Utilizing free samples in the prescription drug market

Kevin E. Pflum*

March 7, 2010

Abstract

We develop a theoretical model of pharmaceutical sampling to provide insight into the strategic use and welfare effects of free drug samples. Consumers have a heterogeneous degree of illness and consumer-to-drug matches are idiosyncratic. A monopolist dispenses samples to mitigate the consumers' risk, and increase profit. We find that the firm's sampling strategy varies with the state of information. When the efficacy of the drug is commonly known the firm's sampling decision is not monotonic in the efficacy. However, when the efficacy is not commonly known, the sampling decision acts as a signal of the efficacy and is monotonic. Moreover, because of the signaling effect, the firm dispenses samples for lower efficacies than is optimal when information is symmetric. The effect on consumer surplus primarily depends on the welfare gains of consumers in the uncovered portion of the market who receive treatment from samples.

JEL Codes: D82, D83, L12, L15

Key Words: product sample, idiosyncratic matching, prescription drugs

1 Introduction

In 1919 Mahon Kline of Glaxo Smith Kline pioneered a new form of pharmaceutical marketing by mailing drug samples directly to physicians across the United States. Kline started a practice which, in 2006, has expanded into a $2.6 billion marketing expense for pharmaceutical firms (Band, 2007).[1] The pharmaceutical industry is not alone in dispensing free samples. Free or trial samples can often be found with beauty products, baby products, pet food, magazines and subscription television services. Goods in these markets represent nondurable *experience goods* (Nelson, 1974) characterized by frequent repurchase. Idiosyncratic tastes make finding the most suitable product match to any particular consumer difficult for consumer and firm alike. However, samples provide a fail-safe means of facilitating consumer learning and product matching.

The provision of free samples are generally welcomed by physicians and consumers and pharmaceutical firms often point out that samples help indigent patients get needed medication. However, the practice is not without controversy; for example, the University of Michigan Health System

*I am grateful to Dan Levin, Paul J. Healy, and Lixin Ye for their advice, guidance and comments on this research project. I would also like to thank Paula Cordero Salas, Michael Sinkey and the participants of the Ohio State applied micro workshop for helpful discussions and comments.

[1]The value reflects the production cost of the samples and not the cost of delivery by representatives, on whom the industry spent $6.7 billion in 2006 (IMS Health, 2008). The retail value of the samples provided exceeds $16 billion and is the value often cited in the popular press.

has completely banned the practice and the University of Pennsylvania and Stanford University medical schools prohibit staff members from accepting samples (Rabin, 2007). Detractors argue that the practice allows pharmaceutical firms to influence a physician's prescribing behavior[2] and believe that there exist better alternatives to helping indigent patients such as voucher programs (Brennan et al., 2006; Greenland, 2009). Detractors further support the claim sampling is used to influence physicians by pointing out that samples are often targeted at more junior physicians, who may be more susceptible to the influence (Rabin, 2007) and, that sampling generally increases the number of prescriptions written for the sampled drug (Gönül et al., 2001; Mizik and Jacobson, 2004). However, as Joseph and Mantrala (2009) point out, these concerns are inconsistent with empirical findings. For example, the intensity of sampling varies substantially across therapeutic categories, countering the notion samples are used to generate goodwill (Backer et al., 2000) as there is no apparent reason samples may generate goodwill in some therapeutic classes but not others.

The objective of the current paper is to develop a simple model of sample dispensation to provide insights into the circumstances leading to and consequences of providing free samples. Specifically, the goals of the paper are to identify what properties of the market and the drug lead to sampling, how the provision of free samples alters the price a firm may set for their drug, and, lastly, to identify the welfare effects of sampling. As Arrow (1963) identifies in his seminal article on medical care, information asymmetries and uncertainty are critical features of health care markets to which its institutions have had to adapt. As consumers face considerable uncertainty concerning their idiosyncratic match with a drug, the premise of this paper is that sample dispensation evolved as a way to mitigate their risk, allowing the firm to charge a higher price and increase its profit. We model consumers as being heterogeneous in their degree of illness, and consequently, in their potential degree of benefit from consuming a drug. The idiosyncrasy of the patient-to-drug match is modeled as a Bernoulli match probability where the efficacy of a drug is defined as the parameter of the Bernoulli distribution. In this way, a drug is either a match or not with a consumer and the match outcome cannot be determined ex ante, even when the drug's efficacy is common knowledge. The firm has the standard objective of maximizing profits and does so by choosing its price and sampling strategy based on the market conditions and consumer beliefs.

Because of the idiosyncratic nature of the drug we contrast the use of samples when there is symmetric information to their use when information regarding the drug's efficacy is asymmetric. When the efficacy of the drug is commonly known, the firm will provide samples to facilitate consumers in learning their ex post match value. The firm's profit curve absent sampling, referred to as the *no-sampling* profit curve, is increasing and convex in the drug's efficacy because efficacy affects profit in two ways. First, a higher efficacy results in a higher number of matches and, thus, a higher volume of repeat purchases; and second, the price the firm can charge is increasing in the efficacy. In contrast, the firm's *sampling* profit curve is linear in efficacy because the firm is able to charge the full information monopoly price for any efficacy and the increase in profit from higher efficacies results from the increase in returning, matched consumers only. Because the no-sampling profit curve is increasing and convex and the sampling profit curve is linear, when the two cross, they cross twice and the sampling strategy is not monotonic in efficacy; i.e., the firm will optimally provide samples for drugs having a moderate efficacy but not for drugs having a sufficiently high

[2]Another marketing practice common in pharmaceutical markets that may be used to influence a physician's prescribing behavior is the giving of small gifts, such as toys or pens, with the names of branded drugs as well as meals or conferences in idyllic locations. Konrad (2002) develops a model of promotional competition where gifting acts as a soft bribe entering the physician's utility function.

or low efficacy.

The firm's sampling strategy changes when information is asymmetric because the consumers' uncertainty is now in two dimensions: uncertainty in the drug's efficacy, and uncertainty in their idiosyncratic match. We show that the firm is unable to signal its efficacy through the price it sets; thus, consumers form a prior based on their knowledge of the efficacy distribution and the firm's sampling choice. Because consumers use the firm's sampling choice in forming their prior, firms may have to provide samples to signal their drug has a high efficacy, even when the firm would optimally not provide samples for a drug having a sufficiently high efficacy under symmetric information. The signaling effect of sample dispensation alters the firm's strategy in two ways. First, it causes the firm to provide samples for drugs having a lower efficacy than it would under symmetric information. Second, the sampling strategy under asymmetric information is monotone in the drug's efficacy as the firm will provide samples for all efficacies above some threshold. We fully characterize the market equilibrium, which may take the form of a pooling or a separating equilibrium depending on the values of the model's parameters.

Finally we characterize the welfare effects of sampling. We show that consumer welfare increases under sampling under one of two circumstances. First, consumer welfare increases whenever the firm's revenue decreases by permitting sampling. This can only occur when information is asymmetric because under symmetric information, the firm's revenue necessarily must increase if it optimally provides samples. Second, consumer welfare can increase when the firm's revenue increases from sampling if there are consumers in the uncovered portion of the market who receive treatment from samples, and their welfare gain exceeds the increase in the firm's profit. Otherwise sampling results in a net transfer of surplus from consumers to the firm. It follows that overall social welfare increases only if there are consumers in the uncovered portion of the market who receive treatment from samples, and their welfare gain exceeds the firm's cost of providing samples.

Ours is the first game-theoretic model of product sampling of which we are aware. However, the model can be thought of as complimentary to the analysis done by Joseph and Mantrala (2009) who look at how physicians utilize samples under diagnostic uncertainty. In their model physicians are able to diagnosis a patient and determine which of two drugs may be a better match before prescribing one; however, because of noisiness in the diagnosis, the physician is unable to determine with certainty which is the more appropriate drug. As a result of the physician's uncertainty and as a means of lowering the overall cost of treatment to the patient, the physician will provide the patient a sample and determine if the provided drug was a match before prescribing the drug. Joseph and Mantrala focus on only the physicians' strategy and not on the firms' strategy, although they show samples give a pharmaceutical firm the ability to capture more of the market than they otherwise could. Choices of prices and the dispensation of samples are both taken as given. The purpose of the current article is, in part, to identify the strategic relationship between the firm's choice of price and sampling decision.

Although there has been surprisingly little work in the theoretical literature specifically on product sampling, the introductory pricing and quality signaling literature has some similarities to the asymmetric information regime of the current article. Seminal articles include Shapiro (1983), Bagwell (1987), Liebeskind and Rumelt (1989) and Bagwell and Riordan (1992). A key distinction between the introductory pricing models and the current model is that the probability of a match is systematic in the introductory pricing models and idiosyncratic in the present model. As shown by Proposition 1, idiosyncrasy in being a match is the *raison d'être* for sampling. In the introductory pricing models the uncertainty is over a vertical quality characteristic and prices serve as a credible signal of the quality level because a mismatch between price and quality will

result in a loss in demand as consumers learn their valuation for the good. That is, everybody is a match, but the value of the match varies from one consumer to the next. In the current model consumers are uncertain that they will be a match, and hence uncertain that the drug will have any value. Proposition 2 shows that the firm cannot use the price to signal the the efficacy of the drug. Although the value of a drug, if it is a match, is constant, the current model could be extended to allow for a vertical quality dimension, in which case the price could be used to signal the ex post quality level as in the introductory pricing literature. The idiosyncratic match probability is the defining characteristic of pharmaceuticals which separate them from most other experience good markets.

Bergemann and Välimäki (2006) consider dynamic pricing of a new idiosyncratic experience good by a monopolist which has some relation to the current model. They utilize a continuous-time model in which the information regarding the consumer's value for the good arrives according to a Poisson process. Absent a signal, consumers balance the informational gain of continuing to purchase with the potential for current losses. The principal difference between the model in Bergemann and Välimäki and the current model is with respect to the nature of the uncertainty. Because, in Bergemann and Välimäki, the information regarding one's match arrives according to a Poisson process, every consumer is essentially a match with the drug if they wait sufficiently long, but they must wait to find out their true ex post match value with the drug. As more of the market becomes informed with the passage of time, the firm will set a price which optimally balances extracting rents from the informed and keeping the uninformed in the market long enough to learn their ex post match value. The consumer's decision to remain in or exit the market is based on the expected match value and the probability of receiving the signal. In the current model, the focus is on the idiosyncrasy of being a match and not on the idiosyncrasy of the ex post match value. Not all consumers will be a match with a particular drug and they learn relatively quickly if they are not a match. Uncertainty in the match probability in the current model is analogous to having uncertainty in the Poisson parameter, λ, of Bergemann and Välimäki.[3] Ideally the firm would like to perfectly price discriminate by setting a distinct price for each consumer based upon the expected match before the consumer learns if he is a match, and then set a different price based on the consumer's ex post valuation for the drug once the consumer learns they are a match. However, like in Bergemann and Välimäki (2006) we assume that the firm can only offer a spot price; thus, does not have the ability to individually price discriminate.

Before introducing the model, it should be noted that there are two unique aspects of the pharmaceutical market which we are omitting: (i) the presence of insurance; and, (ii) the agency of the physician. Insurance is omitted for simplification, but also because nearly half of the amount spent on prescription drugs in the United States is out of pocket compared to 20% for other healthcare expenditures.[4] Consequently, as with any good, the price of prescription drugs plays an important role in consumers' decision to purchase. The presence of a physician adds complexity because of her role as gate-keeper to prescription drugs and agent to the patient. Physicians make the prescription decision, but do not directly experience the benefit, or any negative effects of the drug. Moreover, once the decision to prescribe is made by a physician, the patient still makes the final decision of whether or not to purchase and consume the drug. To simplify the exposition, we assume a perfect agency between physician and patient. Consequently the consumers in the

[3]The relationship between the match probability of the current model and the Poisson parameter in Bergemann and Välimäki (2006) is not perfect, but both have the same affect on the consumer's option value of consumption today.

[4]U.S. Department of Health and Human Services, Medical Expenditure Panel Survey, 2008

current model embody both the role of physician and patient. We acknowledge that agency issues within the physician-patient relationship may be important to fully characterize the marketplace, but because the analysis is focused on the strategic use of samples to facilitate learning, augment beliefs, increase profits, and/or induce treatment, regardless of to which party these effects can be attributed, we feel the simplification for expositional clarity is acceptable.

The remainder of the article is organized as follows. In Section 2 we describe the model. In Section 3 the firm's pricing strategy given its sampling choice is derived and the relationship between third degree price discrimination and sampling is discussed. In Section 4 the equilibria are solved for both the symmetric and asymmetric information regimes. Section 5 examines the firm's strategy when a portion of the market has symmetric information and knows the drug's efficacy whereas the remaining portion of the market does not know the drug's efficacy. In Section 6 we examine the welfare implications of sampling. Finally, in Section 7 we conclude with some remarks and suggestions for further research.

2 The Model

Consider a market where every period a continuum of mass 1 new consumers enter with the desire to purchase one unit of a particular drug per period for treatment of an illness. The drug is produced by only one firm and there are no suitable substitutes. The efficacy $\phi \in [0,1]$ of the drug is defined as a Bernoulli-distributed match probability,[5] which the firm is unable to select. We will analyze the firm's problem when the efficacy is common knowledge and when it is only known by the firm. When the efficacy is not commonly known, the consumers' uncertainty for the efficacy is represented by the distribution G having strictly positive density g, which is known by the firm. The marginal cost of production is independent of the efficacy level of the drug, so is normalized to 0. Dispensing samples has a total cost of c_S per period.[6,7] The firm's profit, which is dependent on its sampling strategy, is defined in the next section where the optimal prices given the firm's sampling strategy are derived.

Consumers are indexed by $i \in [0,1]$. Consumer i's degree of illness is measured by $\theta_i \in [\underline{\theta}, \overline{\theta}]$ and is known by the consumer, but not the firm. However the firm knows the distribution of severities, F, and its properties. This mirrors the observation that pharmaceutical firms leave samples with physicians, so do not observe the illness severity of any given patient, but through extensive market research are well aware of the distribution of the illness within the population. The distribution F is assumed to be continuously differentiable and has strictly positive density f over the supports. To further insure that the profit maximization problem is well behaved, F is assumed to satisfy the monotone hazard rate property $d\{F(\theta)/f(\theta)\}/d\theta > 0$ for all θ in the support.

[5] The efficacy of a drug is predominantly learned by the firm and the Federal Drug Administration (FDA) during the trials phase of development. For the FDA to approve a drug, among other factors, it must be demonstrably more effective than a placebo, therefore very low efficacy drugs are prevented from entering the market. However, for expositional simplicity, we ignore this restriction by allowing the realized efficacy to be as low as 0.

[6] If the total cost includes a fixed cost f_S and a variable cost v_S, then, when x consumers are given a sample the total cost of samples is $c_S = f_S + v_S x$. However, because a firm does not know any particular consumers' type it will choose to either provide all new consumers a sample or no consumer a sample. When it gives all new consumers a sample $x = 1$ and the distinction between total cost and variable costs of sampling is not important.

[7] The cost of providing samples includes the costs of the actual samples, the special packaging and instructions that accompany the samples, and the number of excess samples that physicians may distribute to indigent patients or take home for personal use (Westfall et al., 1997).

A consumer learns if he is a match to the drug after the first period of treatment.[8] If the consumer is a match with the drug, then he continues the treatment by purchasing the drug for a total of T treatment periods.[9] At the conclusion of the treatment the consumer is cured and exits the market. Because every treatment period a unit mass of new consumers enter and a unit mass of consumers exit, the size of the market remains constant. Moreover, T generations of consumers are in the market in any given treatment period. The presence of overlapping generations precludes the idea that all consumers in the market are in the same period of treatment and learning the efficacy of the drug simultaneously; consequently, the proportion of consumers who know their match value remains constant across time. It is further assumed that the drug has been on the market for at least T periods in order to keep the mass of consumers constant.[10]

The consumer-to-drug match value $m_i \in \{0, 1\}$ is 0 in the event of a non-match—the same value the consumer receives from not consuming the drug—and 1 in the event of a match. Consumers are risk-neutral[11] with per period, indirect utility specified as

$$U_{i,t} = \theta_i m_i - p_t \text{ for all } t \in \{1, \ldots, T\},$$

where $p_t \geq 0$ is the price set by the firm.

Consumers are perfectly rational so the expected value of treatment contains an option value of all future expected consumption. Consumers and the firm have the common discount factor $\delta \in [0, 1]$.[12] When consumer i knows the efficacy of the drug then i's total discounted expected utility of trying the drug is given as

$$EU_i = (\phi\theta_i - p_1) + \phi \sum_{t=2}^{T} \delta^{t-1}(\theta_i - p_t). \tag{1}$$

The first term of the consumer's expected utility identifies the expected utility from the first period of treatment. The benefit is stochastic while the cost, p_1, is not. The second term on the right-hand-side of (1) is the option value of future consumption. The consumer will still purchase the drug for the first period of treatment when the first term is negative as long as the option value is sufficient to ensure the total discounted expected utility is positive.

When ϕ is not commonly known, consumer i's total discounted expected utility is given as

$$EU_i = \int_0^1 \left((\phi\theta_i - p_1) + \phi \sum_{t=2}^{T} \delta^{t-1}(\theta_i - p_t)\right) dG(\phi). \tag{2}$$

In this way the consumers' expected utility from trying the drug has two dimensions of uncertainty. First there is uncertainty regarding the overall efficacy of the drug reflected by the distribution G, and second, there is uncertainty regarding whether or not the drug will be a match with any given individual reflected by the Bernoulli parameter ϕ.

[8] A period of treatment is thus defined as the length of time required to learn if a drug is working (e.g., a few days, a week, or a month) and not defined as a single dose or a prescription period such as 30 or 90 days.

[9] The number of treatment periods need not be deterministic. For example, after consuming the drug for one unit of time, it may be the case that the consumer needs to continue the treatment with probability ρ, which is independent of the drug's efficacy or characteristics but is known to the firm. When the consumer continues the treatment with probability ρ, then $T \equiv \frac{1}{1-\rho}$ is the expected total number of treatment periods.

[10] This assumption has no affect on the firm's pricing strategy, but greatly simplifies the analysis.

[11] Adding risk-aversion should not change the qualitative results of the model. Moreover, Chintagunta et al. (2008) provide evidence physicians and patients are risk-neutral in their choice of drug.

[12] Assigning different discount factors to the consumers and firm will not change the results. The affect the discount factor has on the firm's strategy is discussed in Section 3.

3 Pricing

In this section we first examine the optimal pricing strategy of the firm given it cannot engage in any kind of price discrimination, thus offers a single spot price. Next we explore the relationship between third degree price discrimination and sampling and derive the optimal price with sampling.

3.1 Prices without Sampling

Because the firm cannot individually price discriminate and the new consumers have a common prior, the spot price will remain constant from one period to the next ($p_t = p$ for all t), and the firm's problem is simply that of choosing the spot price which maximizes its profit given the beliefs of the uninformed and the distribution of illness severity.

Let μ represent the consumers' common prior for the drug's efficacy and define $\Sigma \equiv \sum_{t=0}^{T-1} \delta^t = (1-\delta^T)/(1-\delta)$. Plugging μ and Σ into (2) and rearranging allows consumer i's total discounted expected utility of consumption to be more simply expressed as

$$EU_i = \mu\Sigma(\theta_i - p) - (1-\mu)p. \tag{3}$$

The first term identifies the discounted present value of utility given the consumer is a match times the expected probability of being a match and the second term identifies the expected loss from trying the drug.

An uninformed consumer, i, will purchase the drug if his expected utility from purchasing the drug is at least as high as his utility of not purchasing, $EU_i \geq 0$. Rearranging (3) allows the participation constraint to be expressed by consumer type as

$$\theta_i \geq \hat{\theta}(p \mid \mu) \equiv p\left(\frac{1+\mu(\Sigma-1)}{\mu\Sigma}\right). \tag{4}$$

Thus, $\hat{\theta}(p \mid \mu)$ identifies the type for which the expected benefit of trying the drug, $\mu\Sigma\hat{\theta}$, just equals the expected cost, $p(1+\mu(\Sigma-1))$. Notably, the cutoff type is dependent on the belief, μ, so any action the firm takes (or does not take) which changes the consumers' belief will also change the cutoff type, ceteris paribus.

The firm's per period demand consists of the demand from new consumers together with the proportion of consumers who are a match with the drug and are continuing the treatment:

$$D(p) = \big(1+\phi(T-1)\big)\big[1-F\big(\hat{\theta}(p \mid \mu)\big)\big]. \tag{5}$$

Although the firm has the same discount factor as the consumers, the discount factor does not affect the firm's sampling decision as it affects the consumers' purchase decision. The firm is assumed to have an infinite horizon while any individual consumer that enters the market will remain for at most T periods (and proportion $1-\phi$ will exit after just one when they find that they are not a match). Both the consumer and the firm will make the choice which maximizes their current discounted utility or profit. However, because the firm has an infinite horizon, it is equivalent for the firm to maximize its average profit,

$$(1-\delta)\sum_{t=0}^{\infty}\delta^t\Pi_t = \Pi_t.$$

Thus the firm's pricing and sampling strategies are independent of its discount factor.

Using (5), the firm's average, expected profit given it does not provide samples is

$$E\Pi^{NS}(p;\phi) = p\big(1+\phi(T-1)\big)\big[1-F\big(\hat{\theta}(p \mid \mu)\big)\big]. \tag{6}$$

The firm's objective is to choose p to maximize (6). Because F satisfies the monotone hazard rate property, (6) is concave in p,[13] therefore, from the first-order condition, the optimal price p^*_{NS} is the unique solution to the equation

$$p^*_{NS}(\mu) = \left(\frac{1-F\big(\hat{\theta}(p^*_{NS} \mid \mu)\big)}{f\big(\hat{\theta}(p^*_{NS} \mid \mu)\big)}\right) \Big/ \hat{\theta}'(p^*_{NS} \mid \mu)\ . \tag{7}$$

The optimal price has the intuitive characteristics of increasing in the drug's actual or expected efficacy, the expected number of periods of treatment and the discount factor:

$$\frac{\partial p^*_{NS}}{\partial \mu} \geq 0,\ \frac{\partial p^*_{NS}}{\partial T} \geq 0, \quad \text{and} \quad \frac{\partial p^*_{NS}}{\partial \delta} \geq 0.$$

The comparative statics are a consequence of the option value of consumption. That is, the higher the option value of consumption the higher the price the firm can charge and the expected efficacy, number of periods of treatment, and discount factor all increase the option value of consumption.

3.2 Prices with Sampling

If the firm does not provide samples, then its optimization program simply consists of choosing the price which maximizes its profits given the consumers' ex ante uncertainty of a match. Once the uncertainty is resolved, however, a consumers' willingness to pay increases to equal the benefit they derive from the drug. Thus, if the firm cannot perfectly price discriminate, it would prefer to minimally discriminate between new and existing consumers to extract some of the consumer surplus available after the uncertainty has been resolved. If the firm could discriminate between new and existing consumers, then its problem consists of choosing a pair of prices $\{p_n, p_e\} \in \mathbb{R}^2_+$ where p_n is the price the firm charges new consumers who do not yet know if they are a match, and p_e is the price it charges consumers who already know they are a match. With this type of third degree price discrimination, the consumer will choose to buy the drug when

$$EU_i = (\mu\theta_i - p_n) + \max\{\mu(\Sigma-1)(\theta_i - p_e), 0\} \geq 0. \tag{8}$$

Rearranging (8) allows the participation constraint to be expressed by consumer type as

$$\theta_i \geq \begin{cases} \tilde{\theta}(p_n, p_e \mid \mu) \equiv \dfrac{p_n + \mu(\Sigma-1)p_e}{\mu\Sigma} & \text{if } \mu p_e \leq p_n, \\ \tilde{\theta}(p_n \mid \mu) \equiv p_n/\mu & \text{otherwise.} \end{cases}$$

When $\mu p_e \leq p_n$ the consumer considers the entire option value of consumption when deciding to make their initial purchase, otherwise, they only consider the value from the first period of

[13] For completeness, let $H(p) = 1-F\big(\hat{\theta}(p)\big)$ and $H'(p) = -f\big(\hat{\theta}(p)\big)\hat{\theta}'(p)$, then the first-order condition of (6) yields $H(p) + pH'(p) = 0$, and the second-order condition yields $2H'(p) + pH''(p)$. The monotone hazard rate condition on $F(\cdot)$ implies $-(H'(p))^2 + H''(p)H(p) < 0$, or $H''(p)H(p) < (H'(p))^2$. Moreover, because $p \leq \overline{\theta}$, $2H'(p) + pH''(p)$ has the same sign as $-2(H'(p))^2 + HH''$. Thus the second order condition is negative and $E\Pi^{NS}(p;\phi)$ is concave.

consumption as it exceeds the value from subsequent consumption. Using the appropriate definition of the cut-off type, $\hat{\theta}$, the firm's average, expected profit is defined as

$$E\Pi_{PD}^{NS}(p_n, p_e; \phi) = \begin{cases} \big(p_n + \phi(T-1)p_e\big)\big[1 - F\big(\tilde{\theta}(p_n, p_e \mid \mu)\big)\big] & \text{if } \mu p_n < p_e, \\ p_n\big[1 - F\big(\tilde{\theta}(p_n \mid \mu)\big)\big] + p_e\phi(T-1)\big[1 - F(p_e)\big] & \text{otherwise.} \end{cases}$$

Let p^m denote the full information monopoly price; that is, p^m is the price solving

$$p^m = \frac{1 - F(p^m)}{f(p^m)}. \tag{9}$$

By taking the first-order conditions of $E\Pi_{PD}^{NS}$, it is easy to show that $\{p_n, p_e\} = \{\mu p^m, p^m\}$ is the unique solution to the firm's maximization problem. The intuition is clear, when the firm can discriminate between new and existing consumers, it maximizes profits by charging the full information monopoly price to those consumers who know they are a match, and for those consumers who do not yet know they are match, the firm maximizes profits by charging the full information monopoly price discounted by the probability of being a match.

On the other hand, if the firm provides a free sample to new consumer i, then because the price the consumer pays for the first period of treatment is zero, consumer i's initial participation constraint is simply

$$\theta_i \mu \geq 0.$$

Because a consumer has limited liability— they cannot be made worse off from trying a free sample of the drug— the participation constraint is satisfied for all consumers.[14]

After trying the drug, if the consumer is a successful match, then his continuation constraint is

$$\theta_i \geq p_S.$$

Consequently, when the firm provides samples, the total (revenue generating) per period demand is

$$D(p_S) = \phi(T-1)\big(1 - F(p_S)\big), \tag{10}$$

which is independent of the expected efficacy μ. A firm cannibalizes the first period of demand from new consumers when it provides new consumers free samples so all of the firm's revenue is from returning consumers. Proportion ϕ of the new consumers will have a successful match, and $1 - F(p_S)$ of these will have a willingness to pay higher than the price, thus will purchase the drug for an expected total of $T-1$ additional periods.

Using the per period demand in (10), the firm's expected, average profit when it provides free samples is

$$E\Pi^S(p_S; \phi) = p_S\phi(T-1)\big(1 - F(p_S)\big) - c_S. \tag{11}$$

From the first-order condition it is clear that $p_S^* = p^m$ is the optimal price when samples are provided. Thus, the firm's revenue under sampling is identical to that under third degree price

[14] A worthwhile extension of the model may be to relax the assumption of limited liability so a bad match can make a person worse off such as happens with an adverse drug reaction. Indeed, perhaps in some therapeutic classes more relevant than learning if a drug is a match is learning that a drug produces no negative effects. However, changing the model to consider a match as the indication of a side effect does not change the insights of the paper so we will consider a match as a positive outcome and not as the absence of a negative outcome.

discrimination for all returning consumers. However, the revenue from a consumer's first period of treatment differs significantly. With third degree price discrimination the firm earns $\mu p^m\big(1-F(p^m)\big)$ but, with sampling, the firm loses c_S in the first period. Thus, third degree price discrimination is always more profitable than sampling. Moreover, the firm always earns a higher profit under third degree price discrimination than with a uniform spot price.[15]

In light of the fact that third degree price discrimination dominates either alternative, it comes as somewhat of a surprise that it is not implemented in any significant way by pharmaceutical firms.[16] An analysis of this particular question is beyond the scope of this paper. However, the reason third degree price discrimination is not used most likely lies with the structure of the industry. For instance, because all drugs are sold through an intermediary, a firm cannot simply set a lower price for new consumers. Instead, the firm must provide a coupon or voucher that the consumer redeems at the pharmacy. Given that pharmacies negotiate their wholesale prices with firms and set the retail price based on a host of factors such as the presence of competition and consumer search costs,[17] implementing third degree price discrimination may simply be untenable in pharmaceutical markets. In contrast, providing samples is relatively straightforward to implement as it does not rely on the downstream reseller and concomitant pricing dynamics.

Because third degree price discrimination dominates both a uniform spot price and sampling but is not practiced to any meaningful degree, it is reasonable to limit the firm's choice set to that of sampling and setting a uniform spot price without sampling. Therefore the analysis of the sampling strategy will concentrate on these two alternatives.

Before leaving this section it will be convenient to formally compare the uniform spot price without sampling, to the price the firm sets when it provides samples. The following lemma makes this comparison.

Lemma 1.

$$p^*_{NS}(\mu) = \frac{\mu\Sigma}{1+\mu(\Sigma-1)}p^*_S. \tag{12}$$

Proof. See appendix □

Supporting the argument by opponents to sampling that sample dispensation leads to higher drug costs (Rabin, 2007), the price the firm can charge by dispensing samples is unambiguously higher than the price it could otherwise charge. More importantly, by plugging Eq. (12) into the no-sampling demand of Eq. (5) it is clear that the firm chooses to cover the same proportion of the market with or without sampling, which has important welfare implications as discussed in Section 6.

[15] See the Appendix for more details and a proof of this statement.

[16] Co-pay cards were recently introduced by many major pharmaceutical companies but because they are designed to cover an insured patient's co-pay, they are similar in spirit to a free sample. From a marketing point-of-view a co-pay card is appealing to pharmaceutical firms because it allows them to more closely track utilization. It will be interesting to see if co-pay card use increases to replace sampling altogether, or if they remain a complementary, and somewhat niche, form of marketing.

[17] See for example Perloff and Salop (1985) for a model of pricing incorporating search costs and spacial differentiation. See Sorensen (2000) for evidence that drug price dispersion correlates with the consumer search costs.

4 Equilibrium Sampling

A firm will provide samples to a consumer if and only if the expected profit from providing samples is higher than the expected profit from not providing samples:

$$E\Pi^S(p_S^*;\phi) > E\Pi^{NS}(p_{NS}^*(\mu);\phi). \tag{13}$$

In Sections 4.1 and 4.2 we analyze the firm's optimal sampling and pricing decisions when the efficacy of its drug is commonly known, and known only to the firm, respectively. The essential strategy difference between the two information regimes is a consequence of how $p_{NS}^*(\mu)$ is affected by the state of information. Under symmetric information p_{NS}^* changes with the efficacy since efficacy is commonly known, but under asymmetric information p_{NS}^* remains constant resulting in significantly different sampling strategies.

4.1 Symmetric Information

When information is symmetric, the value of ϕ is common knowledge. However, consumers still do not know ex ante if they are a match with the drug. By providing samples the firm allows consumers to learn their match before they pay for the drug. Consequently the firm can charge the full information monopoly price without holding up new purchases. In order for sampling to be optimal, (13) implies the following must be true

$$p_S^*\phi(T-1)\big(1-F(p_S^*)\big) - c_S > p_{NS}^*(\mu)(1+\phi(T-1))\big[1-F\big(\hat{\theta}(p_{NS}^*(\mu)\mid\phi)\big)\big]. \tag{14}$$

The left-hand-side of (14) is the profit the firm earns in one period if it provides samples and the right-hand-side is the profit the firm earns if it does not provide samples. Because the efficacy is common knowledge $\mu = \phi$.

Let $R(p^m) \equiv p^m\big(1-F(p^m)\big)$, the unit revenue of matched consumers. By substituting p_S^* for p_{NS}^* the profit comparison of (14) can be expressed as

$$\phi(T-1)R(p^m) - c_S > \left(\frac{\phi\Sigma}{1+\phi(\Sigma-1)}\right)(1+\phi(T-1))R(p^m). \tag{15}$$

The set of efficacies, Φ, for which the firm will find it optimal to provide samples are all ϕ satisfying (15). It is clear from (15) that the profit curve under sampling is linear in the efficacy whereas the no-sampling profit curve is convex in the efficacy. The no-sampling profit curve's convexity is a result of two effects from increasing efficacy. First, with a higher efficacy, more consumers will be a successful match increasing the revenue in subsequent periods; and second, because the expectation of being a match is higher, the firm can charge a higher price, further increasing revenues.

From inspection of (15) we can gain two insights into the firm's sampling decision under symmetric information. First, it is clear that, because there is a cost to providing samples, the following is true:

$$E\Pi^S(p_S^*;0) < E\Pi^{NS}(p_{NS}^*(0);0).$$

Thus, because of the continuity of the profit functions, if there are efficacies for which the firm would like to optimally provide samples, then they must be higher efficacies. Second, it is also clear that because when everybody is a match with the drug the firm can charge the full-information price, the following is true:

$$E\Pi^S(p_S^*;1) < E\Pi^{NS}(p_{NS}^*(1);1).$$

Thus, if there are efficacies for which the firm would like to optimally provide samples, then they must be for efficacies in the interior of the interval $[0, 1]$. This makes intuitive sense, because, when the efficacy is common knowledge, a firm with a drug having a high efficacy will be able to charge a price sufficiently close to the full information price that the potential gains to providing samples are small, hence eliminated by the cost of providing the samples. A drug with a high efficacy may be characterized as more of a *search good* than an *experience good* because a consumer will have a high degree of confidence of being a match. On the other hand, despite being able to charge a higher price by providing samples, a firm with a drug having a lower efficacy will also choose not to provide samples. When a drug has a low efficacy, too few consumers will return to continue treatment, resulting in a loss in profit which exceeds the gain from the attendant higher price.

As a result of the fact that the no-sampling profit is increasing and convex in ϕ and the sampling profit is a linear function of ϕ, when the two functions cross, they cross twice within the interval $[0, 1]$. Consequently the sampling decision is not monotonic in efficacy. The following proposition formally defines the sampling set.

Proposition 1. *A monopolist with a drug of known efficacy will provide samples if and only if* $\phi \in \Phi = (\underline{\phi}, \overline{\phi}) \subset [0, 1]$ *where*

$$0 < \underline{\phi} = \frac{\Omega - \sqrt{\Omega^2 - 4c_S(T-1)R(p^m)}}{2(T-1)R(p^m)}, \quad \overline{\phi} = \frac{\Omega + \sqrt{\Omega^2 - 4c_S(T-1)R(p^m)}}{2(T-1)R(p^m)} < 1,$$

and $\Omega \equiv \big[(T-1) - \Sigma\big]R(p^m) + c_S(\Sigma - 1)$.

Proof. See appendix. □

By multiplying both sides of (15) by $1 + \phi(\Sigma - 1)$ it is clear that the sampling decision is a quadratic function of efficacy. Solving this quadratic reveals the interval of efficacies identified by Proposition 1. Moreover, because the decision is quadratic in ϕ, there exists the possibility that there is no real solution to the problem; i.e., when the discriminant of the quadratic is negative the values of the parameters are such that the firm will optimally not provide samples for any realization of the efficacy. By inspection of the discriminant, it can be seen that this occurs when the cost to sampling is too high, there are too few periods of continued treatment by the consumer, or the consumer does not discount the future too heavily permitting the firm to charge a high no-sampling price.

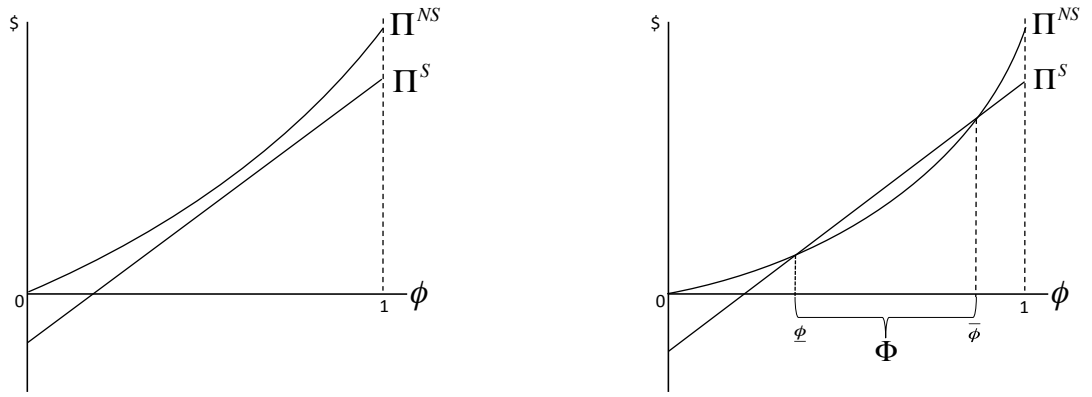
Figure 1 illustrates Proposition 1 by showing two example market conditions. Figure 1(a) depicts a market where the consumers have a sufficiently high discount factor such that Φ is empty and Figure 1(b) depicts a market in which there do exist efficacies for which a firm will provide samples. Reflecting the option value of consumption, the convexity of the no-sampling profit function is decreasing with the discount factor and increasing in the number of periods.[18] The convexity allows $E\Pi^{NS}$ to fall below $E\Pi^{S}$, which is linear in efficacy.

4.2 Asymmetric Information

Before deriving the sampling equilibrium under asymmetric information, it is important to note that a firm cannot use the drug's price to signal the drug's efficacy in a similar manner to how a firm

[18] Formally,

$$\frac{\partial}{\partial \delta}\left\{\frac{\partial^2 \Pi^{NS}}{\partial \phi^2}\right\} < 0 \text{ and } \frac{\partial}{\partial T}\left\{\frac{\partial^2 \Pi^{NS}}{\partial \phi^2}\right\} > 0.$$



(a) No Sampling

(b) Sampling

Figure 1: Examples of a monopolist which will provide samples and a monopolist which will not. In (b) the no-sampling profit line is sufficiently convex therefore a monopolist is better off providing samples for every $\phi \in \Phi$.

can signal quality through price in the introductory pricing literature. This limitation is important because, if a firm could signal the efficacy level of a drug through its choice of price, then the sampling equilibrium would be similar, if not identical, to the symmetric information equilibrium and the act of providing samples has no other affect on the equilibrium. However, we show that because the price cannot signal efficacy, the act of providing samples functions as a signal of higher efficacy altering the consumers' belief, μ.

Proposition 2. *Under asymmetric information, a monopolist cannot signal the efficacy of the drug through price.*

Proof. Suppose the firm can signal the value of ϕ with a price $p(\phi)$, then announcing a price is the same as announcing the efficacy, ϕ. If the price announcement is to credibly signal the efficacy, then it must be the case that the firm cannot do better with an alternative price announcement. That is, if a firm's drug has efficacy ϕ, then an equilibrium must satisfy $E\Pi\big(p(\phi),\phi\big) \geq E\Pi\big(p(\hat{\phi}),\phi\big)$ for all $\phi, \hat{\phi} \in [0,1]$ where $\phi \neq \hat{\phi}$, and

$$E\Pi\big(p(\phi);\phi\big) = p(\phi)\big(1-\phi(T-1)\big)\big[1-F\big(\hat{\theta}(p(\phi)\mid\phi)\big)\big], \tag{16}$$

$$E\Pi\big(p(\hat{\phi});\phi\big) = p(\hat{\phi})\big(1-\phi(T-1)\big)\big[1-F\big(\hat{\theta}(p(\hat{\phi})\mid\hat{\phi})\big)\big]. \tag{17}$$

Therefore, using (16) and (17), $E\Pi\big(p(\phi);\phi\big) \geq E\Pi\big(p(\hat{\phi});\phi\big)$ implies

$$p(\phi)\big[1-F\big(\hat{\theta}(p(\phi)\mid\phi)\big)\big] \geq p(\hat{\phi})\big[1-F\big(\hat{\theta}(p(\hat{\phi})\mid\hat{\phi})\big)\big],\ \forall\hat{\phi},\phi\in[0,1]. \tag{18}$$

However, if a firm's drug has efficacy $\hat{\phi}$, then $E\Pi\big(p(\hat{\phi});\hat{\phi}\big) \geq E\Pi\big(p(\phi);\hat{\phi}\big)$ for all $\hat{\phi},\phi \in [0,1]$ implying

$$p(\hat{\phi})\big[1-F\big(\hat{\theta}(p(\hat{\phi})\mid\hat{\phi})\big)\big] \geq p(\phi)\big[1-F\big(\hat{\theta}(p(\phi)\mid\phi)\big)\big]\ \forall\hat{\phi},\phi\in[0,1]. \tag{19}$$

Both (18) and (19) are correct if and only if $p(\phi) = p(\hat{\phi})$ for all $\phi, \hat{\phi} \in \Phi$, a contradiction. Therefore price cannot credibly signal a drug's efficacy. □

The intuition behind why a firm cannot utilize price to signal efficacy is because the firm does not incur a penalty for misreporting. In the introductory pricing literature the quality level affects the ex post valuation of the good; consequently, a firm setting a high price for a low quality good is penalized with a large loss in demand after consumers learn that the quality level is insufficient for the price. In the current model, the ex post match value of the drug is independent of the realization of ϕ. Note that this result is not a consequence of the deterministic ex post match value. If the ex post match value was itself a random variable, then the price could be used to credibly signal some characteristic, such as the mean of the distribution of match values. In this way, it is the idiosyncrasy of being a match which distinguishes pharmaceuticals from most other experience good markets.

When the firm is not able to provide samples, perhaps because of regulation, then the consumers' beliefs will simply be the unconditional prior, $\mu = \int_0^1 \phi \; dG(\phi)$. However, if a firm has the ability to provide samples, then whether or not it dispenses samples furnishes consumers with additional information with which to condition their belief. Because the firm's optimal strategy must now take into account how its actions affect the consumers' beliefs, and how those beliefs restrict what the firm can do, the appropriate equilibrium concept is a perfect Bayesian equilibrium.

Definition 1. *A perfect Bayesian equilibrium of the sampling game consists of the tuple*

$$\langle \mu^*(s^*), p^*(\phi), s^*(\phi) \rangle,$$

where $\mu(s^*)$ is the consumers' common equilibrium belief about the drug's efficacy given $s^* \in \{S = \text{Sample}, NS = \text{No Sample}\}$, p^* is the firm's equilibrium price, and s^* is the firm's equilibrium sampling choice.

When the firm has the ability to provide samples, then the consumers' prior is conditioned on the sampling decision of the firm. Define the consumers' belief given the firm does not provide samples as $\mu_{NS} = E[\phi \mid s = NS]$. Define $\hat{\phi} \in [0, 1]$ as the efficacy solving

$$E\Pi^S\big(p_S^*; \hat{\phi}\big) = E\Pi^{NS}\big(p_{NS}^*(\mu_{NS}); \hat{\phi}\big).$$

That is, $\hat{\phi}$ represents the the efficacy level where the firm is just indifferent between providing samples and not providing samples. To see how the firm's optimal strategy changes under asymmetric information, observe that the firm faces two problems that were not present when information is symmetric.

The first problem the firm encounters when information is asymmetric is that it cannot signal when the drug has a very high efficacy. Recall that when the efficacy is common knowledge, the firm finds it optimal to not provide samples for very high efficacy drugs. However, now that consumers cannot distinguish when the firm is not providing samples because its drug has a very high efficacy from when the firm is not providing samples because its drug has a lower efficacy, the firm must optimally provide samples for all high efficacy drugs. Because the firm finds it optimal to provide samples for all high efficacy drugs, when it exists, the cut-off efficacy $\hat{\phi}$ will be unique, and the sampling strategy is monotonic in efficacy. The following lemma formally states this.

Lemma 2. *The cut-off efficacy $\hat{\phi}$ is unique, and the firm will provide samples for any efficacy above the cut-off.*

The second problem the firm encounters when information is asymmetric is that the act of dispensing samples furnishes information about the efficacy. To see why this introduces a problem, define $\lambda(\hat{\phi})$ as the unique[19] $\lambda \in (0,1)$ satisfying

$$\lambda(\hat{\phi})\hat{\phi} = E[\phi \mid \phi < \hat{\phi}] = \int_0^{\hat{\phi}} \phi \; dG(\phi) \Big/ G(\hat{\phi}) \; . \tag{20}$$

The right-hand-side of (20) identifies the consumers' expectation of the drug's efficacy given the firm has not provided samples; i.e., given that the efficacy must be less than the cut-off $\hat{\phi}$. The coefficient $\lambda(\hat{\phi})$ can be interpreted as a measure of the degree to which not providing samples lowers the consumers' belief. A heavier lower tail for the distribution G results in a $\lambda(\hat{\phi})$ closer to zero. Consequently the price the firm can charge will naturally be lower when it is more likely that the drug has a low efficacy. Because the firm must now take the consumers' beliefs into account when making its sampling decision, the firm finds it optimal to provide samples for lower efficacies than it would have if the consumers knew the efficacy. This argument is captured in the following lemma.

Lemma 3. *The sampling cut-off efficacy is always lower under asymmetric information than under symmetric information; i.e.,* $\hat{\phi} < \underline{\phi}$.

However, as when information is symmetric, the firm may not find it optimal to provide samples for any efficacy in $[0,1]$. As with symmetric information, this occurs when there is an excessive cost to sampling, there are too few number of periods of repeat purchase by the consumer, or the consumer does not discount the future too heavily permitting the firm to charge a high no-sampling price.

The following proposition combines lemmas 2 and 3 and formalizes the above discussion. The proposition shows that there is a unique sampling equilibrium, and when $\hat{\phi}$ exists, it is unique and the equilibrium is separating, otherwise the equilibrium is pooling as the firm will not provide samples for any efficacy of the drug.

Proposition 3. *If the efficacy, ϕ, of a drug is not known by consumers and has distribution G, with strictly positive density g over the supports $[0,1]$, then there exists a unique perfect Bayesian equilibrium.*

The equilibrium is separating when the parameters satisfy

$$(T-1) - \lambda(\hat{\phi})\big[\Sigma - c_S(\Sigma - 1)\big] > 2\sqrt[+]{\lambda(\hat{\phi})c_S(T-1)R(p^m)},$$

and will have the following form:

1. *Consumers have conditional beliefs*

$$\mu^*_{NS} = \lambda(\hat{\phi}^*)\hat{\phi} = \frac{\int_0^{\hat{\phi}} \phi \; dG(\phi)}{G(\hat{\phi})} \qquad \text{and} \qquad \mu^*_S = \frac{\int_{\hat{\phi}^*}^1 \phi \; dG(\phi)}{1 - G(\hat{\phi})}.$$

2. $s^* = S$ *whenever* $\phi > \hat{\phi}$, *otherwise* $s^* = NS$.

3. $p^* = p^m$ *when* $s^* = S$ *and* p^* *is as in (7) when* $s^* = NS$.

[19] Recall G has strictly positive density over the supports $[0,1]$.

4. *The sampling set is* $\Phi = (\hat{\phi}, 1]$.

5. *The cut-off efficacy is* $\hat{\phi} = \frac{\tilde{\Omega} - \sqrt[+]{\tilde{\Omega}^2 - 4\lambda(\hat{\phi})c_S(T-1)R(p^m)}}{(T-1)R(p^m)}$, *where* $\tilde{\Omega} \equiv \big((T-1) - \lambda(\hat{\phi})\Sigma\big)R(p^m) - \lambda(\hat{\phi})c_S(\Sigma - 1)$.

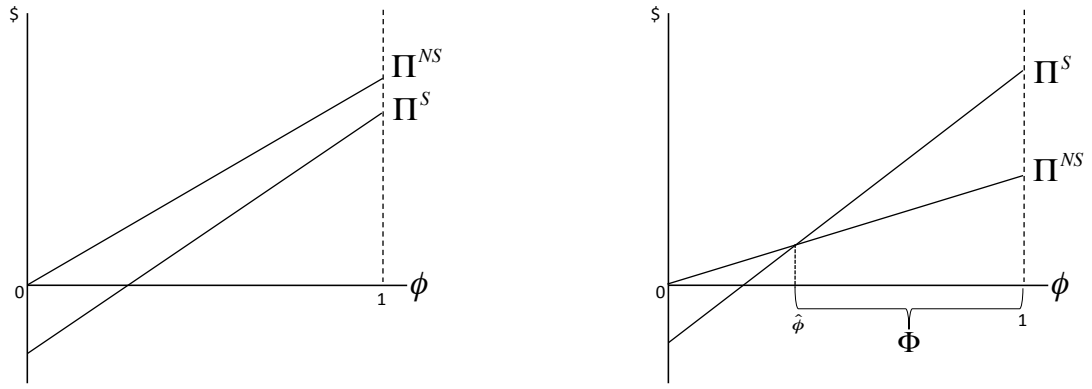
The equilibrium is pooling when the parameters satisfy

$$(T-1) - \lambda(\hat{\phi})\big[\Sigma - c_S(\Sigma - 1)\big] \leq 2\sqrt[+]{\lambda(\hat{\phi})c_S(T-1)R(p^m)}.$$

In the pooling equilibrium the sampling set of efficacies is empty ($\Phi = \emptyset, s^* = NS$), *consumers have the unconditional belief* $\mu_{NS} = \int_0^1 \phi \; dG(\phi)$, *and* p^* *is as in (7).*

Proposition 3 shows that the information asymmetry results in a monotonic sampling strategy. Venkataraman and Stremersch (2007) provide supporting evidence by observing that samples are more likely to be provided for drugs with a high efficacy. Sampling for low efficacy drugs is not profitable and that fact drives consumers' beliefs.

Figure 2 highlights how the firm's sampling strategy changes under asymmetric information by again illustrating the sampling strategy in two example markets. Figure 2(a) depicts a market where the consumers have a sufficiently high discount factor such that Φ is empty. Figure 2(b) depicts a market with a lower discount factor in which there are some efficacies with which a firm will provide samples. The key distinction between the examples in Figures 1 and 2 is the shape of the no-sampling profit curve. When the efficacy of the drug is not commonly known, the no-sampling profit curve is linear since the consumers' beliefs result in a single price $p^*_{NS}(\mu)$ for all realizations of the efficacy. The linearity of the firm's no-sampling profit results in a single-crossing of the profit curves and, thus, a unique cut-off efficacy $\hat{\phi}$.



(a) No Sampling

(b) Sampling

Figure 2: Examples of a monopolist which will provide samples and a monopolist which will not when the efficacy of the drug is not common knowledge.

5 Sampling with Informed and Uninformed Consumers

An interesting extension to the analysis is to consider the firm's sampling strategy when some consumers are informed and others are not. For example, physicians may have different levels of experience or receive evaluative information about drugs from different sources such as studies published in medical journals or word of mouth from colleagues. Consequently some physicians may be better informed than others about any particular drug's efficacy. Pharmaceutical firms spend billions of dollars annually on physician detailing in which they discuss their firm's drugs with physicians and it is reasonable to assume that the representatives develop a sense of the physicians' beliefs regarding the efficacy of their drug through the detailing process. With this information the representative can decide whether or not to leave any particular physician with samples. Supporting the notion that pharmaceutical firms partially base their sampling decision on observable characteristics of physicians, Rabin (2007) observes that less experienced physicians were more likely to be given samples and Boltri et al. (2002) report that less experienced physicians tend to change their prescription choices more than experienced physicians when samples are present. If less experienced physicians are uncertain of a drug's efficacy because of their lack of experience with a drug, then the firm may find it optimal to provide them samples. In contrast there may be no advantage to a firm providing experienced physicians samples when those physicians already know the efficacy of the drug.

To show how a mix of consumer knowledge alters the firm's sampling decision, consider the following addition to the model. Suppose a proportion γ of the consumers know the efficacy of the drug (call them type-s consumers), whereas the proportion $1-\gamma$ do not know the drug's efficacy but have the prior G (call them type-a consumers). Capturing the assumption that pharmaceutical firms learn the phsycians' priors through the detailing process, we assume that the firm knows each consumer's type, but consumers do not know the other consumers' prior, or their proportions. Let Φ_a denote the optimal sampling set for type-a consumers, and let Φ_s denote the optimal sampling set for type-s consumers. As lemmas 2 and 3 show, the sampling set of efficacies for type-a consumers is larger than, and a superset of the sampling set for type-s consumers; i.e., $\Phi_s \subset \Phi_a$.

Let p^j_{NS} for $j \in \{a, s\}$ denote the optimal price the firm can charge either type of consumer given it does not provide samples. Instead of a simple sample/no-sample strategy space, because the sampling set for the less informed consumers is a superset of the sampling set for the informed consumers, the firm now has four feasible strategies which we define as follows:

1. *Persistent Sampling* ($\phi \in \Phi^s$): The firm can choose to serve both types by providing samples to everybody and setting the price to p^*_S. The firm will utilize this strategy when it is optimal to provide samples under symmetric information. A drug in this category will be heavily sampled throughout its product life. Examples may include drugs from the top four therapeutic categories, asthma and allergy remedies, anti-infective agents, analgesics and anti-inflammatory medications, and antihypertensive drugs, which collectively account for more than 63 percent of all drug samples dispensed (Backer et al., 2000).

2. *Introductory Sampling* ($\phi \in \Phi^a$ and $\phi \notin \Phi^s$): The firm can choose to serve both types of consumer by providing samples to the type-a consumers and setting the price to p^s_{NS}. The firm would select this option when it is not optimal to provide samples under symmetric information. Drugs in this category will exhibit heavy sampling when introduced to the market and the proportion of uninformed consumers is large, but the volume of sampling will decline over time as physicians gain experience with the drug and learn its efficacy. Drug

sampling will also be higher with residents and physicians who are early in their career than more experienced physicians. The majority of drug sampling follows this strategy as for most therapeutic classes sampling is heaviest right after a new drug has been introduced to the market (Joseph and Mantrala, 2009).

3. *Established Pricing* ($\phi \notin \Phi^a$ and $(1-\gamma)E\Pi^{NS}(p^s_{NS}) > E\Pi^{NS}(p^a_{NS})$): The firm can choose to not provide samples and only serve the consumers who know the efficacy by setting the price to p^s_{NS}. Because this strategy creates a hold-up problem for the uninformed, it will only be useful for well established drugs where the proportion of consumers who are uninformed is likely to remain low. The firm will select this strategy option when it is not optimal to provide samples under symmetric or asymmetric information.

4. *Introductory Pricing* ($\phi \notin \Phi^a$ and $E\Pi^{NS}(p^a_{NS}) > (1-\gamma)E\Pi^{NS}(p^s_{NS})$): The firm can choose to not provide samples and serve both types by setting the price to p^a_{NS}. This outcome is similar to *Established Pricing*. The firm's choice would hinge on the size of the uninformed market and the differential between p^a_{NS} and p^s_{NS}; i.e., the firm will choose this strategy over the previous strategy if $p^a_{NS} > \gamma p^s_{NS}$. As the proportion of informed consumers increases the firm's strategy will flip over to the established pricing strategy.

Predicting which strategy a firm will use is of course an empirical problem. The model contains several parameters, many of which will likely vary considerably from one drug to another, and from one therapeutic class to another. However, given the values of the parameters for a given drug or therapeutic class, the model predicts behavior inline with observations of sampling behavior in the previously cited medical literature.

6 Welfare

As the concern with sample dispensation has primarily been on how the practice affects consumers, we proceed by first focus on consumer welfare before examining overall social surplus. The welfare analysis should be viewed as a baseline as there are important factors not explicitly modeled here that will also affect welfare.[20]

6.1 Consumer Surplus

When the sampling set, Φ, is empty and the equilibrium takes the form of the pooling equilibrium described in Proposition 3, then the firm will not provide samples for any realization of the efficacy. The fact that the firm does not provide samples provides no additional information to consumers so there is no change in consumer or producer surplus from permitting sampling.

However, when the sampling set is nonempty then consumer surplus will be altered depending on whether or not sampling is permitted. To see how consumer surplus is altered, let μ_0 represent the consumers' unconditional prior when sampling is not permitted; that is, let $\mu_0 \equiv \int_0^1 \phi \; dG(\phi)$.

[20]For example, there is some concern that sample package labeling and instruction is insufficient, thus resulting in unsafe drug consumption (Backer et al., 2000), reducing the consumer benefit to sampling. However, because sample dispensation is more profitable for the firm, the practice increases the firm's available funds for drug development. An increase in the firm's profit will potentially increase the rate at which new drugs are developed and brought to market, increasing the consumer benefit to sampling.

Suppose the efficacy of the firm's drug is in the sampling set, then the per period change in consumer surplus from permitting samples is

$$\begin{aligned} CS_S - CS_{NS} &= \Delta U_1 + \phi \sum_{t=2}^{T} \Delta U_t \\ &= \int_0^1 \phi\theta \; dF(\theta) - \int_{\hat{\theta}\left(p^*_{NS}(\mu_0)|\mu_0\right)}^1 \left(\phi\theta - p^*_{NS}(\mu_0)\right) \; dF(\theta) - \phi(T-1) \int_{p^*_S}^1 \Delta p^* \; dF(\theta) \\ &= \underbrace{\int_0^{\hat{\theta}\left(p^*_{NS}(\mu_0)|\mu_0\right)} \phi\theta \; dF(\theta)}_{\text{(A)}} + \underbrace{\int_{\hat{\theta}\left(p^*_{NS}(\mu_0)|\mu_0\right)}^1 p^*_{NS}(\mu_0) \; dF(\theta)}_{\text{(B)}} - \underbrace{\phi(T-1) \int_{p^*_S}^1 \Delta p^* \; dF(\theta)}_{\text{(C)}}, \end{aligned} \tag{21}$$

where $\Delta p^* = p^*_S - p^*_{NS}(\mu_0)$. The change in consumer surplus from permitting sampling comes from three components. The first component, (21.A), identifies the consumer surplus derived from the uncovered market which gains one period of treatment from receiving a free sample.[21] The second component, (21.B), identifies the surplus change from the proportion of the market which is covered, but now does not have to pay for their first period of consumption. Finally, (21.C) identifies the surplus change from the higher price paid in all subsequent periods of consumption. The first two components represent a gain in consumer surplus and the third component represents a loss.

Eq. (21) can be rearranged to derive the following result.

Proposition 4. *Suppose a monopolist's drug has unknown efficacy ϕ and Φ is the sampling set. Consumer welfare is increased by permitting sampling under the following conditions:*

1. *When $\phi \in \Phi$, consumer welfare is increased if and only if*

$$\phi \int_0^{p^m} \theta \; dF(\theta) > p^m \phi(T-1)\big(1 - F(p^m)\big) - p^*_{NS}(\mu_0)\big(1 + \phi(T-1)\big)\big(1 - F(p^m)\big) \tag{22}$$

2. *When $\phi \notin \Phi$, welfare is unchanged when Φ is empty and welfare is always increased when Φ is nonempty.*

The left-hand-side of (22) represents the consumer surplus derived from the uncovered portion of the market and the right-hand-side is the difference in the firm's revenues from providing samples and the firm's revenue when sampling is not permitted. In words, Proposition 4 says that when the firm chooses to provide samples, consumer surplus is increased if and only if the consumer surplus derived from the uncovered portion of the market exceeds the difference in the firm's revenue from providing samples to the firm's revenue when sampling is not permitted. Note that the firm may find it optimal to provide samples when permitted to do so, but receive a lower revenue than when sampling is not permitted. This counterintuitive result may occur because, when sampling is permitted but the firm does not provide samples, consumers infer the drug has a low efficacy and

[21] It is not unreasonable to assume that a consumer who will not continue the treatment after the first period still tries a free sample. For instance, the consumer may not realize ex ante that he will not continue the treatment even if he is a match. The consumer may first try the sample and then, only upon going to the pharmacy and learning the price of the drug, decide not to continue the treatment.

the firm is restricted to setting the price $p^*_{NS}(\mu_{NS})$, which is less than the optimal no-sampling price based on the unconditional prior, $p^*_{NS}(\mu_0)$. Sampling may thus raise the firm's profit over the profit it can earn with $p^*_{NS}(\mu_{NS})$ but not over the profit it earns with $p^*_{NS}(\mu_0)$. When the firm's revenue decreases, consumer surplus is increased by sampling regardless of whether or not the welfare accounting considers the uncovered portion of the market.

When information is symmetric, then the no-sampling revenue can never be less than the sampling revenue if the firm optimally chooses to provide samples and consumer welfare is increasing only if the surplus generated by providing samples to the uncovered portion of the market exceeds the firm's increase in revenue. In pharmaceutical markets consumers in the uncovered portion of the market would include indigent patients suggesting, if sampling increases consumer welfare, then it is largely driven by the welfare gain from those consumers who the firm would not cover with or without samples. The gain in consumer welfare identified by Proposition 4 represents a floor to the gain in consumer surplus in that the left-hand-side of Eq. (22) represents the welfare from consumers initially offered a sample but do not continue with the treatment because of an insufficient willingness-to-pay. However, some physicians may provide sufficient samples to indigent consumers that they can complete their entire treatment and this represents behavior not directly captured by the model.[22] Proposition 4 implies that the social value in sampling will depend on the number of indigent patients actually receiving free treatment via samples, highlighting the need to measure the number of patients who, because of samples, are able to receive treatment. To our knowledge, nobody has yet looked at the number of patients who have received treatment via free samples as research has primarily focused on what proportion of those receiving free samples are poor, or without insurance.[23] These studies fail to indicate if the poor or uninsured were fully treated by the samples provided, or if they were only given a starter pack of samples only to discontinue treatment after the samples were consumed.[24]

The change in consumer welfare from permitting samples is straightforward when the efficacy of the drug is sufficiently low that the firm chooses not to provide samples. Consumers observe that the firm does not provide samples so (correctly) infer that the efficacy of the drug is low. Consequently, the price the firm can charge will be lower than if sampling was not permitted. Thus, permitting sampling unambiguously increases consumer welfare when the efficacy is sufficiently low that the firm does not provide samples.

6.2 Socially Efficient Sampling

If a social planner is responsible for determining whether or not a firm should provide samples, then what would the planner choose? If the social planner is a hospital administrator interested in maximizing consumer welfare, then the planner will permit sampling only when the conditions of Proposition 4 are met. However, if the planner's interest is in maximizing total social surplus, then she may still choose to permit sampling, even if it results in lower consumer surplus.

[22] Recall that the cost of sampling, c_S, implicitly accounts for any excess samples given to indigent patients to cover their entire course of treatment; a quantity the firm is assumed to know.

[23] See for example, Cutrona et al. (2008).

[24] The largest data set capturing sample dispensation is the Medical Expenditure Panel Survey (MEPS) which identifies when a consumer receives a sample. However the MEPS does not identify the quantity of samples a consumer receives making it unsuitable for the kind of study necessary to completely identify the welfare affects of sample dispensation.

The change in producer surplus from permitting sampling is given as follows:

$$\Delta\Pi = \begin{cases} E\Pi^S\big(p_S^*;\phi\big) - E\Pi^{NS}\big(p_{NS}^*(\mu_0);\phi\big) & \text{if } \phi \in \Phi, \\ E\Pi^{NS}\big(p_{NS}^*(\mu_{NS});\phi\big) - E\Pi^{NS}\big(p_{NS}^*(\mu_0);\phi\big) & \text{if } \phi \notin \Phi. \end{cases} \tag{23}$$

When $\phi \in \Phi$, the change in producer surplus is the differences between the profits from providing samples and not providing samples and is always positive. When the firm does not provide samples, the change in producer surplus comes from the difference in the price the firm can charge. If the conditional prior μ_{NS} is less than the unconditional prior μ_0 then producer surplus will be lower. Using the definitions of $E\Pi^S$, $E\Pi^{NS}$, and the change in consumer surplus identified in (21) it is straightforward to derive the following result.

Proposition 5. *Suppose a monopolist's drug has unknown efficacy ϕ and Φ is the sampling set. If $\phi \in \Phi$, then social surplus is increased by sampling if and only if*

$$\phi \int_0^{p_S^*} \theta dF(\theta) > c_S. \tag{24}$$

Otherwise permitting sampling is socially wasteful. If $\phi \notin \Phi$, then there is no change in social surplus.

When a drug's efficacy is in the sampling set Φ, then any increase in social welfare is a result of a welfare gained by those consumers who are in the uncovered portion of the market, $\theta < p_S^*$. It is thus important to identify the number of patients receiving treatment via samples that would not otherwise receive treatment. If the welfare gain of these consumers exceeds the total cost of sample dispensation, then sample dispensation is socially efficient.

When the efficacy of the drug is not in the sampling set, there is no change in social welfare. Consumers infer a lower level of efficacy for the drug given the firm does not provide samples resulting in a lower price. The gain in consumer surplus from the lower price is a zero-sum transfer from the firm to consumers.

7 Conclusion

We have examined the effects of sampling in the prescription drug market by developing a simple, stylized model of sampling as a dynamic game of incomplete information. By assuming samples are used to mitigate the consumers' uncertainty in their drug-to-patient match and by solving for the firm's sampling strategy under symmetric and asymmetric information regimes, the model provides qualitative results consistent with observations found in the medical literature. For example, the model explains why sample dispensation varies across therapeutic classes, why it results in higher drug prices, why physicians early in their career are more likely to receive samples, and why sample dispensation changes over a drug's life cycle.

The model shows that the idiosyncratic nature of the consumer-to-drug match is the distinguishing characteristic that separates prescription drugs from other experience goods. This follows because sampling is not *cheap-talk* and the firm cannot otherwise signal the drug's efficacy with the price alone. Under asymmetric information the firm dispenses samples to facilitate learning and signal that its drug has a higher efficacy whereas under symmetric information, the firm provides samples simply to facilitate the learning of one's match outcome. The signaling aspect of sampling

under asymmetric information alters the firm's strategy in two ways. First, the firm finds that it is necessary to provide samples for drugs having a very high efficacy, even though it would not do so when information is symmetric, resulting in a monotonic sampling strategy. And second, the signaling aspect of sample dispensation causes the firm to optimally provide samples for drugs having a lower efficacy than it otherwise would when information is symmetric. Consequently the sampling set of efficacies is larger under asymmetric information. Using this property of the sampling sets, the firm's pricing and sampling strategies were identified when the market exhibits a mix of knowledge.

As sampling permits consumers to learn their match outcome before paying to consume the drug, sample dispensation generally results in a welfare transfer from consumers to firm. This follows because the firm is able to charge the full information monopoly price when it dispenses samples. However, the welfare transfer is in the consumers' favor when sampling reduces the firm's revenue in relation to the baseline case of sample prohibition. In either case, consumers in the uncovered portion of the market will gain if some are able to be treated or partially treated via samples alone as when samples are not available the uncovered portion of the market will not receive any treatment. Thus, a policy maker that is concerned with either consumer surplus or total social welfare should determine the degree to which indigent patients are treated with samples. Although most samples are not given to the indigent,[25] if some indigent patients are able to receive treatment that they would not otherwise have received had sampling been prohibited, then consumer surplus and/or social welfare can be improved by sampling. In this way, it is not contradictory to assert that sample dispensation is both a marketing tool used to secure higher profit, and a means to improve overall consumer and/or social welfare. On the other hand, if a social planner's concern is for the covered portion of the market, then there is no doubt that sampling results in a higher price and welfare is transfered from consumers to the firm.

There are two important extensions to the analysis performed here left for future work. First, because the roles of physician and patient are combined there is no room to model learning of the efficacy and hence the transition from an asymmetric information regime to a symmetric information regime. By separating the roles of physician and patient, the physician can be modeled as a Bayesian, updating her posterior with every observation of a patient's match outcome. Firms will face a variant of the one-armed bandit problem in determining the optimal number of samples to provide to facilitate learning whereas not unnecessarily cannibalizing demand. The rate at which beliefs converge to the true efficacy will provide insights into the sampling duration exhibited by introductory sampling.

A second important extension to this paper is to develop a model for sampling competition among multiple firms. Because a significant number of new drugs are "me-too" drugs, new patented drugs only slightly improved, at best, from earlier or competitor's versions, competition between patented drugs is an important characteristic of pharmaceutical markets. A physician faces a multi-armed bandit problem when she is presented with many alternative drugs which may be prescribed for the same condition and it is not clear how samples may be strategically used to aid in or alter the physician's choice.

Finally, much additional empirical research should be performed in order to provide a better picture of how and under what conditions samples are used. Are there other important characteristics of therapeutic classes that drive sample usage not captured in the model? Is efficacy the

[25] For example, Cutrona et al. (2008) report that only 10.8 percent of low income persons received a free sample (in 2003) compared to 12.3 percent for middle income and 12.8 percent for high-income persons ($P < .001$).

most important characteristic, or is the potential for adverse side-effects more important?[26] The answer to this question probably varies by therapeutic class. Finally, are a physician's beliefs and characteristics more important than a patient's? For example do some physicians exhibit a higher willingness to experiment with more treatments, or are less likely to prescribe without first being able to use samples and to what degree do patients concur with their physicians and their beliefs?[27] We leave these questions to further research.

Appendix

PROOF OF LEMMA 1

Because $\hat{\theta}(p)$ is homogeneous of degree 1 a rearrangement of (7) gives

$$\hat{\theta}(p^*_{NS}) = \frac{1 - F\big(\hat{\theta}(p^*_{NS})\big)}{f\big(\hat{\theta}(p^*_{NS})\big)}.$$

Therefore $p^m = p^*_S = \hat{\theta}(p^*_{NS})$ and

$$p^*_{NS} = \frac{\mu\Sigma}{1 + \mu(\Sigma - 1)} p^*_S.$$

Notice $\hat{\theta}(p^*_{NS}) > p^*_{NS}$ for all $\mu < 1$ therefore $p^*_S > p^*_{NS}$ for all $\mu < 1$ and all distributions of types F.

PROOF OF PROPOSITION 1

Condition (13) is both a necessary and sufficient condition for a firm to provide samples. Thus

$$\Pi^S > \Pi^{NS}$$

$$\Leftrightarrow \; p^*_S \phi(T-1)\big(1 - F(p^*_S)\big) - c_S > p^*_{NS}(1 + \phi(T-1))\big[1 - F\big(\hat{\theta}(p^*_{NS})\big)\big] \tag{A-1}$$

Recall that $R(p^m) \equiv p^m\big(1 - F(p^m)\big)$ $p^*_S = p^m$, $p^*_S = \hat{\theta}(p^*_{NS})$, and $\mu = \phi$ under symmetric information. Plugging $R(p^m)$ and Eq. (12) into (A-1) and rearranging yields

$$-\phi^2(T-1)R(p^m) + \phi\big[\big((T-1) - \Sigma\big)R(p^m) - c_S(\Sigma - 1)\big] - c_S > 0. \tag{A-2}$$

Let $\Omega \equiv \big((T-1) - \Sigma\big)R(p^m) - c_S(\Sigma - 1)$, then the set of ϕ which will lead to sampling are all of the $\phi \in (\underline{\phi}, \overline{\phi})$ where $\underline{\phi}$ and $\overline{\phi}$ are.

$$\underline{\phi} = \frac{\Omega - \sqrt{\Omega^2 - 4c_S(T-1)R(p^m)}}{2(T-1)R(p^m)} \tag{A-3}$$

$$\overline{\phi} = \frac{\Omega + \sqrt{\Omega^2 - 4c_S(T-1)R(p^m)}}{2(T-1)R(p^m)} \tag{A-4}$$

[26] Venkataraman and Stremersch (2007) perform some work comparing the importance of efficacy versus the number and likelihood of side effects on the effectiveness of physician detailing and likelihood of sample dispensation. The study finds that physicians are more likely to provide samples for higher efficacy drugs and drugs with fewer side effects as this model predicts, however it is not clear from the study if the firms were themselves are providing more samples for drugs with these characteristics.

[27] For example Backer et al. (2000) found significant differences in physician sample dispensation behavior. In one two physician practice they found that one physician dispensed samples in 41.9 percent of all observed patient encounters whereas the other physician only dispensed samples in 3.2 percent of the observed patient encounters.

A real solution exists if and only if the discriminants of (A-3) and (A-4) are positive; i.e., if and only if $\big[\big((T-1)-\Sigma\big)R(p^m)-c_S(\Sigma-1)\big]^2 > 4c_S(\Sigma-1)R(p^m)$.

If both $\underline{\phi},\overline{\phi} < 0$, then there does not exist any set $\Phi \subseteq [0,1]$ such that the firm will provide samples for any $\phi \in \Phi$. If both $\underline{\phi},\overline{\phi} > 0$, then $\big((T-1)-\Sigma\big)R(p^m)-c_S(\Sigma-1) > 0$.

To see that $\overline{\phi} < 1$ when $\underline{\phi} > 0$, start with (A-4):

$$\overline{\phi} = \frac{\Omega + \sqrt{\Omega^2 - 4c_S(T-1)R(p^m)}}{2(T-1)R(p^m)} \gtrless 1.$$

Rearranging yields

$$\sqrt{\big[\big((T-1)-\Sigma\big)R(p^m)-c_S(\Sigma-1)\big]^2 - 4c_S(T-1)R(p^m)} \gtrless \big((T-1)+\Sigma\big)R(p^m)+c_S(\Sigma-1).$$

Because $\big((T-1)+\Sigma\big)R(p^m)+c_S(\Sigma-1) > \big((T-1)-\Sigma\big)R(p^m)-c_S(\Sigma-1)$, it must be the case the the RHS is greater than the LHS of the above inequality thus $\overline{\phi} < 1$ and $\Phi = [\underline{\phi},\overline{\phi}] \subset [0,1]$.

Therefore a firm will provide samples if and only if $\phi \in \Phi = [\underline{\phi},\overline{\phi}] \subset [0,1]$.

PROOF OF LEMMA 2

The proof for this statement follows directly from the proof for Proposition 3. As the proof for Proposition 3 shows the firm's sampling strategy and the consumers' beliefs must be consistent in equilibrium. The only beliefs and sampling strategy which are consistent require a unique sampling cutoff, above which the firm always provides samples and below which it never provides samples. Finding the unique cutoff requires solving a quadratic equation, thus when the discriminant of the quadratic is negative there does not exist a real solution and there is no cutoff efficacy.

PROOF OF LEMMA 3

Let $\mathcal{G}$ be the family of distributions having positive density everywhere on the supports $[0,1]$. From Lemma 3 there exists a unique cutoff $\hat{\phi}$ such that for every $\phi > \hat{\phi}$ the firm will provide samples. Therefore, for any $G \in \mathcal{G}$ we have

$$\hat{\phi} > \lambda(\hat{\phi})\hat{\phi} = E[\phi \mid \phi < \hat{\phi}] = \int_0^{\hat{\phi}} \phi \; dG(\phi)/G(\hat{\phi}).$$

It necessarily follows that $\lambda(\hat{\phi}) < 1$ for all $G \in \mathcal{G}$.

Define $\hat{\phi}_s$ as the solution to

$$\text{(A-5)} \qquad \hat{\phi}_s(T-1)R(p^m) - c_S = \left(\frac{\hat{\phi}_s\Sigma}{1+\hat{\phi}_s(\Sigma-1)}\right)(1+\phi(T-1))R(p^m),$$

and define, $\hat{\phi}_a$ as the solution to

$$\text{(A-6)} \qquad \hat{\phi}_a(T-1)R(p^m) - c_S = \left(\frac{\lambda(\hat{\phi}_a)\hat{\phi}_a\Sigma}{1+\lambda(\hat{\phi}_a)\hat{\phi}_a(\Sigma-1)}\right)(1+\hat{\phi}_a(T-1))R(p^m).$$

The LHS of Eqs. (A-5) and (A-6) are the firm's average, expected profit given it dispenses samples and the RHS of Eqs. (A-5) and (A-6) are the firm's average, expected profit given it does not dispense samples under symmetric and asymmetric information, respectively.

First, it is clear from Eqs. (A-5) and (A-6) that $\hat{\phi}_s = \hat{\phi}_a$ if and only if $\lambda(\hat{\phi}_a) = 1$. Second, by substituting $\lambda \in (0,1)$ for $\lambda(\hat{\phi}_a)$ into (A-6) and totally differentiating gives

$$\frac{d\hat{\phi}_a}{d\lambda} > 0 \text{ for all } \lambda, \hat{\phi}_a \in (0,1). \tag{A-7}$$

Therefore, because $\hat{\phi}_a$ is monotonically increasing with λ, and $\hat{\phi}_s = \hat{\phi}_a$ if and only if $\lambda(\hat{\phi}_a)$, it must be the case that the cut-off efficacy level under asymmetric information is less than the cut-off efficacy level under symmetric information for all distributions $G \in \mathcal{G}$.

PROOF OF PROPOSITION 3

When information is asymmetric the sampling decision of the firm partially signals the efficacy to the consumers. The two cases for the proposition follow from the fact that the unique cut-off efficacy is the lower root of a quadratic equation. When the discriminant of the quadratic is negative there is no real solution and the equilibrium is pooling, when it is positive the equilibrium is separating with the firm providing samples for all efficacies above the lower root of the quadratic, and not providing samples for any efficacy below the lower cutoff. We proceed with the proof by assuming that the cutoff is unique and show that the conditions of the proposition are an equilibrium. We then show that it is the unique perfect Bayesian equilibrium.

Recall $\lambda(\hat{\phi})$ is the unique $\lambda \in (0,1)$ solving $\lambda(\hat{\phi})\hat{\phi} = \left[\int_0^{\hat{\phi}} \phi \; dG(\phi)\right] \Big/ \left[G(\hat{\phi})\right]$. If the conditional beliefs are

$$\mu^*_{NS} = E[\phi \mid \phi \leq \hat{\phi}] = \frac{\int_0^{\hat{\phi}} \phi \; dG(\phi)}{G(\hat{\phi})} \text{ and}$$

$$\mu^*_{S} = E[\phi \mid \phi > \hat{\phi}] = \frac{\int_{\hat{\phi}}^1 \phi \; dG(\phi)}{1 - G(\hat{\phi})},$$

then the value of $\hat{\phi}^*$ must satisfy $E\Pi^S(p^*_S; \hat{\phi}) = E\Pi^{NS}\big(p^*_{NS}(\mu_{NS}); \hat{\phi}\big)$. Recalling that $p^*_S = \hat{\theta}(p^*_{NS})$, then $E\Pi^S(p^*_S; \hat{\phi}) = E\Pi^{NS}\big(p^*_{NS}(\mu_{NS}); \hat{\phi}\big)$ implies:

$$\hat{\phi}(T-1)R(p^m) - c_S = \left(\frac{\lambda(\hat{\phi})\hat{\phi}\Sigma}{1 + \lambda(\hat{\phi})\hat{\phi}(\Sigma - 1)}\right)(1 + \hat{\phi}(T-1))R(p^m). \tag{A-8}$$

Multiplying both sides by $1 + \lambda(\hat{\phi})\hat{\phi}(\Sigma - 1)$ and rearranging gives

$$-\lambda(\hat{\phi})\hat{\phi}^2(T-1)R(p^m) + \hat{\phi}\big[\big(T - 1 - \lambda(\hat{\phi})\Sigma\big)R(p^m) - \lambda(\hat{\phi})c_S(\Sigma - 1)\big] - c_S = 0. \tag{A-9}$$

Additionally, the existence of a positive solution to Eq. (A-9) requires

$$\big((T-1) - \lambda(\hat{\phi})\Sigma\big)R(p^m) - \lambda(\hat{\phi})c_S(\Sigma - 1) > 0. \tag{A-10}$$

If (A-10) is not satisfied then both roots are less than 0 and there does not exist a set $\Phi \subseteq [0,1]$ such that the firm will provide samples for any $\phi \in \Phi$.

When the lower root is greater than 0, it is a cutoff point where, for any $\phi \leq \hat{\phi}$, the firm will not provide samples and for some $\phi > \hat{\phi}$ it will. The lower root is given as

$$\hat{\phi} = \frac{\tilde{\Omega} - \sqrt{\tilde{\Omega}^2 - 4\lambda(\hat{\phi})c_S(T-1)R(p^m)}}{\lambda(\hat{\phi})(T-1)R(p^m)}, \tag{A-11}$$

where $\tilde{\Omega} \equiv \big((T-1) - \lambda(\hat{\phi})\Sigma\big)R(p^m) + \lambda(\hat{\phi})c_S(\Sigma - 1)$.

Again the existence of a real root requires that the discriminant of (A-11) is non-negative; i.e., the existence of a real root requires

$$\big[\big((T-1) - \lambda(\hat{\phi})\Sigma\big)R(p^m) - \lambda(\hat{\phi})c_S(\Sigma-1)\big]^2 - 4\lambda(\hat{\phi})c_S(T-1)R(p^m) \geq 0. \tag{A-12}$$

Combining (A-10), and (A-12) provides the following sufficient condition for the existence of a sampling set $\Phi \subset [0,1]$:

$$\big((T-1) - \lambda(\hat{\phi})\Sigma\big)R(p^m) - \lambda(\hat{\phi})c_S(\Sigma-1) \geq 2\sqrt[+]{\lambda(\hat{\phi})c_S(T-1)R(p^m)}. \tag{A-13}$$

When condition (A-13) is not met, the firm will not find it optimal to provide samples for any efficacy in $[0,1]$. Consequently, the firm will not provide samples for any realization of the efficacy and consumers will hold the unconditional prior $\mu = E\phi = \int_0^1 \phi \; dG(\phi)$. The price the firm sets when it does not provide samples was derived in the text in Section 3.

It remains to be shown that when condition (A-13) $\hat{\phi}$ exists, the firm will provide samples for all $\phi > \hat{\phi}$. To see that it will, consider how the firm's profit changes with the efficacy of its drug. By the definition of $\hat{phi}$, $E\Pi^S(\hat{\phi}) = E\Pi^{NS}(\hat{\phi})$. However, unlike when information is symmetric, the no-sampling price is fixed at the expected efficacy given that the firm did not provide samples; i.e., at $\lambda(\hat{\phi})\hat{\phi}$ and p^*_{NS} can be expressed as

$$p^*_{NS} = \left(\frac{\lambda(\hat{\phi})\hat{\phi}\Sigma}{1 + \lambda(\hat{\phi})\hat{\phi}(\Sigma-1)}\right) p^*_S < p^*_S \tag{A-14}$$

Thus, it is clear that

$$\frac{dE\Pi^S(\phi)}{d\phi} \geq \frac{dE\Pi^{NS}(\phi)}{d\phi} \quad \forall \phi \geq \hat{\phi},$$

And it is optimal for the firm to provide samples for all $\phi > \hat{\phi}$.

Therefore, the firm is best-responding to the consumers' beliefs and the consumers' beliefs are consistent with the firm's actions. The values for p^*_S and p^*_{NS} were derived in the text. Finally, all of the preceding analysis holds for all $\lambda(\phi) \in (0,1)$.

It has been shown that the conditions of the proposition constitute an equilibrium. The uniqueness of the separating equilibrium is a consequence of the beliefs. If the specified beliefs are the unique set of consistent beliefs, then the equilibrium is unique because the sampling set Φ, derived from the beliefs is unique.

The following shows these are in fact the unique set of beliefs given the priors of the consumers. The cut-off value was set to the lower root of the quadratic because it was in $(0,1)$. However, the upper root may also be less than 1 for sufficiently high values of Σ, or conversely, low values of T and when $\phi > \overline{\phi}$ the firms profits are higher not providing samples. Taking into account when $\overline{\phi} < 1$, the belief regarding the drug's efficacy when the firm provides samples may alternatively be defined as

$$\mu_S = \frac{\int_{\underline{\phi}}^{\overline{\phi}} \phi \; dG(\phi)}{G(\overline{\phi}) - G(\underline{\phi})}.$$

In this way, the belief regarding the drug's efficacy when the firm does not provide samples is

$$\mu_{NS} = \frac{\int_0^{\underline{\phi}} \phi \; dG(\phi)}{G(\underline{\phi})} + \frac{\int_{\overline{\phi}}^1 \phi \; dG(\phi)}{1 - G(\overline{\phi})}.$$

Beliefs are not important for the sampling profit as the firm will set price equal to the full information price, but for the no-sampling profit prices are a function of beliefs. For beliefs to be consistent, a firm with a very high quality $\phi > \overline{\phi}$ should prefer to not provide samples. At $\phi = \overline{\phi}$ the profits are the same. Taking the derivative with respect to ϕ at $\phi = \overline{\phi}$ we get:

$$\frac{dE\Pi^S}{d\phi} = (T-1)R(p^m) > \frac{dE\Pi^{NS}}{d\phi} = \left(\frac{\mu_{NS}\Sigma}{1+\mu_{NS}(\Sigma-1)}\right)(T-1)R(p^m),$$

Therefore a firm with $\phi > \overline{\phi}$ will still prefer to provide samples and the beliefs are not consistent. Thus the sampling decision is monotonic in ϕ and $\hat{\phi}$ is the unique cut-off efficacy level.

PROOF OF PROPOSITION 4

There are two parts to the proposition. Let ϕ be the efficacy of the firm's drug and let Φ be the sampling set of efficacies.

1.) When $\phi \in \Phi$, then condition (22) of the proposition is a result of plugging in (12) into (21) and finding when the resulting expression is greater than 0.

2.) When $\phi \notin \Phi$, consumer welfare is still increased because the conditional equilibrium belief given the firm does not dispense samples is less than the unconditional equilibrium belief when the firm is not permitted to provide samples, $\mu_{NS} < \mu_0$, thus the optimal no sampling price is less when sampling is permitted.

PROOF OF PROPOSITION 5

Let ϕ be the efficacy of the firm's drug and let Φ be the sampling set of efficacies. If $\phi \in \Phi$, then the result is found by determing when the summation of Eqs. (21) and (23) are greater than 0.

PROOF THAT THIRD DEGREE PRICE DISCRIMINATION IS MORE PROFITABLE

To completely prove the statement that third degree price discrimination is more profitable than both sampling and setting a uniform price we must compare them when information is symmetric and asymmetric. The comparison between sampling and third degree price discrimination was done in the text so we need only compare third degree price discrimination with a uniform price.

When information is symmetric the comparison is straight-forward as there are no beliefs to consider. The profit comparison is as follows.

$$\begin{aligned}
E\Pi^{NS} &\gtrless E\Pi^{NS}_{PD} \\
p^*_{NS}\big(1+\phi(T-1)\big)\big(1-F(p^*_{NS})\big) &\gtrless \big(\phi p^m + \phi p^m(T-1)\big)\big(1-F(p^m)\big) \\
\left(\frac{\phi\Sigma}{1+\phi(\Sigma-1)}\right)\big(1+\phi(T-1)\big) &\gtrless \phi T \\
\phi\Sigma\big(1+\phi(T-1)\big) &\gtrless \phi T\big(1+\phi(\Sigma-1)\big) \\
\phi &\gtrless 1.
\end{aligned}$$

Thus third degree price discrimination is more profitable for all $\phi < 1$ when information is symmetric.

To compare a uniform price under with price discrimination when information is asymmetric let μ_{PD} represent the consumers' beliefs given the firm price discriminates and let μ_{UP} represent the consumers' beliefs given the firm charges a uniform spot price. To find when the firm will price discriminate we start by assuming there is some cutoff efficacy such that the firm will price

discriminate for all efficacies to one side of the cutoff and charge a uniform price for all efficacies on the other side of the cutoff. The profit comparison is as follows.

$$
\begin{aligned}
E\Pi_{UP}^{NS} &\gtrless E\Pi_{PD}^{NS} \\
p_{NS}^{*}(\mu_{UP})\big[1+\phi(T-1)\big)\big(1-F(p_{NS}^{*}(\mu_{UP}))\big] &\gtrless \big(\mu_{PD}p^{m}+\phi p^{m}(T-1)\big)\big(1-F(p^{m})\big) \\
\left(\frac{\mu_{UP}\Sigma}{1+\mu_{UP}(\Sigma-1)}\right)\big[1+\phi(T-1)\big] &\gtrless \big(\mu_{PD}+\phi(T-1)\big) \qquad \text{(A-15)}
\end{aligned}
$$

From (A-15) it is clear that if there are any efficacies with which the firm would prefer to price-discriminate, then they must be for higher efficacies since $\partial E\Pi_{PD}^{NS}/\partial\phi > \partial E\Pi_{UP}^{NS}/\partial\phi$. If the firm prefers to price discriminate at $\phi = 1$ then rational consumers will infer that when the firm is price discriminating the efficacy is higher than if it was not. Thus, $\mu_{PD} > \mu_{UP}$. However, if $\mu_{PD} > \mu_{UP}$, then

$$
\mu_{PD} > \mu_{UP} > \frac{\mu_{UP}\Sigma}{1+\mu_{UP}(\Sigma-1)}, \qquad \text{(A-16)}
$$

and the firm would prefer to price discriminate for all efficacies. Moreover, it follows that the firm will price discriminate when $\phi = 1$ because the profit is higher, thus third degree price discrimination is preferred over charging a uniform price for all efficacies $\phi \in [0, 1]$.

References

Arrow, Kenneth J., "Uncertainty and the Welfare Economics of Medical Care," *American Economic Review*, 1963, *53* (5), 941–973.

Backer, E.L., J.A. Van Tonder, and B.F. Crabtree, "The value of pharmaceutical representative visits and medication samples in community-based family practices," *Journal of Family Practice*, 2000, *49*, 811–816.

Bagwell, Kyle, "Introductory Price as a Signal of Cost in a Model of Repeat Business," *Review of Economic Studies*, 1987, *54*, 365–84.

_ **and Michael H. Riordan**, "High and Declining Prices Signal Product Quality," *The American Economic Review*, 1992, *81* (1), 224–239.

Band, Ron, "Product Sampling in a New Era of Cost Efficiencies," *Product Management Today*, 2007, *18* (8), 28–29.

Bergemann, Dirk and Juuso Välimäki, "Dynamic Pricing of new Experience Goods," *Journal of Political Economy*, 2006, *114* (4), 713–743.

Boltri, J. M., E. R. Gordon, and R. L. Vogel, "Effect of antihypertensive samples on phyisician prescribing habits," *Family Medicine*, 2002, *34* (10), 729–731.

Brennan, Troyen A., David J. Rothman, Linda Blank, David Blumenthal, Susan C. Chimonas, Jordan J. Cohen, Janlori Goldman, Jerome P. Kassirer, Harry Kimball, James Naughton, and Neil Smelser, "Health industry practices that create conflicts of interest: a policy proposal for academic medical centers," *Journal of the American Medical Association*, 2006, *295* (4), 429–433.

Chintagunta, Pradeep, Renna Jiang, and Giner Z. Jin, "Information, Learning, and Drug Diffusion: the Case of Cox-2 Inhibitors," June 2008. Under Review: Quantitative Marketing and Economics.

Cutrona, Sarah L., Steffie Woolhandler, Karen E. Lasser, David H. Bor, Danny McCormick, and David U. Himmelstein, "Characteristics of Recipients of Free Prescription Drug Samples; a Nationally Representative Analysis," *American Journal of Public Health*, 2008, *98* (2), 284–289.

Gönül, Fusun F., Franklin Carter, Elina Petrova, and Kannan Srinivasan, "Promotion of Prescription Drugs and Its Impact on Physicians Choice Behavior," *Journal of Marketing*, 2001, *65* (3), 79–90.

Greenland, Philip, "Time for the Medical Profession to Act," *Archives of Internal Medicine*, 2009, *169* (9), 829–831.

IMS Health, "National Prescription Audit," Technical Report, IMS Health 2008.

Joseph, Kissan and Murali Mantrala, "A model of the role of free drug samples in physicians' prescription decisions," *Marketing Letters*, 2009, *20*, 15–29.

Konrad, Kai A., "Competition for Prescription Drugs," 2002. Working Paper.

Liebeskind, Julia and Richard P. Rumelt, "Markets for Experience Goods with performance Uncertainty," *The RAND Journal of Economics*, 1989, *20* (4), 601–621.

Mizik, Natalie and Robert Jacobson, "Are Physicians "Easy Marks"? Quantifying the Effects of Detailing and Sampling on New Prescriptions," *Management Science*, 2004, *50* (12), 1704–1715.

Nelson, Philip, "Advertising as Information," *The Journal of Political Economy*, 1974, *82*, 729–754.

Perloff, Jaffrey M. and Steven Salop, "Equilibrium with Product Differentiation," *Review of Economic Studies*, 1985, *52* (1), 1.

Rabin, R. C., "Free drug samples? Bad idea, some say," *New York Times*, May 2007.

Shapiro, Carl, "Optimal Pricing of Experience Goods," *Bell Journal of Economics*, 1983, *14*, 497–507.

Sorensen, Alan T., "Equilibrium Price Dispersion in Retail Markets," *Journal of Political Economy*, 2000, *108* (4), 833–850.

Venkataraman, S. and S. Stremersch, "The debate on influencing doctor's decisions: Are drug charateristics the missing link?," *Management Science*, 2007, *53* (11), 1688–1701.

Westfall, J.M., J. McCabe, and R.A. Nicholas, "Personal Use of Drug Samples by Physicians and Office Staff," *Journal of the American Medical Association*, 1997, *278* (2), 141–143.