on game will typically have many other equilibria with higher and lower profit levels. The reason is that one can deter deviations from many first-period prices by assuming that firms set $p_{iU} = w/\alpha_h$ on the equilibrium path, but revert to the equilibrium with $p_{iU} = w/\alpha_l$ following any deviation. These equilibria seem unreasonable because there is no natural reason why second-period play should shift from one equilibrium to the other following a small change in first-period play. Essentially, the shifts between equilibria are being used as arbitrary punishments to deter deviators as in the finite horizon folk theorem of Benoit and Krishna [1985].

I now present a couple of corollaries characterizing profits and welfare. To avoid repeating lengthy phrases, I will refer to the equilibrium of the standard pricing game given in part (a) of Proposition 2 as the "discriminatory equilibrium of the standard pricing game" and write $\pi^{s,d}$ for the profits each firm receives in this equilibrium. I refer to the equilibrium described in part (b) of Proposition 2 as the "nondiscriminatory equilibrium of the standard pricing game" and write $\pi^{s,nd}$ for the profits in it. I refer to the equilibrium described in part (a) of Proposition 3 as the "add-on pricing equilibrium" and write π^{a} for the profits in it.

There are two main results on profits. First, the invention of good L can increase profits even if firms do not practice add-on pricing—profits in the discriminatory equilibrium of the standard pricing game are higher than the profits in the nondiscriminatory equilibrium of the standard pricing game. Second, the profits in the add-on pricing equilibrium are even higher.

COROLLARY 1. Suppose that $\alpha_l/\alpha_h \in [3.2,10]$ and $w \in (\underline{w},\bar{w})$. Then,

$$\pi^{a} > \pi^{s,d} > \pi^{s,nd}$$

with

$$\pi^{a} - \pi^{s,d} = \frac{\alpha_l - \alpha_h}{4\bar{\alpha}\alpha_h}(w - \underline{w}), \quad \pi^{a} - \pi^{s,nd} = \frac{\alpha_l - \alpha_h}{4\bar{\alpha}\alpha_h} w,$$

$$\text{and} \quad \pi^{s,d} - \pi^{s,nd} = \frac{\alpha_l - \alpha_h}{4\bar{\alpha}\alpha_h} \underline{w}.$$

Proof. In the discriminatory equilibrium of the standard pricing game, each firm's profit is

$$\pi^{s,d} = \frac{1}{2}\left(\frac{1}{\alpha_l} + \frac{1}{\alpha_h}\right).$$

In the nondiscriminatory equilibrium of the standard pricing game, each firm's profit is

$$\pi^{s,nd} = 1/\bar{\alpha}.$$

In the add-on pricing equilibrium, each firm's profit is

$$\pi^{a} = \frac{1}{2}\left(\frac{1}{\bar{\alpha}} - \frac{w}{2\bar{\alpha}}\right) + \frac{1}{2}\left(\frac{1}{\bar{\alpha}} - \frac{w}{2\bar{\alpha}} + \frac{w}{\alpha_h}\right).$$

Taking differences and simplifying gives the desired results. QED

One can get some intuition for why profits are higher in the add-on pricing equilibrium than in the discriminatory equilibrium of the standard pricing game by thinking about how the equilibrium might be reached by alternately iterating the best response functions of the two firms starting from the equilibrium prices in the standard pricing game. Suppose that firm 2 sets $p_{2L} = c + 1/\alpha_l$ and $p_{2H} = c + 1/\alpha_h$. In the standard pricing game, firm 1's best response is to match these prices. In the add-on pricing game, those prices are not feasible because firm 1 must set $p_{1H} = p_{1L} + w/\alpha_h$. The $w > \underline{w}$ assumption is the condition for w/α_h to be greater than $1/\alpha_h - 1/\alpha_l$, i.e., for firm 1 to be constrained in the add-on pricing game to choose prices that are farther apart than it would like. Profits decrease as p_{1L} is moved away from $c + 1/\alpha_l$ and as p_{1H} is moved away from $c + 1/\alpha_h$, so firm 1 will choose $p_{1L} < p_{2L}$ and $p_{1H} > p_{2H}$. Why do average prices increase? Roughly, prices are reduced less in the small market because cutting prices to the low types is more costly than increasing prices to the high types.

Formally, the constrained best-response prices satisfy the first-order condition: $(d\pi_{1L}/dp_{1L})(p_{1L}) = -(d\pi_{1H}/dp_{1H})(p_{1H})$. Approximating the derivatives in a neighborhood of p_{2L} and p_{2H} using a second-order Taylor expansions gives

$$\frac{p_{2L} - p_{1L}}{p_{1H} - p_{2H}} = \frac{d^2\pi_{1H}/dp_{1H}^2}{d^2\pi_{1L}/dp_{1L}^2} = \frac{Q''_H(p_{1H})(p_{1H} - c) + 2Q'_H(p_{1H})}{Q''_L(p_{1L})(p_{1L} - c) + 2Q'_L(p_{1L})}.$$

In the competition-on-a-line model, firm-level demand curves are linear, so the Q'' terms are zero and the fact that the low types' demand is more price-sensitive implies that $|p_{1L} - p_{2L}| < |p_{1H} - p_{2H}|$, i.e., firm 1's low quality price is moved down from p_{2L} less than its high-quality price is moved up from p_{2H}, and hence firm 1's best response is to set an average price above the average price in the equilibrium of the standard pricing game. Firm 2's unconstrained best response has a higher average price when firm 1's average price is higher, so one would expect that iterating the best responses would lead to an equilibrium with higher average prices. For more general demand curves, the calculation suggests that similar results should obtain unless there are substantial curvatures of a particular form in the demand curves; i.e., unless $Q''_H(p_{1H})$ is very different from $Q''_L(p_{1L})$ or both are large and negative.

A good way to think about the difference in profits between the add-on pricing game and the nondiscriminatory equilibrium of the standard pricing game is to regard it as resulting from the firms' having created an adverse-selection problem. In both games firms will choose their prices so that the profits earned from the marginal customers attracted by a dp price cut are exactly offset by the loss of revenue on inframarginal customers. The revenue loss is identical across games—it is equal to Qdp, and in each game each firm attracts half of the customers. Hence, the profits on the marginal consumers attracted by a price cut must also be identical in the two games. The number of consumers attracted by a small price cut is again identical across games, so per-consumer profit on the marginal consumers is also identical. In the standard pricing game the firms' profits on the marginal consumers are the same as their profits on the average consumer: firms make $p_H^* - c$ on every consumer. In the add-on pricing game the profits on the marginal consumers are much lower than the profits on the average consumer because of the adverse selection effect—the marginal consumers attracted by a small price cut are disproportionately low types, whereas the full customer pool is equally split. Hence, when marginal profits are equal in the two games, average profits are higher in the add-on pricing game. This intuition suggests that the $\pi^{\mathrm{a}} > \pi^{\mathrm{s,nd}}$ result may hold under fairly general conditions.

The assumption that high valuations for the add-on are associated with more intense horizontal preferences is clearly playing a central role in this argument. If I had instead assumed that high and low types were equally willing to switch between firms to take advantage of interfirm price differences, then the marginal customer pool would have the same type of distribution as the overall customer pool, and profits earned on the add-ons would be fully competed away.

Deneckere and McAfee's [1996] discussion of damaged goods price discrimination by monopolies emphasizes that the invention of a damaged good can provide a Pareto improvement: inventing and selling good L can increase the surplus of both low- and high-type consumers (as well as increasing the monopolist's profits). They mention that in other cases the more standard welfare trade-off occurs: price discrimination helps low-type consumers but hurts the high-type consumers. In the competitive situation considered here the outcome is different: the invention of good L makes both low- and high-type consumers worse off.

COROLLARY 2. Suppose that $\alpha_l/\alpha_h \in [3.2,10]$ and $w \in (\underline{w},\bar{w})$. Then,

(a) Both low- and high-type consumers are worse off in the discriminatory equilibrium of the standard pricing game than they would be if good L did not exist.

(b) Both low- and high-type consumers are worse off in the add-on pricing equilibrium than they would be if good L did not exist.

Proof. When good L does not exist, the model is the standard Hotelling model, and all consumers buy good H at $p_H^* = c + 1/\bar{\alpha}$.

(a) High types are obviously worse off in the discriminatory equilibrium because they pay more for the same good: $c + 1/\alpha_h > c + 1/\bar{\alpha}$. Low types pay $1/\bar{\alpha} - 1/\alpha_l$ less in the discriminatory equilibrium, but receive a good that is w/α_l less valuable. They are worse off because

$$\frac{w}{\alpha_l} > \frac{\underline{w}}{\alpha_l} = \frac{\alpha_l - \alpha_h}{\alpha_l^2} > \frac{\alpha_l - \alpha_h}{\alpha_l(\alpha_l + \alpha_h)} = \frac{2}{\alpha_l + \alpha_h} - \frac{1}{\alpha_l} = \frac{1}{\bar{\alpha}} - \frac{1}{\alpha_l}.$$

(b) High types are again worse off because they pay more. Low types pay $w/(\alpha_l + \alpha_h)$ less in the add-on pricing equilibrium, but get a good that is w/α_l less valuable and therefore are worse off. QED

Remarks

1. The view of add-on pricing that consumers should have in light of the equilibrium effects of add-on pricing is counter to what one often hears from consumer groups. For example, there was great popular uproar when, in the midst of the electricity crisis of 2001, some hotel chains started adding a fixed daily energy surcharge to every bill. Proposition 1 suggests that such a fee is irrelevant. High prices for minibar items and in-room movies seem to be regarded as less outrageous because consumers can avoid paying the high prices by not consuming the add-ons. The results of this section, however, indicate that it is precisely the voluntary nature of such fees that leads to lower consumer surplus.
2. A comparison of the discriminatory equilibrium of the standard pricing model and the add-on pricing equilibrium would reveal the standard welfare trade-off. Practicing add-on pricing constrains firms to charge more for the

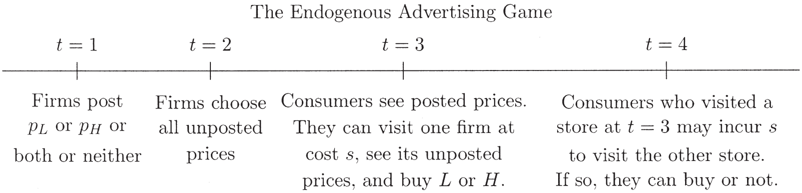

FIGURE II
Timeline for the Endogenous Advertising Game

add-on. Firms react by raising the price for good H and lowering the price for good L. This helps low types and hurts high types.

V. Why Do Firms Adopt Add-on Pricing?

The previous sections examined how the joint adoption of add-on pricing practices affects profits and consumer surplus. In this section I turn to the question of whether firms will adopt add-on pricing practices.

My first observation is that there is something to explain: in the simplest game one can write down with an endogenous choice of what to advertise, practicing add-on pricing is not individually rational.[14] Specifically, consider the "endogenous advertising game" in Figure II. It is a hybrid of the standard- and add-on pricing games in which firms post as many prices as they like at $t = 1$ and then choose the nonposted prices at $t = 2$. Recall that one intuition for the higher profits of the add-on pricing equilibrium is that firms benefit from a constraint that forces them to keep their low- and high-quality prices farther apart. The endogenous advertising game removes the constraint. Hence, if a firm's rival is playing as in the add-on pricing equilibrium, then the firm will have an incentive to advertise both prices and move them closer together.

PROPOSITION 4. Suppose that $\alpha_l/\alpha_h \in [3.2,10]$ and $w \in (\underline{w},\bar{w})$. Then, the endogenous advertising game does not have an equilibrium in which firms play as in the add-on pricing

14. This is different from what happens in Lal and Matutes [1994], where firms are indifferent to advertising one or two prices when there are no per-product advertising costs.

equilibrium; i.e., there is no equilibrium in which the firms advertise the price of good L at $t = 1$ and price the add-on at w/α_h at $t = 2$ whenever this is the best continuation equilibrium.

Proof. If both firms are playing as in the add-on pricing equilibrium, then they post $p_{iL} = p_L^* \equiv c + 1/\bar{\alpha} - w/2\bar{\alpha}$ at $t = 1$ and choose $p_{iH} = p_{iL} + w/\alpha_h$ at $t = 2$ whenever that is the best continuation equilibrium. Consider a deviation where firm 1 posts a slightly higher price for the low-quality good, $p_{1L} = p_L^* + \epsilon$ and also posts $p_{1H} = p_L^* + w/\alpha_h - \epsilon$. For a sufficiently small ϵ, the unique best-response for firm 2 at $t = 2$ is to price its upgrade at w/α_h. Hence, firm 1's high-quality price is ϵ less than firm 2's. The change in firm 1's profits from this deviation is approximated to first-order by $\Delta\pi_{1L} + \Delta\pi_{1H}$, with

$$\Delta\pi_{1L} = \epsilon \frac{\partial}{\partial p_{1L}} (p_{1L} - c)\left(\frac{1}{2} + \frac{p_{2L} - p_{1L}}{2}\alpha_l\right)\Bigg|_{p_{iL}=p_L^*}$$

$$\Delta\pi_{1H} = -\epsilon \frac{\partial}{\partial p_{1H}} (p_{1H} - c)\left(\frac{1}{2} + \frac{p_{2H} - p_{1H}}{2}\alpha_h\right)\Bigg|_{p_{iH}=p_L^*+w/\alpha_h}.$$

Simplifying and using $w > \underline{w} = (\alpha_l - \alpha_h)/\alpha_l$ gives

$$\Delta\pi_{1L} = \epsilon \frac{(\alpha_l + \alpha_h) - 2\alpha_l + w\alpha_l}{4\bar{\alpha}} > \epsilon \frac{\alpha_h - \alpha_l + \underline{w}\alpha_l}{4\bar{\alpha}} = 0$$

$$\Delta\pi_{1H} = -\epsilon \frac{-2\alpha_h + w\alpha_h + (\alpha_l + \alpha_h) - w(\alpha_l + \alpha_h)}{4\bar{\alpha}}$$

$$> \epsilon \frac{\alpha_h - \alpha_l + \underline{w}\alpha_l}{4\bar{\alpha}} = 0.$$

Hence, for a small enough ϵ the deviation is profitable. This shows that the profile is not an equilibrium. QED

Note that the proposition covers only the parameters for which I previously showed that the add-on pricing increases profits. It is precisely the fact that add-on pricing acts as a constraint on pricing that makes it not individually rational. In the situation considered in Proposition 1, in which whether firms practice add-on pricing is irrelevant, there are equilibria in which firms do and do not practice add-on pricing.

Proposition 4 shows that there is always a profitable deviation from the add-on pricing equilibrium. It is worth noting, however, that the deviation may not dramatically increase profits. The reason is that undercutting a nonposted price can be more difficult than undercutting a posted price. Consider, for example, the add-on pricing model with $\alpha_l/\alpha_h = 3$ and $w = 10/3$. The add-on pricing equilibrium has $p_L^* = c - 1/\alpha_l$ and $p_H^* = c + 3/\alpha_h$. If firm 2 was committed to these prices and firm 1 was capable of posting two prices at $t = 1$, its optimal deviation would be to dump all the unprofitable low types on the other firm and steal all of the high types by setting $p_{1L} \geq c$ and $p_{1H} = c + 2/\alpha_h$. In the add-on pricing model, however, this does not work. There is no continuation equilibrium in which firm 1 gets all the high types and $p_{2H} = p_{2L} + w/\alpha_h$ and because in that case firm 2 would be visited exclusively by low types and would therefore deviate at $t = 2$ to $p_{2H} = p_{2L} + w/\alpha_l$. It turns out that the only pure-strategy equilibrium of the continuation game is for firm 2 to set $p_{2H} = p_{2L} + w/\alpha_l$ at $t = 2$. At this price, firm 2 sells to all of the low types and all of the high types, and firm 1's large deviation ends up yielding it zero profits. In the proof of Proposition 4 I avoided this problem by considering only ϵ deviations. Somewhat larger deviations will yield even higher profits, but firm 1 does need to be sure to leave its rival with enough high types so that it remains an equilibrium for firm 2 to choose a high add-on price at $t = 2$. (In this example it must ensure that $q_{2H} \geq q_{2L}/2$.) This limits the gains to deviating.

How can one account for the use of add-on pricing strategies? My view is that this is a practically important question, but not one I should dwell on in this paper. There are a number of ways in which I could modify the endogenous advertising model to provide an explanation for why add-on pricing occurs without affecting my conclusions about the effects of add-on pricing. Some of the explanations are fairly standard and some are less so. In each case, however, the arguments seem sufficiently straightforward so that discussing them verbally in a paragraph or two probably conveys most of the insights one would get from a longer formal development. I will now briefly discuss four of these.

V.A. *Per-Product Advertising Costs*

The most obvious way to produce a model in which unobserved add-on pricing occurs endogenously is to simply assume that add-on prices are inherently very costly to advertise. This seems like a reasonable explanation for many applications. For example, if a particular hotel decided to make its lobby restaurant slightly less overpriced than lobby restaurants of the nearby hotels, it would be very difficult to credibly convey this information to potential guests.

In other applications advertising add-on prices would incur some incremental costs, but these are probably not large relative to the cost of advertising the base good. For example, it would not be too difficult for Best Buy to add a couple lines to the descriptions of computers in its sales circulars listing prices for add-ons like more memory or a higher capacity disk drive, nor for a hotel to advertise the cost of in-room Internet access along with its room rates. Lal and Matutes [1994] first pointed out that any incremental per-product advertising costs can provide a reason to advertise just one good and leave other prices unadvertised. A similar albeit weaker argument applies in my model. If the incremental cost of advertising the price of a second product is greater than the amount that a firm can gain by choosing a somewhat lower price for good H and a somewhat higher price for good L, then it will be individually rational for the firms to advertise just one price.

To argue that incremental advertising costs account for the use of unobservable add-on pricing in some market, it is not enough to argue that the incremental cost of advertising two prices instead of one is above the appropriate threshold. One would also need to argue that the firms cannot profitably deviate by posting a price for just good H instead of for just good L, e.g., advertising the price of the computer including the extra memory and higher capacity disk drive. To think about this, suppose that a firm only posted a price for good H at $t = 1$. It would then only sell good H in equilibrium, because at any price for good L that makes positive sales the firm will want to deviate and increase its good L price slightly given the search costs. The profits from selling good H to both populations are bounded above by the profits the firm receives when it chooses the price p_{1H} to maximize

$$\pi_1(p_{1H}) = \left(\frac{1}{2} + \frac{\alpha_l}{2}\left(c + \frac{1}{\bar{\alpha}} - \frac{w}{2\bar{\alpha}} + \frac{w}{\alpha_l} - p_{1H}\right)\right)(p_{1H} - c)$$
$$+ \left(\frac{1}{2} + \frac{\alpha_h}{2}\left(c + \frac{1}{\bar{\alpha}} - \frac{w}{2\bar{\alpha}} + \frac{w}{\alpha_h} - p_{1H}\right)\right)(p_{1H} - c).$$

This expression is maximized at $p_{1H} = c + 1/\bar{\alpha} + w/4\bar{\alpha}$, with the maximized value being $(1 + w/4)^2/\bar{\alpha}$. For the parameter values considered in Section IV, this is less than the equilibrium profit. Hence, incremental advertising costs can provide a complete explanation for why firms practice add-on pricing without the need to assume high-quality products are inherently more difficult to advertise.

The prices given in Proposition 3 are an equilibrium of the add-on pricing game for a larger set of parameter values than is covered by the hypotheses of the proposition. For some of these (e.g., when w is very large) the prices would fail to be an equilibrium of the endogenous advertising game because the firms would want to deviate and advertise good H instead.

V.B. Advertising Costs Determined by Consumer Search Patterns

A related explanation for why firms may choose not to advertise add-on prices is that in some industries firms sell so many products that advertising the cost of *any* product is prohibitively expensive. For example, it would have made no sense for Avis to conduct a nationwide media campaign to tell a few potential consumers that its rate for a three-day rental of a Pontiac Grand Am at the Detroit airport on August 2, 2002, was $74.97. What happens instead in industries like hotels and rental cars is that consumers learn about prices by looking for prices that firms have posted.

In such an environment firms can only practically inform consumers of prices that the consumers are looking for. Each of the main Internet travel Web sites, for example, is only designed to let consumers search for the base price for a rental, not for the price of a rental including insurance, prepaid gasoline, and other add-on charges. Looking only at low-quality prices can be perfectly rational for consumers if there is no dispersion in add-on prices. If most consumers only look for prices for good L, add-on pricing will be individually rational for firms. Cutting the price of good H lowers the firm's margin on all good H sales and does not attract consumers who only look for good L prices.

This obviously is a multiple equilibrium explanation. Practicing unobservable add-on pricing is an equilibrium, but there would also be an equilibrium in which firms post and consumers examine all prices.

V.C. Exploitation of Boundedly Rational Consumers

In some applications advertising costs do not seem like a plausible explanation for the use of unobservable add-on pricing. For example, it would not be very costly for Hertz to list insurance rates in its advertisements, nor for Best Buy to list prices for extended warranties in its sales circulars, nor for Capital One to make late payment fees more prominent in its mass mailings. The most promising explanation for why add-on prices are not advertised in these applications may be a behavioral one.

I mentioned earlier that the add-on pricing model can be given a behavioral interpretation: some or all of the high types could be unsophisticated consumers who are not as good at making price comparisons across firms and who are also easier to talk into buying add-ons at the point of sale. For example, they might be people who do not always compare prices from competing rental car companies before making a reservation and who also do not think in advance about the fact that they will be offered rental car insurance at the counter.

One reason why firms adopt add-on pricing policies may be that they somehow “trick” some unsophisticated consumers into paying more than they would if the firms advertised their add-on prices. For example, suppose that Hertz decided to augment its traditional advertisement of a \$97 weekly rental rate in Florida with a note saying that it was making full insurance available at a small discount off its current \$244 weekly rate. It seems plausible that this could reduce the profits Hertz earns on unsophisticated consumers via several mechanisms: some customers who make a bad decision to buy the insurance when under time pressure might make a better decision if they thought about it in advance; some customers might be spurred to gather information and learn that the insurance is largely unnecessary given the coverage they have through their regular auto insurance policy; and some customers might decide to make other plans and not rent a car when confronted with the total cost of a rental plus insurance in advance.

A simple modification to the endogenous advertising model that can make add-on pricing individually rational is to assume that a fraction of the high types are highly unsophisticated consumers who will buy good H if it is presented to them as add-ons are, but who will not buy good H if advertising informs them about its price in advance. This would make add-on pricing individually rational if the increase in profits from rational consumers when a firm moves its low- and high-quality prices closer together is more than offset by the decrease in profits that results from not tricking highly unsophisticated consumers. In light of the difficulty of undercutting a nonposted price noted above, such an argument might be made even if only a moderate fraction of consumers are irrational types.

V.D. Tacit Collusion

The main conclusion of Section IV was that the joint adoption of add-on pricing policies increases profits. This makes it possible to apply another standard explanation for why firms might practice add-on pricing: tacit collusion. To complete this story, one would want to explain why firms only collude on using add-on pricing rather than colluding directly on prices. Colluding on price would be more profitable, so this requires arguing that colluding on using add-on pricing is somehow easier than colluding on price. Colluding on the monopoly price can be difficult for many reasons: firms need to coordinate on changing prices in response to cost or demand shocks; different firms may prefer different prices; and monitoring deviations from optimal pricing may be difficult if the optimal pricing policy involves dynamically changing prices in response to privately known cost shocks and capacity constraints. A tacit agreement to use add-on pricing avoids all of the complexity, coordination, and monitoring issues: the firms just need to agree to and monitor that no one is advertising the price of good H.

To make this story more convincing, one would also want to argue not just that full collusion is impossible, but also that there are not easy strategies for colluding on prices that are less than fully collusive but still are more profitable than the equilibrium prices in the add-on pricing game. See Athey, Bagwell, and Sanchirico [2004] for a discussion of partially collusive pricing schemes in a model where firms have private information.

VI. The Cheapskate Externality

How do cheapskates affect markets? The question may be of current interest given that the Internet makes it much easier for cheapskates to find and exploit small price differences. The standard answer would be that cheapskates play an important role in keeping prices near cost. Frankel [1998], for example, proposes that the desire to live where budget-conscious consumers keep prices low may be one reason why wealthy and poor households are often found in close proximity in the United States. In this section I note that the traditional view of cheapskates is turned on its head in the add-on pricing model.

The model of this section is a slight variant of the previous add-on pricing model that I will refer to as the "cheapskate model." The only differences are that I assume that there is only an ϵ mass of cheapskates (rather than a unit mass) and that I will focus on what happens when α_l is much larger than α_h.

Proposition 5 contrasts the outcome of the cheapskate version of the add-on pricing game with what would happen if firms were selling a single good to the same population. Part (a) illustrates that the ordinary intuition about the effects of cheapskates on other consumers and on firms is borne out in a one-good model, which can be obtained as a special case of the cheapskate model by assuming that $w = 0$. Part (b) notes that the ordinary comparative statics are reversed in the cheapskate model when w is large enough to act as a constraint forcing firms to keep prices for good L and H apart.[15] One can thus think of add-on pricing as a clever competitive strategy that firms can use to turn the presence of cheapskates from a curse into a blessing. At the same time the presence of cheapskates reduces the utility of normal consumers.

The intuition for the contrast is that whereas firms in the one-good model are tempted to slightly undercut each other to attract cheapskates, firms in the add-on pricing model are tempted to slightly overcut each other. When w is large, firms are losing money on the cheapskates and would like to dump all of their cheapskate customers on the other firm. When w is not quite so large, the firms earn positive profits on the cheapskates. However, if they were to leave the high price unchanged and sell L at $c + 1/\alpha_h - w/\alpha_h$, they would be selling L for less than $c + 1/\alpha_l$ and hence would prefer to serve fewer cheapskates at a higher margin.

PROPOSITION 5. Suppose that $\alpha_l/\alpha_h > 2$. Define $\overline{\alpha_\epsilon} \equiv (\alpha_h + \epsilon\alpha_l)/(1 + \epsilon)$.

(a) In the one-good version of the cheapskate model obtained

15. As in Propositions 2 and 3 the requirement is that the upgrade price w/α_h be larger than what the difference between p_H and p_L would be if the firms competed separately for the low and high types.

by setting $w = 0$, for sufficiently small ϵ the unique symmetric equilibrium has $p^* = c + 1/\overline{\alpha_\epsilon}$, and prices and profits are decreasing in ϵ.
(b) If $w > \underline{w}$, then for sufficiently small ϵ the unique symmetric equilibrium of the cheapskate version of the add-on pricing model has

$$p_H^* = c + \frac{1}{\alpha_h} + \left(\frac{w}{\alpha_h} - \left(\frac{1}{\alpha_h} - \frac{1}{\alpha_l}\right)\right)\frac{\epsilon\alpha_l}{\alpha_h + \epsilon\alpha_l},$$

and profits and the price paid by high types are increasing in ϵ.

Proof. (a) In a neighborhood of any symmetric equilibrium price p^* firm 1's profits are

$$\pi_1(p_1) = \left(\frac{1+\epsilon}{2} + \frac{\alpha_h + \epsilon\alpha_l}{2}(p^* - p_1)\right)(p_1 - c).$$

The first-order condition for maximizing this implies that the only possible symmetric pure strategy equilibrium is $p^* = c + 1/\overline{\alpha_\epsilon}$. To verify that this is indeed an equilibrium, one must also check that firm 1 cannot profitably deviate to a higher price at which it serves no low types. The price that maximizes firm 1's profits from sales to high types is $p_1 = c + 1/2\overline{\alpha_\epsilon} + 1/2\alpha_h$. The profits from the high types at this price are $\alpha_h(1/\alpha_h + 1/\overline{\alpha_\epsilon})^2/8$. One can show that this is less than the equilibrium profit level for sufficiently small ϵ by evaluating the derivatives of this expression and the expression for the equilibrium profits with respect to ϵ at $\epsilon = 0$. Intuitively, if the firm abandons the low market, it gives up a potential profit that is first-order in ϵ, whereas the profits that a firm sacrifices in the high market when it also serves the low types are second-order in ϵ by the envelope theorem (because the price is approaching the optimal price in the high submarket).

The expression for the equilibrium price is clearly decreasing in ϵ. Equilibrium profits are given by $(1 + \epsilon)^2/(\alpha_h + \epsilon\alpha_l)$. Evaluating the derivative of this expression with respect to ϵ at $\epsilon = 0$ shows that profits are decreasing in ϵ in a neighborhood of $\epsilon = 0$ if $\alpha_l > 2\alpha_h$.

(b) Let p_L^* be the price set at $t = 1$ in a pure strategy equilibrium. When ϵ is small, both firms will set $p_{iH} = p_{iL} + w/\alpha_h$ at $t = 2$ whenever the first period prices are in some neighborhood of p_L^*. Hence, if firm 1 deviates to a price in a neighborhood of p_L^* its profits are given by

$$\pi_1(p_{1L}) = \left(\frac{1}{2} + \frac{\alpha_h}{2}(p_L^* - p_{1L})\right)\left(p_{1L} + \frac{w}{\alpha_h} - c\right)$$
$$+ \epsilon\left(\frac{1}{2} + \frac{\alpha_l}{2}(p_L^* - p_{1L})\right)(p_{1L} - c).$$

The fact that any equilibrium price p_L^* must be a solution to the first-order condition for maximizing this expression gives that the only possible equilibrium is to have p_L^* equal to w/α_h less than the expression given in the statement of the proposition. The expression for p_H^* is clearly increasing in ϵ. A first-order approximation to the profits when the firms charge prices p_L^* and p_H^* is

$$\pi^*(\epsilon) = \frac{1}{2\alpha_h} + \frac{1}{2}\left(\frac{\alpha_l}{\alpha_h}\left(\frac{w}{\alpha_h} - \left(\frac{1}{\alpha_h} - \frac{1}{\alpha_l}\right)\right) + \frac{1-w}{\alpha_h}\right)\epsilon + O(\epsilon^2).$$

The coefficient on ϵ in this expression is positive when $w = \underline{w} = \alpha_h(1/\alpha_h - 1/\alpha_l)$, and the coefficient is increasing in w. Hence, for all w satisfying the hypothesis of part (b), profits are increasing in ϵ when ϵ is small.

To complete the proof of part (b), it remains only to show that the prices derived above are an equilibrium and not just the solution to the first-order condition. Deviating to a higher price cannot be profitable. The concave profit function above applies as long as sales to the low types are nonzero. Hence firm 1's profits decline as it raises its price from p_L^* to $p_L^* + 1/\alpha_l$. Any price increases beyond that point would further decrease profits, since profits from sales to the high types are decreasing in p_{1L} at p_L^* and all higher prices. No deviation to a lower price will be profitable if firm 1 makes positive sales to the low types at the price which maximizes its profits on sales to the high types (by the concavity of the profit function). The difference between p_H^* and the price that maximizes profits from sales to the high types (setting $p_{1H} = \frac{1}{2}(p_H^* + c + 1/\alpha_h)$) is of order ϵ. Hence for ϵ small it is within $1/\alpha_l$ of the equilibrium price, and we can conclude that the profile is an equilibrium. QED

VII. Related Literature

This paper is related to several literatures. One is the literature on loss leaders in multigood settings. It focuses on the question of why firms set low prices for some goods and high prices for others. Holton [1957] is the seminal paper here. It notes that “The margin sacrificed on the loss leader is, of course, a promotion expense incurred to boost the sales of the other products of the store” and argues that high margins on the “other” products can be rationalized because “the supermarket enjoys a spatial monopoly on that item *once the consumer is in the store.*” Lal and Matutes [1994] formalizes Holton’s argument. It uses a Hotelling model of differentiation and models ex post monopoly power with a mechanism like that of Diamond [1971]. Verboven [1999] is another formalization in which the same thing—high prices for add-ons—happens for the same reason. Its model has both vertical and horizontal consumer heterogeneity, but the increased similarity to this paper is more apparent than substantive. Verboven does not consider the possibility of vertical tastes being correlated with the strength of horizontal preferences, which is the driving force behind my results. Gabaix and Laibson [2004] have developed a behavioral model of add-on pricing that proceeds very much along the lines of the model I sketched in subsection V.C. The trade-off that determines whether add-on pricing is individually rational in their model is similar to what I described in subsection V.C, albeit with one difference due to their assuming that firms engage in Bertrand competition—the loss from not tricking unsophisticated consumers must be larger than the improvement in efficiency that a firm could generate (and extract from the homogeneous rational consumers) by pricing add-ons at cost. Gabaix and Laibson do not address the impact of add-on pricing on profits: in their Bertrand model firms receive zero profits regardless of how they advertise.

This paper also belongs to the broader literature on competitive price discrimination.[16] Much of this literature examines third-degree discrimination, i.e., models in which firms can identify whether a consumer is a “high” or “low” type and charge different prices accordingly. Borenstein [1985] and Holmes [1989] provide some of the most basic results. Borenstein notes that differences across groups in the interfirm price sensitivity are needed to generate price discrimination. Holmes notes that banning price discrimination will lower prices in one market and raise them in the other; the net effect on profits is ambiguous. One paper with a result superficially similar to mine is Corts [1998]. It emphasizes that price discrimination can lead to reduced profits in all markets, but also shows that price discrimi-

16. Stole [2004] provides an excellent survey.

nation can lead to higher prices in all markets. The papers are not closely related, however. Corts' model is of third-degree discrimination and relies on strong asymmetries to generate the uniform price changes. Indeed, he shows that banning price discrimination always helps one group of consumers and hurts the other unless the groups or firms are sufficiently asymmetric so that one firm wants to price high to the first group and the other firm wants to price high to the second group.

The literature on competitive second-degree discrimination is smaller. Within this literature few papers analyze models with both vertical and horizontal differentiation, no doubt because it is difficult to construct models that are sufficiently tractable to allow closed-form solutions. (Borenstein [1985] presented his results using numerical simulations.) Two notable exceptions are Armstrong and Vickers [2001] and Rochet and Stole [2002]. Among other contributions, each of these papers derives a nondiscrimination theorem. They show that when brand preferences are of the type generally assumed in discrete-choice models and brand preferences are independent of consumers' valuations for quality, then the outcome of the competitive second-degree price discrimination model is that firms do not use quality levels to discriminate: all quality levels are offered at the same dollar markup over cost. As a contribution to this literature, my paper can be seen as providing a first analysis of a case that has been largely left aside since Borenstein's [1985] initial work: the case when willingness to pay for brand preferences is correlated with willingness to pay for higher quality. I think that this is an important case to study. The result on price discrimination being self-reinforcing is another contribution.

Another related literature is the literature on switching costs.[17] Although the early switching cost papers stressed applications where consumers buy the same product in multiple periods, many arguments are equally applicable to situations where the product purchased in the second period is different from the product bought in the first period. For example, Klemperer's [1987a] discussion of situations where profits with infinite switching costs are identical to profits with no switching costs is essentially the same as Lal and Matutes' irrelevance result, and a number of papers have used similar frameworks to discuss market power in aftermarket service, e.g., Shapiro [1995] and Boren-

17. Farrell and Klemperer [2004] provides an excellent survey.

stein, MacKie-Mason, and Netz [2000]. The most basic result in the switching cost literature is that switching costs can increase or decrease profits because they usually make first-period prices (think base goods) lower and second-period prices (add-ons) higher. The literature contains several well-known arguments about why switching costs may tend to raise profits, for example, Farrell and Shapiro [1988], Klemperer [1987b], and Beggs and Klemperer [1992]. The arguments in the above papers are inapplicable to add-ons, however, because they require an assumption that firms cannot differentiate between new and old customers; i.e., that the firm cannot choose an add-on price different from the price for good L. As a contribution to this literature, my paper can be seen as presenting a new argument for why switching costs may tend to raise profits in situations where firms can distinguish between old and new consumers. It also runs counter to some of the aftermarkets literature in that it provides an argument for why it might be advantageous to mandate that aftermarket service contracts be bundled with base goods.

There are other papers on loss-leaders that take very different approaches. Simester [1995] provides a signaling explanation for loss leaders in a model where retailers have heterogeneous costs. Lazear [1995] develops a monopoly model of bait-and-switch advertising. Hess and Gerstner [1987] develop a model in which firms sometimes stock out on advertised products and offer rain checks because consumers buy "impulse goods" whenever they visit a store to buy an advertised product.

This paper is also loosely related to all papers discussing a strategic investment that softens competition. Chapter 8 of Tirole [1988] reviews a number of such papers. A classic example is Thisse and Vives [1988], which notes that firms are better off competing in FOB prices than in delivered prices, because when they choose separate delivered prices for each location they end up being in Bertrand competition for the consumers at each location. As in this paper, they also note that FOB pricing is not individually rational in an extended game in which firms first choose pricing policies, and then compete in prices.

The one closely related empirical paper is Ellison and Ellison [2004], which analyzes demand and markups at a retailer using an add-on strategy when selling computer parts on the Internet. It provides evidence in support of this paper in two ways: it provides evidence that this paper's assumptions about demand reflect reality in at least one market; and it provides evidence in support of this paper's conclusions. The evidence relevant to the assumptions are estimates of how the demand for products of several quality levels depends on the prices of the other qualities. Specifically, loss leaders are shown to attract a large number of customers who end up buying upgraded products at higher price, and there is evidence of the adverse selection effect—the customer pool of attracted by a low-priced loss leader is shown to have a much higher percentage of customers who do not upgrade. Supporting evidence for the conclusion that add-on pricing softens competition comes from a straightforward analysis of price and cost data. The firm is estimated to earn average markups over marginal cost of about 10 to 15 percent even though the elasticity of demand for the base goods is between -25 and -40.

There is surprisingly little other empirical evidence on loss-leader pricing. The one standard empirical reference in marketing seems to be Walters [1988]. It examines the impact of loss leaders on store traffic by estimating a system of simultaneous equations. The key equation essentially regresses the total number of customers visiting a supermarket in a week on dummy variables for whether a product in each of eight categories is featured in a sales circular and offered at a discount of at least 15 percent. Walters finds little evidence that loss leaders affect store traffic. Chevalier, Kashyap, and Rossi [2003] use data from a Chicago supermarket chain to examine the pricing and demand for products that have large seasonal peaks in demand. Several findings are consistent with these products serving as loss leaders: the retail margin of a product tends to decline during the period of its peak demand even if this does not coincide with a peak in aggregate supermarket demand; aggregate margins do not decrease during aggregate demand peaks; reductions in item prices during product-specific demand peaks do not appear to be due to changes in demand elasticities; and reductions in item prices during product-specific demand peaks are associated with increases in product-specific advertising. Verboven [1999] uses a hedonic regression to compare markups for base model cars and cars with more powerful engines and finds that percentage markups on the premium engines are higher in some car classes but not in others.

VIII. Conclusion

The add-on pricing strategy described in this paper could be practiced in almost any business. Firms just need to be able to invent lower-quality versions of their products; the lower-quality products need not be any cheaper to produce. The key assumption about the consumer pool is that consumers who are more sensitive to interfirm price differences are less likely to purchase costly add-ons. This seems plausible given a number of sources of heterogeneity, e.g., rich versus poor consumers, individual versus business customers, or sophisticated versus unsophisticated shoppers. The key informational assumption needed for unobserved add-on pricing is that the add-on prices must be unobserved at the time when consumers compare base-good prices.

The idea of intentionally creating an adverse selection problem to limit competition seems robust. I imagine that my result about unadvertised add-on prices yielding higher profits than nondiscrimination could be substantially generalized. The idea could perhaps also be applied in other contexts.

For firms the main consequence of add-on pricing is that profits are higher than they otherwise would be given the degree of product differentiation. This effect may be generally important to our understanding of how firms maintain sufficient markups to survive in a world where fixed costs are often substantial. In the long run, of course, entry would be expected to reduce the degree of differentiation between adjacent firms and bring profits into line with fixed costs. What add-on pricing may help us understand is thus why we observe so many firms in various industries; e.g., how is it that National, Budget, and Thrifty rental cars exist?

I have not discussed social welfare extensively. Models like mine with unit demands are poorly suited to welfare analyses. For example, social welfare in the add-on pricing model is identical to that in the discriminatory equilibrium of the standard pricing model—in both models all low types buy one unit of L and all high types buy one unit of H. In a more realistic setup, the lower price for good L would increase consumption of L and the higher price for the add-on would reduce consumption of H. How the losses and gains would trade off is not clear.[18] The welfare comparison between the add-on pricing model and the one-good

18. See, for example, Klemperer [1987a] and Borenstein, MacKie-Mason, and Netz [2000].

model obtained by eliminating good L may be more straightforward. I noted that both the high and low types pay more relative to their valuation in the add-on pricing game than in the one-good model. If this is also true in a model with continuous aggregate demand functions, deadweight loss would presumably be unambiguously larger in the add-on model. (Welfare is unambiguously lower in the add-on pricing game with unit demands because it is inefficient for the low types to buy L rather than H.)

Appendix

Proof of Proposition 1. (a) Consider first the possibility of a symmetric pure strategy Nash equilibrium where all consumers buy good H at a price of p_H^*. This requires that $p_{iL} \geq p_H^* - w/\alpha_l$. If firm 1 deviates to a price p_{1H} in a neighborhood of p_H^* (and raises p_{1L} at the same time if need be), then firm 1's profits are

$$\pi_1(p_{1H}) = \left(1 + \frac{\alpha_l + \alpha_h}{2}(p_H^* - p_{1H})\right)(p_{1H} - c).$$

A necessary condition for Nash equilibrium is that the derivative of this expression be zero at $p_{1H} = p_H^*$. This gives $p_H^* = ½(c + 1/\bar{\alpha} + p_H^*)$, which implies that the only possible equilibrium of this form is $p_{1H} = p_{2H} = p_H^* = c + 1/\bar{\alpha}$.

To show that it is indeed a SPE for both firms to set $p_{iH} = c + 1/\bar{\alpha}$ and $p_{iL} \geq c + 1/\bar{\alpha} - w/\alpha_l$ (with all consumers buying good H from the closest firm) requires that we check that various possible deviations do not increase a firm's profits.

Consider first a deviation to prices p_{1L} and p_{1H} at which consumers only buy good H. To show that such a deviation cannot increase firm 1's profits, I will make a few observations in succession.

Observation 1: If firm 1 sells good H to some but not all consumers in each population, then the deviation does not increase profits.

To see this, note that in this case the formula above gives firm 1's profits. The expression is a quadratic in p_{1H}, and hence the solution to the first-order condition is the maximum.

Observation 2: If firm 1 sells good H to everyone in the cheapskate population, then the deviation does not increase profits.

With such prices, firm 1's profits are smaller than what one

gets from plugging p_{1H} into the profit formula above, which in turn is smaller than the profits from setting $p_{1H} = p_H^*$.
Observation 3: If firm 1 makes no sales in the cheapskate population, then the deviation is not profitable.

If firm 1 chooses $p_{1H} > p_H^* + 1/\alpha_l$, then it makes sales only to the high types, and its profits are

$$\pi_1(p_{1H}) = \left(\frac{1}{2} + \frac{\alpha_h}{2}(p_H^* - p_{1H})\right)(p_{1H} - c).$$

Taking the first-order condition, we see that the global maximum of this expression occurs at

$$p_{1H} = c + 1/2\alpha_h + 1/2\bar{\alpha}.$$

The firm would sell to low types at this price if

$$c + \frac{1}{2\alpha_h} + \frac{1}{2\bar{\alpha}} \leq c + \frac{1}{\bar{\alpha}} + \frac{1}{\alpha_l}.$$

A straightforward calculation shows that this is the case if $\alpha_l/\alpha_h \leq (3 + \sqrt{17})/2 \approx 3.562$, which is true given the assumption of the proposition. Hence, we can conclude that the profits from any price that sells only to the high types are at most equal to the profits received from the high types by setting $p_{1H} = c + 1/2\alpha_h + 1/2\bar{\alpha}$, which in turn is less than the profits received from setting this price and selling to members of both populations, which by observation 1 are less than what firm 1 receives by setting $p_{1H} = p_H^*$.

Taken together, observations 1–3 imply that any deviation which involves only selling good H is not profitable: if firm 1 deviates to $p_{1H} < p_H^*$, then firm 1 makes more sales to cheapskates than to high types so either observation 1 or observation 2 applies; if firm 1 deviates to $p_{1H} > p_H^*$, then firm 1 makes more sales to high types than to cheapskates and observation 1 or observation 3 applies.
Observation 4: Any deviation to prices p_{1L} and p_{1H} at which firm 1 sells only good L is not profitable.

To see this, note that firm 1 would sell at least as many units (and get a higher price on each at no higher cost) by setting prices $p'_{1L} = \infty$ and $p'_{1H} = p_{1L} + w/\alpha_l$. We have already shown that these prices do not increase firm 1's profit.

Finally, consider a deviation to prices p_{1L} and p_{1H} at which firm 1 sells good L to the cheapskates and good H to the high types. If there were no IC constraints so firm 1 could simply choose the optimal prices in each population, its choices would be $p_{1H} = c + 1/2\bar{\alpha} + 1/2\alpha_h$ and $p_{1L} = c + 1/2\bar{\alpha} + (1 - w)/2\alpha_l$. If $w < (\alpha_l - \alpha_h)/(2\alpha_l - \alpha_h)$, however, these prices would lead the high types to buy good L. If $w > (\alpha_l - \alpha_h)/\alpha_h$, these prices would lead the low types to buy good H. Accordingly, I will consider separately the optimal deviation of this form when w is small (with the high type's IC constraint binds), intermediate, and high (with the low type's IC constraint binding). I do this by presenting an additional series of observations.

Observation 5: If $w \leq (\alpha_l - \alpha_h)/(2\alpha_l - \alpha_h)$, then a deviation that sells L to the low types and H to the high types is not profitable.

In this case the constraint that $p_{1H} - p_{1L} \leq w/\alpha_h$ binds. Define $\pi_1(p_{1H},w)$ by

$$\pi_1(p_{1H},w) \equiv \left(\frac{1}{2} + \frac{\alpha_h}{2}\,(p_H^* - p_{1H})\right)(p_{1H} - c) + \left(\frac{1}{2} + \frac{\alpha_l}{2}\left(p_H^* - \frac{w}{\alpha_l} - \left(p_{1H} - \frac{w}{\alpha_h}\right)\right)\right)\left(p_{1H} - \frac{w}{\alpha_h} - c\right).$$

Let $\pi_1^d(w) = \max_{p_{1H}} \pi_1(p_{1H},w)$, and write p_{1H}^* for the price that maximizes this expression. The maximum profit achievable by a deviation of this form is at most $\pi_1^d(w)$ as long as the best possible deviation of this form has $p_{1H} - w/\alpha_h \geq c$. (In the opposite case the deviation cannot increase profits because firm 1 would be better off not selling good L and we have already seen that such deviations do not increase firm 1's profits.) From the envelope theorem we have

$$\frac{d\pi_1^d}{dw} = \frac{\partial \pi_1}{\partial w} = \frac{1}{2\alpha_h}\left((2\alpha_l - \alpha_h)(p_{1H}^*(w) - c) - \frac{2w(\alpha_l - \alpha_h)}{\alpha_h} - \frac{\alpha_l}{\bar{\alpha}} - 1\right).$$

To show that $\pi_1^d(w) < 1/\bar{\alpha}$ for all $w \in (0, (\alpha_l - \alpha_h)/(2\alpha_l - \alpha_h))$, it suffices to show that the derivative is negative for all w in the interval. For this it suffices to show that

$$(2\alpha_l - \alpha_h)(p_{1H}^*(w) - c) < 1 + \alpha_l/\bar{\alpha}.$$

If the high type's IC constraint were not binding, firm 1 would choose $p_{1H} = c + 1/2\bar{\alpha} + 1/2\alpha_h$. Given the constraint, the optimal $p^*_{1H}(w)$ will be smaller. Plugging this upper bound into the equation above gives that a deviation is not profitable if

$$\frac{1}{2}(2\alpha_l - \alpha_h)\left(\frac{\alpha_h + \bar{\alpha}}{\alpha_h \bar{\alpha}}\right) < \frac{\bar{\alpha} + \alpha_l}{\bar{\alpha}}.$$

Multiplying through and collecting terms, this is equivalent to

$$2\alpha_l^2 - \alpha_l\alpha_h - 5\alpha_h^2 < 0,$$

which holds provided that $\alpha_l/\alpha_h < (1 + \sqrt{41})/4 \approx 1.851$.
Observation 6: If $(\alpha_l - \alpha_h)/(2\alpha_l - \alpha_h) \leq w \leq (\alpha_l - \alpha_h)/\alpha_h$, then a deviation that sells L to the low types and H to the high types is not profitable.

In this case, the IC constraints are not binding, and the optimal deviation of this form is to $p_{1L} = c + 1/2\bar{\alpha} + (1 - w)/2\alpha_l$ and $p_{1H} = c + 1/2\bar{\alpha} + 1/2\alpha_h$. With these prices, profits from high type consumers are independent of w, and profits from low type consumers are decreasing in w. To see that the deviation is not profitable for any w in the interval, it therefore suffices to show that the deviation is not profitable when $w = (\alpha_l - \alpha_h)/(2\alpha_l - \alpha_h)$. This follows from observation 5.
Observation 7: If $(\alpha_l - \alpha_h)/\alpha_h \leq w$, then a deviation that sells L to the low types and H to the high types is not profitable.

In this case, the IC constraint of the low type is binding. The optimal deviation of this type has $p_{1L} = p_{1H} - w/\alpha_l$. This cannot increase firm 1's profits, because the type L consumers would also be willing to buy good H at price p_{1H}. Hence, firm 1 could do better selling only good H, and we have already seen that there is no profitable deviation of this form.

This concludes the argument to show that there are subgame perfect equilibria with $p_{2H} = p_{2H} = c + 1/\bar{\alpha}$, $p_{iL} > c + 1/\bar{\alpha} - w/\alpha_l$ and all consumers buying H from the closest firm at $t = 3$.

To prove the uniqueness claim of part (a), we must also show that there are no other symmetric pure strategy equilibria in the standard pricing game. It is obvious that there are no equilibria in which all consumers buy good L. A firm could increase its profits by setting $p'_{1L} = \infty$ and $p'_{iH} = \min(c, p_{iL} + w/\alpha_l)$. There are no equilibria where the low types buy good H and high types buy good L because the high types will strictly prefer to buy H whenever the low types weakly prefer H.

The final more serious possibility to consider is whether there is an equilibrium in which low types buy good L and high types buy good H. We can think of three possible cases: equilibria where low types and high types both strictly prefer to purchase the good they are purchasing, those where the high types are indifferent to buying good L, and those where the low types are indifferent to buying good H. The last of the three cases is not possible—each firm could increase its profits by not offering good L (because its low type consumers would buy H instead at the higher price). I will first discuss the first case.

In a discriminatory equilibrium where low types strictly prefer good L and high types strictly prefer good H the first-order conditions for each firm's profits imply that the only possible equilibrium is $p_{1L} = p_{2L} = c + 1/\alpha_l$ and $p_{1H} = p_{2H} = c + 1/\alpha_h$. Low types prefer good L at these prices only if $p_{iL} < p_{iH} - w/\alpha_l$. This requires $w \leq (\alpha_l - \alpha_h)/\alpha_h$. High types prefer good H at these prices only if $p_{iL} > p_{iH} - w/\alpha_h$. This requires $w \geq (\alpha_l - \alpha_h)/\alpha_l$. Assume that w does satisfy these conditions.

Suppose that firm 1 deviates to $p'_{1L} = \infty$ and $p'_{1H} = c + 1/\bar{\alpha} + w/4\bar{\alpha}$. One can verify that $p'_{1H} > p_{2H} - 1/\alpha_h$ and $p'_{1H} > p_{2L} + w/\alpha_l - 1/\alpha_l$ whenever $\alpha_l/\alpha_h < (3 + \sqrt{17})/2$. Hence, after the deviation, firm 1 sells to a subset of each population, and firm 1's profits are bounded below by the standard expression for profits in a competition-on-a-line model. Omitting much algebra, this gives that the profits from the deviation are at least

$$\left(\frac{1}{2} + \frac{\alpha_h}{2}(p_{2H} - p'_{1H})\right)(p'_{1H} - c) + \left(\frac{1}{2} + \frac{\alpha_l}{2}\left(p_{2L} - \left(p'_{1H} - \frac{w}{\alpha_l}\right)\right)\right)(p'_{1H} - c) = \left(1 + \frac{w}{4}\right)^2 \frac{1}{\bar{\alpha}}.$$

This is a profitable deviation from the hypothesized equilibrium profile if

$$\left(1 + \frac{w}{4}\right)^2 \frac{1}{\bar{\alpha}} > \frac{1}{2\alpha_l} + \frac{1}{2\alpha_h}.$$

Using the fact that $w \geq (\alpha_l - \alpha_h)/\alpha_l$, this shows that there is no equilibrium of this form if

$$\left(1 + \frac{\alpha_l - \alpha_h}{4\alpha_l}\right)^2 \frac{1}{\bar{\alpha}} > \frac{1}{2\alpha_l} + \frac{1}{2\alpha_h}.$$

Expanding the formula above, we can see that this is true if and only if

$$\left(\frac{\alpha_l}{\alpha_h} - 1\right)\left(4\left(\frac{\alpha_l}{\alpha_h}\right)^2 - 13\,\frac{\alpha_l}{\alpha_h} + 1\right) < 0.$$

This is true for

$$1 < \frac{\alpha_l}{\alpha_h} < \frac{13 + \sqrt{153}}{8} \approx 3.171.$$

The final analysis necessary to complete the proof of part (a) is a demonstration that there are also no discriminatory equilibria with $p_{iL} = p_{iH} - w/\alpha_h$ with the parameter restrictions of part (a). Firm 1 could deviate from such an equilibrium by raising or lowering p_{1L} and changing p_{1H} by exactly the same amount (i.e., setting $p_{1H} = p_{1L} + w/\alpha_h$). For a small enough change in prices, firm 1 would continue to sell L to a fraction of the low types and H to a fraction of the high types. Firm 1's profit would then be

$$\pi_1(p_{1L}) = \left(\frac{1}{2} + \frac{\alpha_l}{2}\,(p_{2L} - p_{1L})\right)(p_{1L} - c)$$
$$+ \left(\frac{1}{2} + \frac{\alpha_h}{2}\,(p_{2L} - p_{1L})\right)\left(p_{1L} + \frac{w}{\alpha_h} - c\right).$$

Considering the first-order condition for maximizing this expression, we can see that the only possible SPE of this form would have $p_{1L} = c + 1/\bar{\alpha} - w/2\bar{\alpha}$ (and $p_{1H} = c + 1/\bar{\alpha} - w/2\bar{\alpha} + w/\alpha_h$). Given the restriction on α_l/α_h in the proposition, it turns out that there is always a profitable deviation from this profile.

If $w > (\alpha_l - \alpha_h)/\alpha_l$, a profitable deviation is to raise p_{1L} by a small amount and leave p_{1H} unchanged. With such a deviation profits from sales to the high types will be unchanged, and firm 1 will sell fewer units of good L to low types (at a higher price). This is profitable if the derivative with respect to p_{1L} of

$$\left(\frac{1}{2} - \frac{\alpha_l}{2}\,(p_{1L} - p_{2L})\right)(p_{1L} - c)$$

is positive when evaluated at $p_{1L} = p_{2L} = c + 1/\bar{\alpha} - w/2\bar{\alpha}$. The derivative is

$$\frac{1}{2} - \frac{\alpha_l}{2}\left(\frac{1}{\bar{\alpha}} - \frac{w}{2\bar{\alpha}}\right),$$

which is positive for $w > (\alpha_l - \alpha_h)/\alpha_l$.

When $w \leq (\alpha_l - \alpha_h)/\alpha_l$, a profitable deviation is to simply raise p_{1L} sufficiently high so that the low types will prefer to buy good H. Firm 1 will sell fewer units with this strategy, but at a higher price. Profits from the high types are unchanged. Profits from sales to the low types change from ½ $(1/\bar{\alpha} - w/2\bar{\alpha})$ to

$$\left(\frac{1}{2} - \frac{\alpha_l}{2}\left(\frac{w}{\alpha_h} - \frac{w}{\alpha_l}\right)\right)\left(\frac{1}{\bar{\alpha}} - \frac{w}{2\bar{\alpha}} + \frac{w}{\alpha_h}\right).$$

The change in profits simplifies to

$$\frac{w}{2}\left(\frac{1}{\alpha_h} - \frac{\alpha_l - \alpha_h}{\alpha_h}\frac{2\alpha_h + w\alpha_l}{\alpha_h(\alpha_l + \alpha_h)}\right).$$

Substituting in the upper bound $(\alpha_l - \alpha_h)/\alpha_l$ for the second w in this expression and simplifying, we find that the change in profits is at least

$$\frac{w}{2}\frac{2\alpha_h - \alpha_l}{\alpha_h^2},$$

which is positive for $\alpha_l/\alpha_h < 2$. This completes the proof that there is no equilibrium in which the firms make sales of good L and thereby completes the proof of part (a) of the proposition.

(b) To analyze the add-on pricing game, I begin with a lemma noting that if the firms' first-period prices are close together, then at $t = 2$ the firms will sell the "upgrade" to all consumers at a price of w/α_l.

LEMMA 1. Assume that $\alpha_l/\alpha_h \leq 1.6$. Suppose that at $t = 1$ the firms choose prices p_{1L} and p_{2L} with $|p_{2L} - p_{1L}| \leq (2\alpha_h - \alpha_l)/\alpha_h^2$ and $c < p_{iL} < (v - w - s - 1/2)/\alpha_l$. Then, the unique equilibrium of the subgame at $t = 2$ has the firms selling the upgrade to all consumers at a price of w/α_l.

A proof of the lemma is presented immediately after the proof of this proposition. Given the result of the lemma, we know that firm 1's profit following a small deviation at $t = 1$ from the symmetric profile $p_{1L} = p_{2L} = p_L^*$ results in its earning a profit of

$$\pi_1(p_{1L}) = \left(\frac{1}{2} + \frac{\alpha_l}{2}(p_{2L} - p_{1L})\right)\left(p_{1L} + \frac{w}{\alpha_l} - c\right) + \left(\frac{1}{2} + \frac{\alpha_h}{2}(p_{2L} - p_{1L})\right)\left(p_{1L} + \frac{w}{\alpha_l} - c\right).$$

Considering the first-order condition for maximizing this expression shows that the only possible first-period price in a symmetric SPE is $p_L^* = c + 1/\bar{\alpha} - w/\alpha_l$. By Lemma 1, at $t = 2$ both firms must set $p_{iH} = c + 1/\bar{\alpha} - w/\alpha_l + w/\alpha_l = c + 1/\bar{\alpha}$ on the equilibrium path, and all consumers must buy good H from the nearest firm. This completes the proof of the uniqueness part of part (b) of the proposition.

To verify that there is indeed a pure strategy SPE of the form described, suppose that both firms set $p_{iL} = c + 1/\bar{\alpha} - w/\alpha_l$ at $t = 1$ and follow some SPE strategy at $t = 2$ and that consumers behave optimally given the firms' equilibrium strategies and purchase good H if they are indifferent between buying H and L.

By definition we know that firm 1 has no profitable deviation at $t = 2$.

To show that there is no profitable deviation at $t = 1$, I will present a series of observations covering various cases.

Observation 1: Firm 1 cannot increase its profits by deviating to any p_{1L} with $|p_{1L} - p_L^*| < (2\alpha_h - \alpha_l)/\alpha_h^2$.

With such a deviation, Lemma 1 implies that firm 2 sets $p_{2H} = c + 1/\bar{\alpha}$ at $t = 2$. Part (a) of the proposition implies that no matter what prices p_{1L} and p_{1H} firm 1 chooses it cannot earn a profit in excess of $1/\bar{\alpha}$ when $p_{2H} = c + 1/\bar{\alpha}$. This includes the prices firm 1 is charging after a deviation here.

Observation 2: Firm 1 cannot increase its profits by deviating to any p_{1L} with $p_{1L} \leq p_L^* - (2\alpha_h - \alpha_l)/\alpha_h^2$.

In this case, regardless of what prices are chosen at $t = 2$, firm 1 will sell at least as many units of good L as of good H. Hence, its profits are bounded above by the profits from selling the same number of units at a price of $p_{1L} + w/\alpha_l$. If $p_{1L} + w/\alpha_l < 0$, then these profits are negative and not a profitable deviation. If $p_{1L} + w/\alpha_l > 0$, then profits are bounded above by the profits firm 1 would receive from selling to all consumers at this price. Given the assumed upper bound on p_{1L}, the gain from the deviation is

$$\pi_1(p_{1L}) - \frac{1}{\bar{\alpha}} \leq 2\left(\frac{1}{\bar{\alpha}} - \frac{2\alpha_h - \alpha_l}{\alpha_h^2}\right) - \frac{1}{\bar{\alpha}}$$

$$= \frac{2}{\alpha_l + \alpha_h} - 2\,\frac{2\alpha_h - \alpha_l}{\alpha_h^2} = \frac{2}{\alpha_l + \alpha_h}\left(\left(\frac{\alpha_l}{\alpha_h}\right)^2 - \frac{\alpha_l}{\alpha_h} - 1\right).$$

This is negative when $\alpha_l/\alpha_h < (1 + \sqrt{5})/2$.

Observation 3: Firm 1 cannot increase its profits by deviating to any p_{1L} with $p_{1L} \geq p_L^* + (2\alpha_h - \alpha_l)/\alpha_h^2$.

In this case, firm 2 will make at least as many sales to low types as to high types. Hence, $p_{2H} = p_{2L} + w/\alpha_l = c + 1/\bar{\alpha}$. Again, part (a) of the proposition implies that the prices p_{1L} and p_{1H} firm 1 ends up charging cannot increase its profits. QED

Proof of Lemma 1. To see that $p_{1U} = p_{2U} = w/\alpha_l$ is an equilibrium, note that when the firms are expected to set the same upgrade price, the mass of group j customers visiting firm 1 is $\frac{1}{2} + (\alpha_j/2)(p_{2L} - p_{1L})$. Profits are

$$\pi_1\left(\frac{w}{\alpha_l}, \frac{w}{\alpha_l}\right) = \sum_{j=1}^{2} \left(\frac{1}{2} + \frac{\alpha_j}{2}(p_{2L} - p_{1L})\right)\left(p_{1L} - c + \frac{w}{\alpha_l}\right).$$

Deviating to a lower upgrade price obviously cannot increase firm 1's profits—the lower price will not lead to any extra sales.

If firm 1 deviates to charge a higher price, no low types will purchase the upgrade. This decreases profits by $(1/2 + \alpha_l(p_{2L} - p_{1L})/2)w/\alpha_l$. Firm 1's sales to high types will be no higher. The upgrade price paid by these customers can be at most w/α_h. Hence the increase in profits on sales to high types is at most $(1/2 + (\alpha_h/2)(p_{2L} - p_{1L}))(w/\alpha_h - w/\alpha_l)$. The change in firm 1's profits from the deviation is thus bounded above by

$$\left(\frac{1}{2} + \frac{\alpha_h}{2}(p_{2L} - p_{1L})\right)\left(\frac{w}{\alpha_h} - \frac{w}{\alpha_l}\right) - \left(\frac{1}{2} + \frac{\alpha_l}{2}(p_{2L} - p_{1L})\right)\frac{w}{\alpha_l}$$

$$= \frac{w}{2}\left[\left(\frac{1}{\alpha_h} - \frac{2}{\alpha_l}\right) + (p_{2L} - p_{1L})\left(\frac{\alpha_h}{\alpha_h} - \frac{\alpha_h}{\alpha_l} - \frac{\alpha_l}{\alpha_h}\right)\right]$$

$$\leq \frac{w}{2\alpha_h\alpha_l}[\alpha_l - 2\alpha_h - (p_{2L} - p_{1L})\alpha_h^2].$$

The bound on $|p_{2L} - p_{1L}|$ assumed in the lemma ensures that this is negative.

I now show that this is the only equilibrium.

First, note that the upper bound on the prices for L ensures that all consumers will visit one of the firms in equilibrium.

Next, note that in any equilibrium all firms choose p_{iU} equal to either w/α_h or $w\alpha_l$. To see this, one first shows that both firms must set $p_{iU} \geq w/\alpha_l$. Otherwise, the firm with the lower price attracts a positive mass of consumers. All of these consumers receive weakly higher ex ante expected utility from visiting that firm. Once they have sunk s visiting that firm, they strictly prefer to buy there at the equilibrium prices. If the firm raises its upgrade price by some amount less than s/α_l and keeps its price less than w/α_l, it will lose no sales. This would be a profitable deviation. The fact that $p_{iU} \geq w/\alpha_l$ implies that consumers in the low group get no surplus from buying the upgrade. Because of this and because the difference in prices for L is assumed to be bounded above by $(2\alpha_h - \alpha_l)/\alpha_h^2$, which is less than $1/\alpha_l$, each firm attracts a positive mass of consumers in any equilibrium. There cannot be an equilibrium with $w/\alpha_l < p_{iU} < w/\alpha_h$ because firm i would gain by raising its price slightly (if it is making any sales of good H) or by dropping its price to w/α_l (if not). There cannot be an equilibrium with $p_{iU} > w/\alpha_h$ because firm i will sell no units of H, but would make positive sales by dropping its price to w/α_l.

There cannot be an equilibrium with $p_{1U} = p_{2U} = w/\alpha_h$ because then the mass of customers from each group visiting firm 1 is exactly the same as when $p_{1U} = p_{2U} = w/\alpha_l$. The calculation above thus implies that firm 1 would increase its profits by deviating to $p_{1U} = w/\alpha_l$. To see that there cannot be an equilibrium with $p_{1U} = w/\alpha_h$ and $p_{2U} = w/\alpha_l$, note that in this case the mass of low-type consumers visiting firm 1 would be exactly the same as in the above calculations, but that firm 1 would be visited by fewer high types. This makes the gain from deviating to $p_{1U} = w/\alpha_l$ even greater. QED

Proof of Proposition 2. The result that $\bar{w} > \underline{w}$ follows from simple algebra:

$$\begin{aligned}\bar{w} > \underline{w} &\Leftrightarrow \frac{4\bar{\alpha}}{\sqrt{\alpha_l \alpha_h}} - 4 > \frac{\alpha_l - \alpha_h}{\alpha_l}\\ &\Leftrightarrow 4(\alpha_l + \alpha_h)^2\alpha_l^2 > \alpha_l\alpha_h(5\alpha_l - \alpha_h)^2\\ &\Leftrightarrow \alpha_l(\alpha_l - \alpha_h)(4\alpha_l^2 - 13\alpha_l\alpha_h + \alpha_h) > 0.\end{aligned}$$

This inequality is satisfied whenever $\alpha_l/\alpha_h > (13 + \sqrt{153})/8 \approx 3.17$.

Another fact that will come in handy is that $\bar{w} < (\alpha_l - \alpha_h)/\alpha_h$. To see this, one can carry out a calculation similar to that above to show that

$$\frac{\alpha_l - \alpha_h}{\alpha_h} > \bar{w} \Leftrightarrow \alpha_h(\alpha_l - \alpha_h)(\alpha_l^2 + 3\alpha_l\alpha_h + 4\alpha_h) > 0.$$

(a) To show that the strategy profile where both firms set $p_{iL} = p_L^* \equiv c + 1/\alpha_l$ and $p_{iH} = p_H^* \equiv c + 1/\alpha_h$ is a sequential equilibrium (when combined with optimal behavior on the part of consumers), note first that the restrictions on w imply that when consumers anticipate that $p_{iL} = p_L^*$ and $p_{iH} = p_H^*$ then all consumers will visit the closest firm: low types will buy good L, and high types will buy good H. (This follows from $\alpha_h(p_{iH} - p_{iL}) = \underline{w} < w$ and $\alpha_l(p_{iH} - p_{iL}) = (\alpha_l - \alpha_h)/\alpha_h > \bar{w} > w$.) Hence, if the firms follow the given strategy profile, each earns a profit of $1/2\alpha_l + 1/2\alpha_h$.

If firm 1 deviates to any prices p_{1L} and p_{1H} at which it sells L to low types and H to high types and sells to some but not all of the customers in each market, then its profits are

$$\pi_1(p_{1L},p_{1H}) = \left(\frac{1}{2} + \frac{\alpha_l}{2}(p_L^* - p_{1L})\right)(p_{1L} - c) + \left(\frac{1}{2} + \frac{\alpha_h}{2}(p_H^* - p_{1H})\right)(p_{1H} - c).$$

This is a concave function uniquely maximized at $p_{1L} = \frac{1}{2}(c + p_L^* + 1/\alpha_l) = c + 1/\alpha_l$ and $p_{1H} = c + 1/\alpha_h$, so the deviation does not increase firm 1's profits.

If firm 1 sells L to low types and H to high types and sells to no or all customers in one (or both) markets, then it is strictly worse off: zero sales earn zero rather than positive profits; and when selling to all customers of type j firm 1's profits from sales to type j consumers are no greater than the profits it would have earned from setting the price $p_{1j} = p_j^* - 1/\alpha_j$, and profits at this price are lower than the equilibrium profits because they are given by the formula above.

There is no profitable deviation which involves selling H to low types and L to high types because the high types will strictly prefer buying H whenever the low types are willing to buy H.

It is not necessary to check separately whether there is a profitable deviation involving selling only good L. If firm 1 has a profitable deviation which involved selling L at a price of p_{1L} to a subset of the consumers, then it also has an even better profitable deviation in which it sells H at a price of $p_{1L} + w/\alpha_l - \epsilon$ to the same set of consumers.

To show that the profile given in (a) is an equilibrium, it therefore remains only to show that there is no profitable deviation involving selling H to both populations. When firm 1 sells H

to at least some of the consumers in each population at a price $p_{1H} > c$, its profits are bounded above by

$$\pi_1(p_{1H}) = \left(\frac{1}{2} + \frac{\alpha_h}{2}(p_H^* - p_{1H})\right)(p_{1H} - c) + \left(\frac{1}{2} + \frac{\alpha_l}{2}\left(p_L^* - \left(p_{1H} - \frac{w}{\alpha_l}\right)\right)\right)(p_{1H} - c).$$

(The expression is only an upper lower bound and not necessarily the actual profit level because the quantity sold in each market is at most one.) This is a quadratic that is maximized at the unique solution to the first-order condition. Differentiating this expression, we find after some algebra that it is maximized for

$$p_{1H} = c + 1/\bar{\alpha} + w/4\bar{\alpha}.$$

Substituting into the profit function, the value at the maximum is $(1 + w/4)^2/\bar{\alpha}$. This is no greater than the equilibrium profit if

$$\left(1 + \frac{w}{4}\right)^2 \frac{1}{\bar{\alpha}} \leq \frac{1}{2\alpha_l} + \frac{1}{2\alpha_h}.$$

This is satisfied for

$$w \leq 4\left(\frac{\bar{\alpha}}{\sqrt{\alpha_l \alpha_h}} - 1\right),$$

which is the assumption in the statement of the proposition that $w < \bar{w}$. This concludes the proof that the discriminatory profile described in part (a) of the proposition gives a sequential equilibrium.

(b) To see that the standard pricing game sometimes has an equilibrium in which all consumers buy H at a price of $c + 1/\bar{\alpha}$, note first that we showed in the proof of Proposition 1 that these prices satisfy the first-order condition for profit maximization. This profile will be an equilibrium if firm 1 cannot gain either by selling good H to the high types and nothing to the low types or by selling H to the high types and L to the low types.

In the proof of Proposition 1, I noted that there is no profitable deviation involving only sales to the high types when $\alpha_l/\alpha_h < (3 + \sqrt{17})/2$ because at the price that maximizes profits from sales to the high types, the firm will sell to some low types as well. When α_l/α_h is larger, firm 1's profit function does have a local maximum at $p_{1H} = c + 1/2\bar{\alpha} + 1/2\alpha_h$. Firm 1's profit when it sets this price and sells to only high types is $\alpha_h(1/2\bar{\alpha} + 1/2\alpha_h)^2/2$. This is larger than $1/\bar{\alpha}$ only if $\alpha_l/\alpha_h > 5 + \sqrt{32} \approx 10.66$. Hence, for the parameter values of the proposition, this deviation is not profitable.

In the proof of Proposition 1, the optimal deviation involving selling both H and L could take any of three forms. Given the restriction on w in Proposition 2, only the second of these (corresponding to observation 6 in the earlier proof) arises, and the optimal deviation of this form is $p_{1L} = c + 1/2\bar{\alpha} + (1 - w)/2\alpha_l$ and $p_{1H} = c + 1/2\bar{\alpha} + 1/2\alpha_h$. The profit from this deviation is

$$\frac{\alpha_h}{2}\left(\frac{1}{2\bar{\alpha}} + \frac{1}{2\alpha_h}\right)^2 + \frac{\alpha_l}{2}\left(\frac{1}{2\bar{\alpha}} + \frac{1-w}{2\alpha_l}\right)^2.$$

A numerical calculation shows that this deviation is never profitable if $\alpha_l/\alpha_h > 6.4$. This is also true when α_l/α_h is smaller if w is closer to $\bar{w}$. In these cases, the specified profile is therefore also an equilibrium.

There can be no other symmetric pure strategy equilibria in which the firms sell good H to everyone because $p_H = c + 1/\bar{\alpha}$ is the unique solution to the first-order condition that arises in this case. There can be no equilibrium where the firms sell L to the high types and H to the low types for the standard sorting reasons. The only remaining possibility for another symmetric pure strategy equilibrium is that there might be an equilibrium where the firms sell H to the high types and L to the low types, but at a price different from those given in part (a) of the proposition.

There can be no such equilibrium with both types strictly preferring to buy the good they are buying because then the first-order conditions for each firm not wanting to raise or lower each price (used in the existence argument) imply that the equilibrium must have $p_{iL} = c + 1/\alpha_l$ and $p_{iH} = c + 1/\alpha_h$. There can be no such equilibrium in which the low types are indifferent to buying H because in that case firm 1 would profit from lowering the price of the upgrade by ϵ and selling it to the low types as well. This leaves only the possibility of an equilibrium in which the high types are buying H and are indifferent to buying L instead. To see that this does not work, note (as in the proof of Proposition 1) that considering the first-order condition for firm 1 deviating and raising or lowering both p_{1L} and p_{1H} by exactly the same amount shows that the only possible equilibrium of this form would be to have $p_{1L} = c + 1/\bar{\alpha} - w/2\bar{\alpha}$ and $p_{1H} = c + 1/\bar{\alpha} -$

$w/2\bar{\alpha} + w/\alpha_h$. At these prices, firm 1 could deviate and raise p_{1L} slightly. This would not affect firm 1's sales to high types. In the low market firm 1's profits (in a neighborhood above $c + 1/\bar{\alpha} - w/2\bar{\alpha}$) are

$$\left(\frac{1}{2} + \frac{\alpha_l}{2}\left(c + \frac{1}{\bar{\alpha}} - \frac{w}{2\bar{\alpha}} - p_{1L}\right)\right)(p_{1L} - c).$$

The derivative of this expression with respect to p_{1L} evaluated at $c + 1/\bar{\alpha} - w/2\bar{\alpha}$ is

$$\frac{1}{2}\left(1 - \alpha_l\left(\frac{1}{\bar{\alpha}} - \frac{w}{2\bar{\alpha}}\right)\right).$$

This is positive if $w > \underline{w}$. Hence, there is no equilibrium of this form. QED

Proof of Proposition 3. (a) Suppose that in a sequential equilibrium both firms set $p_{iL} = p_L^*$ at $t = 1$. The first thing to note is that at $t = 2$ the optimal continuation equilibrium for the firms involves the add-on being now sold for a price of w/α_h (both in equilibrium and following small deviations).

Claim: If $|p_{1L} - p_L^*| < 1/\alpha_h$ and $p_{2L} = p_L^*$, then there is a sequential equilibrium in which both firms choose $p_{iU} = w/\alpha_h$ at $t = 2$. This is the best equilibrium for the firms.

To see this, note again that because of the structure of the consumer search problem the only possible equilibrium upgrade prices will be w/α_l and w/α_h. If both firms set $p_{iU} = w/\alpha_h$, then at $t = 2$ the firm that chose a lower price at $t = 1$ will be visited by at least half of the low types and by at most all of the low types. Hence, at least one-third of the consumers visiting the low priced firm are high types and the assumption of the proposition that $w/\alpha_h > 3w/\alpha_l$ ensures that this firm is better off selling to just the high types. The firm that set the higher price at $t = 1$ will be visited by more high types than low types and is thus also better choosing the high upgrade price.

If firm 1 deviates from the equilibrium and chooses a price p_{1L} with $|p_{1L} - p_L^*| < 1/\alpha_l$ and the firm-optimal continuation equilibrium is played at $t = 2$, then firm 1's profits are

$$\pi_1(p_{1L}) = \left(\frac{1}{2} + \frac{\alpha_l}{2}(p_L^* - p_{1L})\right)(p_{1L} - c) + \left(\frac{1}{2} + \frac{\alpha_h}{2}(p_L^* - p_{1L})\right)\left(p_{1L} + \frac{w}{\alpha_h} - c\right).$$

This is a quadratic maximized at the solution to the first-order condition. The derivative is

$$\frac{d\pi_1}{dp_{1L}} = 1 - 2\bar{\alpha}p_{1L} + \bar{\alpha}p_L^* + \bar{\alpha}c - \frac{w}{2}.$$

Setting $p_{1L} = p_L^*$ and solving, we see that the only possible symmetric equilibrium of this form is $p_L^* = c + 1/\bar{\alpha} - w/2\bar{\alpha}$. This completes the proof of the uniqueness claim of the proposition.

The calculation above also implies that no deviation from this profile with $|p_{1L} - p_L^*| < 1/\alpha_l$ will increase firm 1's profits. To complete the proof that this is indeed an equilibrium, one needs to verify that larger deviations (for which the expression above is not the correct profit function) also do not increase firm 1's profits.

To see that no deviation to a price $p_{1L} > p_L^* + 1/\alpha_l$ can increase firm 1's profits, note that for prices in this range firm 1's profits (if they are nonzero) are given by

$$\pi_1(p_{1L}) = \left(\frac{1}{2} + \frac{\alpha_h}{2}(p_L^* - p_{1L})\right)\left(p_{1L} + \frac{w}{\alpha_h} - c\right).$$

The derivative of this expression is

$$\frac{d\pi_1}{dp_{1L}} = \frac{1}{2} - \alpha_h p_{1L} + \frac{\alpha_h}{2}p_L^* + \frac{\alpha_h}{2}c - \frac{w}{2}.$$

The derivative is decreasing in p_{1L}, and after some algebra one can show that it is negative when evaluated at $p_L^* + 1/\alpha_l$ when $w \geq \underline{w}$. Hence, profits from any deviation in this form are less than the profits from a deviation to $p_{1L} = c + 1/\alpha_l$, which are less than the putative equilibrium profit by the above argument. (Apart from the algebra the result in this case should also be obvious: firms are keeping p_{1L} and p_{1H} farther apart than is optimal. It would make no sense to increase the already too-high price in market H and abandon market L.)

To see that there is no profitable deviation with $p_{1L} < p_{2L} - 1/\alpha_h$, note that with such a price firm 1 sells to all of the low and high type consumers. (There cannot be an equilibrium where firm 2 attracts some high types by charging a low upgrade price because firm 2 will attract no low types and hence would always raise its upgrade price by s once consumers visit it.) Its profits are bounded above by $(p_L^* - 1/\alpha_h - c) + (p_L^* - 1/\alpha_h + w/\alpha_h - c)$. This is less than the equilibrium profit of $p_L^* + w/2\alpha_h - c$ if

$$p_L^* - c < \frac{2}{\alpha_h} - \frac{w}{2\alpha_h} \Leftrightarrow \frac{2-w}{2\bar{\alpha}} < \frac{4-w}{2\alpha_h} \Leftrightarrow w < \frac{4\alpha_l}{\alpha_l - \alpha_h}.$$

The restrictions that $w < \bar{w}$ and $\alpha_l/\alpha_h < 10$ imply that the left-hand side is less than four. The right-hand side is always greater than four, so the deviation is never profitable.

Finally, to see that there is no profitable deviation with $p_{1L} \in (p_L^* - 1/\alpha_h, p_L^* - 1/\alpha_l)$, note that firm 1's profits with such a price are

$$\pi_1(p_{1L}) = (p_{1L} - c) + \left(\frac{1}{2} + \frac{\alpha_h}{2}(p_L^* - p_{1L})\right)\left(p_{1L} + \frac{w}{\alpha_h} - c\right).$$

The profits from such a deviation cannot be profitable if this expression does not have a local maximum in the interval because we have already seen that deviations to either endpoint of the interval are not profitable. The solution to the first-order condition for maximizing the expression above is

$$p_{1L} = c + \frac{3}{2\alpha_h} + \frac{1}{2\bar{\alpha}} - \frac{w}{4\bar{\alpha}} - \frac{w}{2\alpha_h}.$$

This fails to be interior if

$$c + \frac{3}{2\alpha_h} + \frac{1}{2\bar{\alpha}} - \frac{w}{4\bar{\alpha}} - \frac{w}{2\alpha_h} > c + \frac{1}{\bar{\alpha}} - \frac{w}{2\bar{\alpha}} - \frac{1}{\alpha_l}.$$

After some algebra one can see that this is the case whenever

$$3 + 3\frac{\alpha_h}{\alpha_l} + 2\frac{\alpha_h^2}{\alpha_l^2} > w,$$

which is true for all $w < \bar{w}$ as long as $\alpha_l/\alpha_h < 10$ because the left-hand side is at least 3.32 and the right-hand side is at most $4(5.5/\sqrt{10} - 1) \approx 2.96$. Hence, the deviation cannot be profitable. (The assumption of the proposition that $\alpha_l/\alpha_h < 10$ could be weakened by computing the profits at the interior optimum when it exists and showing that they remain below the equilibrium profit level for a broader range of parameter values.)

Part (b) of the proposition is proved in the text. QED

MASSACHUSETTS INSTITUTE OF TECHNOLOGY AND NATIONAL BUREAU OF ECONOMIC RESEARCH

References

Armstrong, Mark, and John Vickers, "Competitive Price Discrimination," *RAND Journal of Economics,* XXXII (2001), 579–605.

Athey, Susan, Kyle Bagwell, and Chris Sanchirico, "Collusion and Price Rigidity," *Review of Economic Studies,* LXXI (2004), 317–349.

Beggs, Alan, and Paul Klemperer, "Multi-Period Competition with Switching Costs," *Econometrica,* LX (1992), 651–666.

Benoit, Jean-Pierre, and Vijay Krishna, "Finitely Repeated Games," *Econometrica,* LIII (1985), 905–922.

Borenstein, Severin, "Price Discrimination in Free-Entry Markets," *RAND Journal of Economics,* XVI (1985), 380–397.

Borenstein, Severin, Jeffrey MacKie-Mason, and Janet Netz, "Exercising Market Power in Proprietary Aftermarkets," *Journal of Economics & Management Strategy,* IX (2000), 157–188.

Chevalier, Judith, Anil Kashyap, and Peter Rossi, "Why Don't Prices Rise During Periods of Peak Demand? Evidence from Scanner Data," *American Economic Review,* XCIII (2003), 15–37.

Corts, Kenneth, "Third-Degree Price Discrimination in Oligopoly: All-out Competition and Strategic Commitment," *RAND Journal of Economics,* XXIX (1998), 306–323.

Deneckere, Raymond J., and R. Preston McAfee, "Damaged Goods," *Journal of Economics and Management Strategy,* V (1996), 149–174.

Diamond, Peter, "A Model of Price Adjustment," *Journal of Economic Theory,* III (1971), 156–168.

Ellison, Glenn, and Sara F. Ellison, "Search, Obfuscation, and Price Elasticities on the Internet," NBER Working Paper No. 10570, 2004.

Farrell, Joseph, and Carl Shapiro, "Dynamic Competition with Switching Costs," *RAND Journal of Economics,* XIX (1988), 123–137.

Farrell, Joseph, and Paul Klemperer, "Coordination and Lock-In: Competition with Switching Costs and Network Effects," in Mark Armstrong and Robert Porter, eds. *Handbook of Industrial Organization, Volume III* (Amsterdam: North-Holland, 2004), forthcoming.

Frankel, David, "A Pecuniary Reason for Income Smoothing," *Journal of Urban Economics,* 44 (1998), 158–169.

Gabaix, Xavier, and David Laibson, "Shrouded Attributes and Information Suppression in Competitive Markets," mimeo, Massachusetts Institute of Technology and Harvard University, 2004.

Hausman, Jerry A., and J. Gregory Sidak, "Why Do the Poor and the Less-Educated Pay More for Long-Distance Calls?" *Contributions to Economic Analysis & Policy,* III (2004), 3.

Hess, James D., and Eitan Gerstner, "Loss Leader Pricing and Rain Check Policy," *Marketing Science,* VI (1987), 358–374.

Holmes, Thomas J., "The Effects of Third-Degree Price Discrimination in Oligopoly," *American Economic Review,* LXXIX (1989), 244–250.

Holton, Richard H., "Price Discrimination at Retail: The Supermarket Case," *Journal of Industrial Economics,* VI (1957), 13–32.

Klemperer, Paul, "Markets with Consumer Switching Costs," *Quarterly Journal of Economics,* CII (1987a), 375–394.

——, "The Competitiveness of Markets with Switching Costs," *RAND Journal of Economics,* XVIII (1987b), 138–150.

Lal, Rajiv, and Carmen Matutes, "Price Competition in Multimarket Duopolies," *RAND Journal of Economics,* XX (1989), 516–537.

Lal, Rajiv, and Carmen Matutes, "Retail Pricing and Advertising Strategies," *Journal of Business,* LXVII (1994), 345–370.

Lazear, Edward P., "Bait and Switch," *Journal of Political Economy,* CIII (1995), 813–830.

Pratt, John, David A. Wise, and Richard J. Zeckhauser, "Price Differences in Almost Competitive Markets," *Quarterly Journal of Economics,* XCIII (1979), 189–211.

Rochet, Jean-Charles, and Lars Stole, "Nonlinear Pricing with Random Participation," *Review of Economic Studies,* LXIX (2002), 277–311.

Shapiro, Carl, “Aftermarkets and Consumer Welfare: Making Sense of Kodak,” *Antitrust Law Journal,* LXIII (1995), 483–511.
Simester, Duncan, “Signalling Price Image Using Advertised Prices,” *Marketing Science,* XIV (1995), 166–188.
Stole, Lars, “Price Discrimination and Imperfect Competition,” in Mark Armstrong and Robert Porter, eds. *Handbook of Industrial Organization, Volume III.* (Amsterdam: North-Holland, 2004), forthcoming.
Thisse, Jacques-Francois, and Xavier Vives, “On the Strategic Choice of Spatial Pricing Policy,” *American Economic Review,* LXXVIII (1988), 122–137.
Tirole, Jean, *The Theory of Industrial Organization* (Cambridge, MA: MIT Press, 1988).
Verboven, Frank, “Product Line Rivalry and Market Segmentation—With an Application to Automobile Optional Engine Pricing,” *Journal of Industrial Economics,* XLVII (1999), 399–425.
Walters, Rochney, “Retail Promotions and Retail Store Performance: A Test of Some Key Hypotheses,” *Journal of Retailing,* LXIV (1988), 153–180.