Sequential Screening

Pascal Courty

Universitat Pompeu Fabra

Li, Hao

University of Hong Kong

Current version: March 11, 1999

Abstract: We present a model of price discrimination where a monopolist faces consumers with unitary demands who learn their valuations over time. Consumers are privately informed at the time of contracting about valuation distribution, but they privately learn their actual valuations after contracting. The monopolist sequentially screens consumers with a menu of contracts: they first choose a contract and then choose the level of consumption according to the terms specified in the contract. A deterministic sequential mechanism is a menu of refund contracts, each consisting of an advance payment and a refund amount in case of no consumption, but general sequential mechanisms may also involve randomization. We characterize the optimal sequential mechanism both when some consumers are more eager than others in the sense of first-order stochastic dominance, and when some face greater valuation uncertainty than others in the sense of mean-preserving-spread. We show that it can be optimal to subsidize consumers with smaller valuation uncertainty through low refund in order to reduce the rent to those with greater uncertainty, who purchase more "flexible" contracts with greater refund. The size of distortion depends on how informative consumers' initial private knowledge is about their valuations from the monopolist's point of view, but not on the size of valuation uncertainty if it affects all consumers.

Acknowledgements: We received helpful comments from Michael Chwe, Jim Peck, Hugo Hopenhayn, Sherwin Rosen, Lars Stole, and Wing Suen. Mark Armstrong sent us a note on the two-type case under first-order stochastic dominance after we finished an earlier version of the paper. We are especially grateful for the suggestion of the Associate Editor and the referees that we investigate situations where first-order stochastic dominance does not hold.

1. Introduction

The mechanism design literature has shed light on many commonly used price discrimination schemes.[1] However, most models developed in this literature are static in that consumers are assumed to make consumption decisions at the same time as they select a contract. This assumption is not innocuous when consumers learn new information about their demand over time. Consider the demand for plane tickets. Travelers typically do not know their valuations for tickets until just before departure, but they know in advance their likelihood to have high and low valuations. A monopolist can wait until the travelers learn their valuations and charge the monopoly price, but more consumer surplus can be extracted by requiring them to reveal their private information sequentially. An illustration of such monopoly practice is a menu of refund contracts, each consisting of an advance payment and a refund amount in case the traveler decides not to use the ticket. By selecting a refund contract from the menu, travelers reveal their private information about the distribution of their valuations, and by deciding later whether they want the ticket or the specified refund, they reveal what they have learned about their actual valuation.

The following example of airplane ticket pricing illustrates sequential price discrimination.[2] Suppose that one-third of all potential buyers are leisure travelers whose valuation is uniformly distributed on $[1, 2]$, and two-thirds are business travelers whose valuation is uniformly distributed on $[0, 1] \cup [2, 3]$. Intuitively, business travelers face greater valuation uncertainty than leisure travelers. Suppose that cost of flying an additional traveler is 1. If the seller waits until travelers have privately learned their valuations, she faces a valuation distribution that is uniform on $[0, 3]$. The optimal monopoly price is 2 with expected profit of $\frac{1}{3}$, thus excluding all leisure travelers and as well as half of business travelers who turned out to have low valuations. Suppose instead the seller offers two contracts before the travelers learn their valuation, one with an advance payment of 1.5 and no refund and the other with an advance payment of 1.75 and a partial refund of 1. Leisure travelers strictly prefer the contract with no refund. Business travelers are indifferent between the

[1] One of the earliest contributions to this literature is Mussa and Rosen [1978]. Wilson [1993] gives an excellent account of applications to real-life pricing problems.

[2] The numbers of this example are chosen to expedite exposition. Later we will formally derive the optimal menu of refund contracts and use the example to illustrate some of main results of this paper.

two contracts so we assume that they choose the contract with refund. The monopolist separates the two types and earns an expected profit of $\frac{2}{3}$, doubling the profits of charging the monopoly price after travelers have learned their valuations.

This paper considers a class of monopolist sequential screening problems where consumers sequentially learn their demand and contracts are signed when consumers only have partial private information. Such pricing problems are not unique to airplane ticket pricing and refund contracts. Sequential mechanisms take different forms in hotel reservations (cancellation fees), car rentals (free mileage vs. fixed allowance), telephone pricing (calling plans), public transportation (day pass), and utility pricing (optional tariffs). Sequential price discrimination can also play a role in contracting problems such as taxation and procurement where the agent's private information is revealed sequentially.

Surprisingly, sequential screening has not received much direct attention in the screening literature. Although sequential mechanisms share the characteristic with two-part tariffs that consumption decisions are made sequentially, and there is an abundant literature on two-part tariffs (see, e.g., Wilson [1993]), the empirical importance of sequential mechanisms suggests that two-part tariffs are more than a simple way of implementing concave nonlinear tariffs. Moreover, sequential mechanisms have a learning feature that the typical textbook example of two-part tariff does not have: when consumers choose a two part-tariff they do not know the quantity they wish to consume or the valuation they place on the good. An implication of this learning feature is that consumers typically suffer from "regret" at the time of consumption: a businessman could have bought the same ticket at a lower advance price had he known that he would fly for sure, or a traveler could have avoided the cancellation fee had he known his itinerary when he reserved the hotel room.

The primary goal of this paper is to show that sequential mechanisms help producers to price discriminate when consumers learn private information about their demand over time. Although sequential mechanisms can take different forms, we restrict our attention to situations where consumers have unit demands as in the airplane ticket pricing problem. In these situations, optimal *ex post* pricing scheme (after consumers have complete private information about their demand) degenerates to standard monopolist pricing. This allows us to focus on the effects of consumer learning on sequential price discrimination. When consumers have unit demand, the monopolist price-discriminates only by choosing the probability that he delivers the good. Although refund contracts constrain the delivery probabilities to zero or one, a general sequential mechanism is a menu of contracts consisting of pairs of delivery probability and payment to the monopolist. Efficiency is achieved by delivering the good if and only if the consumer's valuation exceeds the production cost, but the optimal sequential mechanism can generate either downward or upward distortions. Downward distortions in sequential mechanisms (i.e. not delivering the good when the consumer's valuation is greater than the cost) are similar to the standard "rationing" result in nonlinear pricing models that under-provision of quality or quantity is used to extract more surplus from more eager consumers (see, e.g., Mussa and Rosen [1978], or Maskin and Riley [1984].) More surprisingly, it can be optimal to "subsidize" consumption by some consumers (i.e. providing the good when the consumer's valuation is below the cost). Inefficient over-production typically arises only in multi-product price discrimination problems when better separation is achieved by subsidizing some goods (see, e.g., Adam and Yellen [1976], or Rochet and Chone [1998].)[3] Although there is only one product in our problem, inefficient over-production can be used effectively as a price discrimination instrument when the production cost is relatively low and consumers differ sufficiently in the degree of valuation uncertainty they face.

In section 2, we consider the problem of designing the optimal menu of refund contracts for two *ex ante* types of potential buyers. To continue with the airplane ticket pricing example, the business traveler is either an eager consumer who is more likely to draw higher valuations than the leisure traveler, in the sense of first-order stochastic dominance, or a consumer who faces greater valuation uncertainty than the leisure traveler, in the sense of mean-preserving-spread (Rothschild and Stiglitz [1970]). In either case, we show that there is no consumption distortion for the business type in the optimal menu of refund contracts. In the case of first-order stochastic dominance, rationing the leisure type is the optimal way of reducing the rent to the business type. In the case of mean-preserving-spread, subsidy as well as rationing can be optimal. Sufficient conditions are provided such that when the production cost is low, subsidizing the leisure type with a refund lower than the cost of the ticket is cost-effective in reducing the rent to the business type. For airplane ticket pricing, the marginal cost is low when capacity constraint is not binding, so

[3] Over-production can also occur in the single-product monopoly models when reservation utility in participation constraints varies with the type of agents. See Lewis and Sappington [1989].

our result that the business type purchases a contract with a higher refund explains why it can be optimal for airlines to offer business travelers more "flexible" contracts.

Section 3 examines the general problem of sequential price discrimination with continuous types. This generalization enables us to discuss how type distribution and consumer learning affect the design of sequential mechanisms. We characterize the optimal sequential mechanism for a case where consumers face the same valuation uncertainty but differ in expected valuation, and a case where consumers have the same expected valuation but differ in valuation uncertainty. In both cases, the delivery rule is shown to be deterministic, and can therefore be interpreted as a menu of refund contracts, or a two-part tariff. The size of distortion depends on the informativeness of consumers' initial private knowledge about their valuations from the monopolist's point of view, as well as on the type distribution. Distortions are small if consumers' initial private knowledge is not informative in that valuation distributions conditional on types do not vary much across different types of consumers. The size of distortion does not depend on any additional valuation uncertainty if it affects all consumers. For example, adding a valuation shock to individual uncertainty that affects all consumers independently does not change the optimal menu of contracts. In the first case, where consumers differ in their expected valuation, consumers with greater expected valuations are less likely to be rationed and choose the refund contract with lower advance payment and lower refund. In the second case, where consumers differ in valuation uncertainty they face, types facing smaller valuation uncertainty have larger consumption distortions. As in section 2, distortions can be either rationing or subsidy, and the latter is optimal when production cost is low.

Section 4 offers further comments on sequential screening. Neither first-order dominance nor mean-preserving-spread is sufficient for reducing the dimension of the design problem. This insufficiency results from the multi-dimensional nature of the sequential mechanism design problem, which also makes it possible that random delivery rules are optimal. We use a two-type example of airplane ticket pricing to illustrate how random delivery rules can fine tune sequential screening when business travelers face greater valuation uncertainty but have a smaller mean valuation than leisure travelers. We show how a tension arises under refund contracts between exploiting the high mean valuation of the leisure type and exploiting the fat tail of the business type, and how randomization can help fine tune sequential discrimination. Section 5 concludes with some discussion on related works and future direction of this line of research.

2. Optimal Menu of refund contracts: Two-Type Case

Consider a monopoly seller of airplane tickets facing two types of travelers, B and L, with proportion f_B and f_L respectively. There are two periods. In the beginning of period one, the traveler privately learns his type which determines the probability distribution of his valuation for the ticket. The seller and the traveler contract at the end of period one. In the beginning of period two, the traveler privately learns his actual valuation v for the ticket, and then travelling may take place. Each ticket costs the seller c. The seller and the traveler are risk-neutral, and do not discount. Throughout this section, we will think of type B as the "business type," which values the ticket more or faces greater valuation uncertainty; type L is a "leisure" traveler. Greater valuation of the business type is captured by first-order stochastic dominance (FSD). The valuation distribution G_B of type B first-order stochastically dominates G_L of the leisure type if $G_B(v) \leq G_L(v)$ for all v in the range of valuations $[\underline{v}, \overline{v}]$. Greater valuation uncertainty of the business type is represented by mean-preserving-spread (MPS, Rothschild and Stiglitz [1970]). The valuation distribution G_B dominates G_L by MPS if they have the same mean and $\int_{\underline{v}}^{v}(G_B(u) - G_L(u))du \geq 0$ for all $v \in [\underline{v}, \overline{v}]$. Let g_B and g_L be the density functions of the two types.

A refund contract consists of an advance payment a at the end of period one and a refund k that can be claimed at the end of period two after the traveler learns his valuation. Clearly, regardless of the payment a, the consumer travels only if he values the ticket more than k. The seller offers two refund contracts $\{a_B, k_B, a_L, k_L\}$. The profit maximization problem can be written as:

$$\max_{k_L, k_B, a_L, a_B} \sum_{t=L,B} f_t(a_t - G_t(k_t)k_t - (1 - G_t(k_t))c)$$

subject to

$$(IR_t) \qquad \forall t = L, B, \ -a_t + k_t G_t(k_t) + \int_{k_t}^{\overline{v}} v dG_t(v) \geq 0,$$

$$(IC_{t,t'}) \qquad \forall t \neq t', \ -a_t + k_t G_t(k_t) + \int_{k_t}^{\overline{v}} v dG_t(v) \geq -a_{t'} + k_{t'} G_t(k_{t'}) + \int_{k_{t'}}^{\overline{v}} v dG_t(v).$$

The first set of constraint (IR) is the individual rationality constraints in period one. The second set (IC) is the incentive compatibility constraints in period one.

Lemma 2.1. *Under either FSD or MPS, IR_L and $IC_{B,L}$ imply IR_B.*

Proof. The two individual rationality constraints can be rewritten as:

$$\forall t = L, B, \; -a_t + \int_{\underline{v}}^{\overline{v}} \max\{k_t, v\} dG_t(v) \geq 0.$$

Then, $IC_{B,L}$ implies

$$-a_B + \int_{\underline{v}}^{\overline{v}} \max\{k_B, v\} dG_B(v) \geq -a_L + \int_{\underline{v}}^{\overline{v}} \max\{k_L, v\} dG_B(v).$$

Since $\max\{k_L, v\}$ is an increasing function of v, if G_B dominates G_L by FSD,

$$\int_{\underline{v}}^{\overline{v}} \max\{k_L, v\} dG_B(v) \geq \int_{\underline{v}}^{\overline{v}} \max\{k_L, v\} dG_L(v).$$

Since $\max\{k_L, v\}$ is a convex function of v, the above condition holds also if $G_B(v)$ dominates $G_L(v)$ by MPS. The lemma then follows from IR_L. *Q.E.D.*

Thus, the business type gets more utility than the leisure type from any refund contract, whether it is defined by greater valuation or by greater uncertainty. Indeed, we can define the business type by combining FSD and MPS. For example, take a distribution G_B that dominates G_L by MPS. Shifting the whole distribution G_B to the right gives a new distribution that has both greater average valuation and greater valuation uncertainty. It is easy to see that the above lemma continues to hold for this combination of FSD and MPS. Note that Lemma 2.1 does not hold under general second-order stochastic dominance, i.e. greater dispersion without the restriction of the same mean. This can be seen from the proof of the lemma. Under general second-order stochastic dominance the integration of the function $\max\{k, v\}$ over $[\underline{v}, \overline{v}]$ can be either greater or smaller for a dominant distribution. An example where the business type faces greater valuation uncertainty but has a smaller mean than the leisure type will be analyzed in section 4.

Lemma 2.1 implies that IR_L binds (holds with equality) in the optimal menu of refund contracts, otherwise increasing both a_L and a_B by the same amount would increase profits. Also, $IC_{B,L}$ binds in the optimal menu of refund contracts, otherwise profits could be increased by increasing a_B. Substituting IR_L and $IC_{B,L}$ into the objective function and ignoring $IC_{L,B}$, we obtain the following "relaxed" problem:

$$\max_{k_L, k_B} \int_{k_L}^{\overline{v}} (f_L(v-c)g_L(v) - f_B(G_L(v) - G_B(v)))dv + \int_{k_B}^{\overline{v}} (f_B(v-c)g_B(v))dv.$$

Let $S(k_L) = \int_{k_L}^{\overline{v}}(v-c)g_L(v)dv$ be the surplus from the leisure type, and $R(k_L) = \int_{k_L}^{\overline{v}}(G_L(v)-G_B(v))dv$ be the rent to the business type, both as function of the refund to the leisure type. Since the choice of refund for the business type is unconstrained, it should be equal to c to maximize the surplus from the business type. Refund for the leisure type should be chosen to maximize the surplus from the leisure type less the rent to the business type, so the solution is given by $\text{argmax}_k f_L S(k) - f_B R(k)$. Now we are ready to state the first main result of this paper.

PROPOSITION 2.2. *Under either FSD or MPS, in the optimal menu of refund contracts,* $k_B = c$ *and* $k_L = \text{argmax}_k f_L S(k) - f_B R(k)$.

PROOF. It suffices to show that the "upward" constraint $IC_{L,B}$ is satisfied by the given $\{k_B, k_L\}$. Since $IC_{B,L}$ binds, we have

$$a_L - a_B = \int_{k_B}^{k_L} G_B(v)dv,$$

which implies

$$\begin{aligned}&-a_L + k_L G_L(k_L) + \int_{k_L}^{\overline{v}} v dG_L(v)\\ =&-a_B + k_B G_L(k_B) + \int_{k_B}^{\overline{v}} v dG_L(v) - \int_{k_B}^{k_L}(G_B(v)-G_L(v))dv.\end{aligned}$$

Thus, $IC_{L,B}$ is satisfied if and only if $\int_{k_B}^{k_L}(G_B(v)-G_L(v))dv \leq 0$. If the solution to the relaxed problem has $k_L = c$, the proposition follows immediately. Suppose that the solution has $k_L \neq c$, and $\int_c^{k_L}(G_B(v)-G_L(v))dv > 0$. Consider an alternative menu where the two types have the same refund c. The surplus $S(c)$ from type L is greater in the alternative menu. The rent to type B is smaller in the alternative, because

$$R(c) = \int_{k_L}^{\overline{v}}(G_L(v)-G_B(v))dv + \int_c^{k_L}(G_L(v)-G_B(v))dv < \int_{k_L}^{\overline{v}}(G_L(v)-G_B(v))dv.$$

This contradicts the assumption that $\{k_L, k_B\}$ solves the relaxed problem. Thus, the solution to the relaxed problem satisfies $IC_{L,B}$. *Q.E.D.*

The derivation of the optimal menu of refund contracts is completed by finding the advance payments a_L and a_B from IR_L and $IC_{B,L}$. Proposition 2.2 shows there is no consumption distortion for the business type, either when it's defined by FSD or MPS, or a combination of the two as described previously. Under FSD, the rent function $R(k_L)$ is decreasing for any k_L. Since the surplus $S(k_L)$ from type L is increasing for any $k_L < c$, the solution to the relaxed problem has $k_L \geq c$. Thus, we have the standard result that there is rationing for the leisure type to lower the rent given away to the business type. Under MPS, the rent $R(k_L)$ is not a monotonically decreasing function of the refund k_L to the leisure type. Instead, $R(k_L)$ is zero at both $\underline{v}$ and $\overline{v}$, and tends to be greater in the middle of the support. In this case, subsidy as well as rationing can be used to reduce the rent to the business type. In order to obtain more insights about the nature of consumption distortion for the leisure type, we need to impose additional restrictions on top of dominance by MPS. Suppose that the rent function $R(\cdot)$ is single-peaked at some $z \in (\underline{v}, \overline{v})$. This is satisfied if for example G_B and G_L differ by a single mean-preserving-spread (Rothschild and Stiglitz [1970]). For simplicity, let's assume that there is no "plateau" at z so that $G_B(v) > G_L(v)$ for all $v < z$ and $G_B(v) < G_L(v)$ for all $v > z$. An example of this is normal distributions with the same mean z and a greater variance for type B.[4]

Under the assumption of single peak, the rent to the business type is the greatest when the refund for the leisure type equals the peak of the distributions, and it falls monotonically on either side of the peak. Whether it is optimal to subsidize (set $k_L < c$) or ration ($k_L > c$) the leisure type depends on how the loss of surplus due to distortions compares with rent reduction. Note that when $R(\cdot)$ is single-peaked at some $z \neq c$, the optimal refund to the leisure type k_L cannot lie between c and z. For example, if $c < z$, then any $k_L \in [c, z]$ cannot be optimal because by decreasing k_L toward c the seller could increase the surplus $S(k_L)$ and decrease the rent $R(k_L)$. Furthermore, setting $k_L = c$ cannot be optimal, since by decreasing k_L slightly below c, surplus from type L is not affected (because $S'(c) = 0$) but the rent to type B decreases. With the values between c and z excluded, intuition suggests that it is optimal to subsidize consumption when the cost is low, and to ration when the cost is high. The reason is that when the cost is below the peak of the rent function, rationing is too costly because it prevents many profitable trades, while when the cost is above the peak, subsidy means too many inefficient trades. The

[4] If U is a random variable with log-concave density function, and V has zero mean and is independent of U, then the distribution functions of $U + V$ and U have the above desired properties. See Shaked and Shanthikumar [1994].

following two results give sufficient conditions under which such patterns of distortions are optimal. The first one assumes symmetry of the density functions; the second one assumes that the proportion of business travelers is sufficiently small and/or the cost is sufficiently different from the peak of the rent function.

PROPOSITION 2.3. *It is optimal to subsidize (ration) the low type when $c < z$ ($c > z$) if at least one of the following two conditions is satisfied: (i) g_B and g_L are symmetric around z; (ii) $f_B R(c) \leq f_L(S(c) - S(z))$.*

PROOF. (i) Suppose that $c < z < k_L$. Since g_B and g_L are symmetric around z, the rent function $R(\cdot)$ is also symmetric around z. If $z < k_L \leq\!< 2z - c$, an alternative menu with $\tilde{k}_L = 2z - k_L$ yields a greater surplus because $c \leq \tilde{k}_L < k_L$, and the same rent by symmetry, a contradiction. Suppose $k_L > 2z - c$. Comparing the slope of $S(\cdot)$ at any $k > 2z - c$ and at $\tilde{k} = 2z - k$, we have

$$-S'(k) = (k - c)g_L(k) = (k - c)g_L(\tilde{k}) > (c - \tilde{k})g_L(\tilde{k}) = S'(\tilde{k}).$$

It follows that $S(\tilde{k}_L) > S(k_L)$ for any $k_L > 2z - c$ and $\tilde{k}_L = 2z - k_L$. Since $R(\tilde{k}_L) = R(k_L)$, the alternative menu yields a greater profit, a contradiction. The proposition then follows from the fact that k_L cannot lie between c and z. The argument is similar when $c > z > k_L$.

(ii) Suppose that $c < z$. Since k_L cannot lie between c and z, either $k_L < c$ or $k_L > z$. The profit of setting $k_L < c$ is at least as great as $f_L S(c) - f_B R(c)$, since choosing k_L slightly below c always reduces the rent without changing the surplus at the margin. On the other hand, the profit of setting $k_L > z$ is at best as great as $f_L S(z)$, with maximum surplus and zero rent. If $f_B R(c) \leq f_L(S(c) - S(z))$, setting $k_L > z$ cannot be optimal. A similar argument holds when $c > z$. *Q.E.D.*

Thus, according to Proposition 2.3, the pattern of consumption distortion is determined by the comparison between the cost of the ticket and the peak of the distributions. When neither condition in Proposition 2.3 holds, the pattern of consumption distortions for the leisure type can be different from the predictions of the proposition. For example, if $c < z$, rationing instead of subsidy for the leisure type can be an optimal way of reducing rent to the business type. However, the case in which leisure travelers are subsidized corresponds to the familiar ticketing pattern where leisure travelers buy discount tickets with lower or even zero refund while business travelers pay more for flexibility. The conclusion of Proposition 2.3 shows that greater uncertainty for business travelers, instead of greater likelihood to have high valuations, is the key to understanding this pattern in terms of monopoly pricing.

To illustrate, let us reconsider the example presented in the introduction. By definition, the surplus function is

$$S(k) = \begin{cases} 1/2 & \text{if} \quad k \in [0,1) \\ k - \frac{1}{2}k^2 & \text{if} \quad k \in [1,2] \\ 0 & \text{if} \quad k \in (2,3], \end{cases}$$

and the rent function is

$$R(k) = \begin{cases} \frac{1}{4}k^2 & \text{if} \quad k \in [0,1) \\ \frac{1}{4} - \frac{1}{2}(k-1)(k-2) & \text{if} \quad k \in [1,2] \\ \frac{1}{4}(k-3)^2 & \text{if} \quad k \in (2,3]. \end{cases}$$

According to Proposition 2.2, the business type buys a contract with refund $k_B = 1$ and the leisure type buys a contract with refund $k_L = \text{argmax}_k \frac{1}{3}S(k) - \frac{2}{3}R(k)$. Figure 1 plots $\frac{1}{3}S(k)$ and $\frac{2}{3}R(k)$. From Figure 1, we find that the optimal refund for the leisure type is $k_L = 0$. Note that the rent function is single-peaked at $z = 1.5$. This example satisfies the first condition in Proposition 2.3 because the density functions are symmetric around 1.5. Recall that the production cost is 1. Since $c < z$, Proposition 2.3 predicts subsidy for the leisure type. Indeed, the optimal refund for the leisure type is 0. But since the leisure type's valuation exceeds the cost with probability one, subsidy never actually occurs. Optimal sequential contracts can be easily determined. Since the leisure type's participation constraint IR_L binds, from $k_L = 0$ we have $a_L = 1.5$. Since $k_B = 1$, from $IC_{B,L}$ we have $a_B = 1.75$. This verifies the optimality of the refund contracts given in the introduction.

A few remarks are in order. First, in the optimal sequential contract, even business type gets only partial refund. Since a sequential contract with full refund is formally identical to buying from the spot market at a price equal the refund, this example shows that it is generally not optimal for a monopoly seller to offer only advance purchase discount and induce the business type to buy from the spot market.[5] Second, business type gets

[5] Dana [1998] shows that consumers with more certain demand buy non-refundable advance tickets. His result is based on the use of advance purchase discount in choosing production capacity by competitive sellers. In his model, ticketing strategies are restricted to very simple forms, and consumers facing greater demand uncertainty wait for the spot market instead of purchasing refund contracts as in our model.

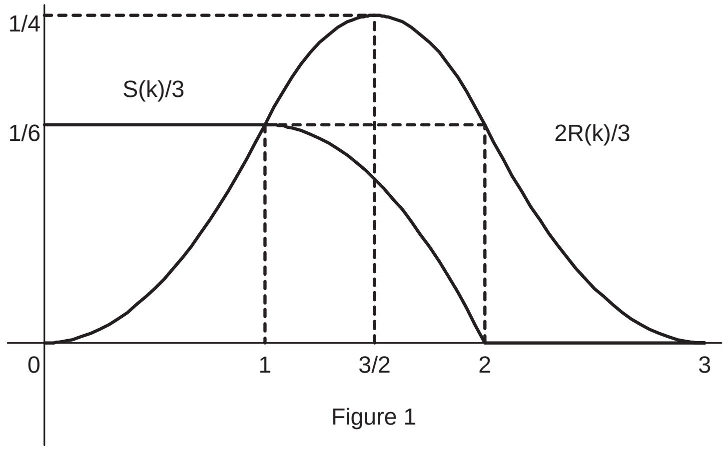

Figure 1

zero surplus from the optimal sequential contract. This is a special result, because by Lemma 2.1, business type should get a greater surplus than leisure type from any refund contract. What happens is that leisure type's contract has zero refund, and since the two types have the same mean, business type gets the same zero expected surplus from the zero-refund contract of leisure type.[6] Since $IC_{B,L}$ binds, business type gets zero expected surplus from its partial-refund contract. Finally, in this example, the first-best outcome is achieved by sequential screening, but this result is also special. If leisure type has a small probability of reaching valuations on the interval $[0, 1]$, then as long as such probability is small enough so that $\frac{1}{3}S(k) - \frac{2}{3}R(k)$ still reaches maximum at $k = 0$, zero-refund for leisure type continues to be optimal. But in this case, there is over-production for leisure type, and efficiency of sequential screening disappears.

3. Sequential Mechanism Design: Continuous Type Case

The analysis in the last section illustrates some general characteristics of sequential price discrimination. In this section, we show that these characteristics carry through in the absence of the restriction to two *ex ante* types and the restriction to menus of refund contracts. Moreover, we will discuss the issue of how type distribution and consumer learning affect the design of sequential mechanisms, which cannot be done satisfactorily in the two-type case. Finally, by solving for optimal sequential mechanisms for a number of simple

[6] Technically, in the proof of Lemma 2.1, the function $\max\{k_L, v\}$ is not strictly convex.

and intuitive parameterizations, we take a first step toward testing implications of the sequential screening model. Readers mostly interested in applications to price discrimination issues may skip the technical treatment of the continuous type case and move directly to after the proof of Lemma 3.4 below.

In this section, we assume that types are continuously distributed over $T = [\underline{t}, \overline{t}]$, with a density function $f(t)$ and cumulative function $F(t)$. Each type t is represented by a distribution of valuations over $[\underline{v}, \overline{v}]$, with a differentiable density function $g(v|t)$ and cumulative function $G(v|t)$. Type information is known only to the consumer. Note that we have assumed that the type space T is one-dimensional for simplicity, but this does not reduce the complexity of the type space, because each type is a probability distribution and can vary in arbitrary ways. In the applications later in this section, private type information will be about the expected valuation or the degree of valuation uncertainty.

As in the standard mechanism problem, the revelation principle (see, e.g., Myerson [1979], Harris and Townsend [1981]) allows us to take a first step toward simplifying the problem of sequential mechanism design. We assume that the conditional distributions $g(v|t)$ have the same support for all $t \in T$. This assumption makes it simpler to write down the incentive compatibility constraints in the optimization problem.[7] For each pair of reports t and v, let $y(t,v)$ be the probability of delivery and $x(t,v)$ be the payment to the monopolist. The monopolist solves the following sequential mechanism design problem:

$$\max_{x(t,v),y(t,v)} \int_{\underline{t}}^{\overline{t}} \int_{\underline{v}}^{\overline{v}} f(t)(x(t,v) - cy(t,v))g(v|t)dvdt$$

subject to constraints:

$$(IC_2) \qquad \forall t, \forall v, v', \quad vy(t,v) - x(t,v) \geq vy(t,v') - x(t,v'),$$

$$(IC_1) \qquad \forall t, t', \quad \int_{\underline{v}}^{\overline{v}} (vy(t,v) - x(t,v))g(v|t)dv \geq \int_{\underline{v}}^{\overline{v}} (vy(t',v) - x(t',v))g(v|t)dv,$$

$$(IR) \qquad \forall t, \quad \int_{\underline{v}}^{\overline{v}} (vy(t,v) - x(t,v))g(v|t)dv \geq 0,$$

[7] The optimization problem does not get more complicated without the assumption of common support as long as supports of different types overlap sufficiently. More precisely, the condition is: for any type t and any two valuations v and v', there is a type t' (possibly t itself) such that v and v' are in the support of type t'. If this condition holds, the optimization problem has incentive compatibility constraints for each type involving all valuations in the union of the all supports.

$$(R) \qquad \forall t, \forall v, \quad 0 \leq y(t,v) \leq 1.$$

The first set of constraint (IC_2) is the incentive compatibility constraints in period two. The second set (IC_1) is the incentive compatibility constraints in period one. The third set (IR) is the individual rationality constraints in period one. The last set of constraint (R) requires the delivery rule to be feasible.[8]

Following the standard treatment of incentive compatibility constraints (see, e.g., Mirrlees [1971]), we can eliminate most of the period-two incentive constraints. Define $u(t,v) = vy(t,v) - x(t,v)$ to be the consumer's *ex post* surplus after he truthfully reports t and then v. The following lemma shows that when the consumer draws a greater valuation, he receives the good with a greater probability and has a greater consumer surplus. It also amounts to "localization" of IC_2 constraints. In searching for the optimal sequential mechanisms, we need only impose local constraints on the sequential mechanisms to ensure that all IC_2 constraints are satisfied. The proof the lemma is standard and therefore skipped (see, e.g., Stole [1996]).[9]

LEMMA 3.1. *The period-two incentive compatibility constraints are satisfied if and only if (i)* $\frac{\partial u(t,v)}{\partial v} = y(t,v)$, *and (ii)* $y(t,v)$ *is non-decreasing in* v *for each* t.

Our sequential mechanism design problem is related to static multi-dimensional price discrimination models (e.g., McAfee and McMillan [1988]). In these problems, the consumer is screened only once but he generally has more than one piece of private information (e.g., willingness to pay for two different goods), and the monopolist generally has more than one instrument of price discrimination (e.g., quantities of the two goods sold to the consumer). In our sequential mechanism design problem, the consumer is screened

[8] Note that there is no period-two individual rationality constraint $vy(t,v) - x(t,v) \geq 0$ for all t and v. This corresponds to situations where up-front deposits are not fully refundable or there are cancellation fees at the consumption date. The absence of this constraint is important for our results. Within the class of deterministic sequential mechanisms (menus of refund contracts), the *ex post* participation constraint implies that in each refund contract the advance payment does not exceed the refund. One can show that the monopolist cannot use the combination of advance payment and refund to price discriminate, and therefore all types have the same contract. The menu of refund contract then coincides with the *ex post* monopolist pricing. This conclusion does not hold if the monopolist is not restricted to deterministic mechanisms, but the presence of the *ex post* participation constraint clearly reduces the monopolist's discriminatory power.

[9] We consider only sequential mechanisms with piece-wise differentiable delivery rule $y(t,v)$.

twice, but since the contract is signed in the first period, we can think of the sequential design problem as a static problem in the first period, where the consumer chooses a contingent package of delivery probabilities and transfer payments. This static problem is multi-dimensional in a sense because, although the consumer has one piece of private information, the monopolist has many discrimination instruments in contingent packages of delivery probabilities and transfer payments. One difference between our problem and the static multi-dimensional problems is that in our problem the second-period screening imposes IC_2 constraints on the instruments that the monopolist can use, as stated in Lemma 3.1, whereas in the static multi-dimensional problems, there is no such *a priori* constraint.

With the interpretation of our sequential mechanism design problem as a static screening problem, it becomes natural to "localize" period-one incentive compatibility constraints as in Lemma 3.1. Define $U(t) = \int_{\underline{v}}^{\overline{v}} u(t,v)g(v|t)dv$ as the expected surplus of consumer of type t and $Y(t,v) = \int_{\underline{v}}^{v} y(t,u)du$ as the cumulative delivery probability.[10]

LEMMA 3.2. *The period-one incentive compatibility constraints are satisfied only if (i)* $\frac{dU(t)}{dt} = -\int_{\underline{v}}^{\overline{v}} y(t,v)\frac{\partial G(v|t)}{\partial t}dv$; *and (ii)* $\int_{\underline{v}}^{\overline{v}} \frac{\partial Y(t,v)}{\partial t}\frac{\partial g(v|t)}{\partial t}dv \geq 0$.

PROOF. By the period-one incentive compatibility constraint,

$$U(t') \geq U(t) + \int_{\underline{v}}^{\overline{v}} (g(v|t') - g(v|t))(vy(t,v) - x(t,v))dv.$$

Exchanging the roles of t and t', we have

$$U(t') - U(t) \leq \int_{\underline{v}}^{\overline{v}} (g(v|t') - g(v|t))(vy(t',v) - x(t',v))dv.$$

To obtain (i), we combine the above two inequalities, divide them by $t' - t$ (assuming $t' > t$), and let t' converge to t. Then,

$$\frac{dU(t)}{dt} = \int_{\underline{v}}^{\overline{v}} \frac{\partial g(v|t)}{\partial t}u(t,v)dv = -\int_{\underline{v}}^{\overline{v}} \frac{\partial G(v|t)}{\partial t}y(t,v)dv,$$

[10] The reason to use the cumulative delivery probability is that optimal sequential mechanisms often have piece-wise constant delivery rules, in which case first-period second-order condition written in derivatives of the delivery probability does not capture the restrictions imposed by the incentive compatibility constraints.

where the last equality uses Lemma 3.1 and integration by parts. Condition (ii) can be obtained similarly by combining the two inequalities, dividing them by $(t-t')^2$ and letting t' converge to t. *Q.E.D.*

Lemma 3.2 parallels Lemma 3.1. The first condition is a local period-one first-order condition (FOC_1), counterpart to the local period-two first-order condition (FOC_2) in Lemma 3.1 that $\frac{\partial u(t,v)}{\partial v} = y(t,v)$; the second condition is a local period-one second-order condition (SOC_1), counterpart to the local period-two second-order condition (SOC_2) in Lemma 3.1 that $y(t,v)$ is non-decreasing in v. However, the two lemmas differ on an important point: the two local conditions in Lemma 3.2 are necessary but not sufficient for IC_1, whereas the two conditions in Lemma 3.1 are both necessary and sufficient for IC_2. Section 4 comments on the difference.

It is well-known that multi-product price discrimination problems are complex when consumers' private information is multi-dimensional (see Armstrong [1996] and Rochet and Chone [1998]). In our model, the consumer's private information is a probability distribution and in general can vary quite arbitrarily. Little can be said about the properties of optimal mechanism without making further assumptions. Since each type is a probability distribution on $[\underline{v}, \overline{v}]$, one natural way of imposing a structure on T is through FSD. In this case, we say that type t is "higher" than t' if $G(v|t) \leq G(v|t')$ for all v, and that T is ordered by FSD if $t > t'$ implies that t is higher than t' for any $t, t' \in T$. Another way to impose a structure on T is through a particular kind of mean-preserving spread where all distributions $G(v|t)$ cross at a single point z. In this case, we say that type t is "higher" than t' if $G(v|t) \geq G(v|t')$ for all $v < z$ and $G(v|t) \leq G(v|t')$ for all $v > z$, and that T is ordered by MPS if $t > t'$ implies that t is higher than t' for any $t, t' \in T$. As in the two-type case, the analyses of these two cases will be similar.

Following the standard practice of mechanism design, we obtain a "relaxed" problem by imposing the two local first-order conditions in Lemma 3.1 and Lemma 3.2 while ignoring the second order conditions (and all but the lowest type IR constraint). By Lemma 3.2, we have

$$\int_{\underline{t}}^{\overline{t}} f(t)U(t)dt = U(\underline{t}) - \int_{\underline{t}}^{\overline{t}} \int_{\underline{v}}^{\overline{v}} (1-F(t))y(t,v)\frac{\partial G(v|t)}{\partial t} dv dt.$$

Define

$$\phi(t,v) = v - c + \frac{(1-F(t))}{f(t)} \frac{\partial G(v|t)/\partial t}{g(v|t)}.$$

The relaxed problem can be then written as $\max_{y(t,v)} \int_{\underline{t}}^{\overline{t}} \int_{\underline{v}}^{\overline{v}} \phi(t,v) y(t,v) g(v|t) f(t) dv dt$ subject to $0 \leq y(t,v) \leq 1$. The solution to the relaxed problem is given by $y(t,v) = 1$ for t and v such that $\phi(t,v) > 0$ and 0 otherwise. There is no randomization.

If transfer payments $x(t,v)$ can be found so that the solution $y(t,v)$ to the relaxed problem given above satisfies all IC_1 and IC_2 constraints, then the sequential mechanism $\{y(t,v), x(t,v)\}$ is optimal. But since we have ignored the local second-order conditions in Lemma 3.1 and Lemma 3.2, and since the two conditions in Lemma 3.2 are generally insufficient for IC_1, we need to impose some condition on $y(t,v)$. The next result states that in the case of FSD, if the solution $y(t,v)$ to the relaxed problem is monotone in both t and v, then transfer payments $x(t,v)$ can be found such that the sequential mechanism $\{y(t,v), x(t,v)\}$ solves the original problem.[11]

Lemma 3.3. *Suppose that T is ordered by FSD. If a delivery rule $y(t,v)$ solves the relaxed problem, and if $y(t,v)$ is non-decreasing in t for all v and in v for all t, then there exist transfer payments $x(t,v)$ such that the sequential mechanism $\{y(t,v), x(t,v)\}$ is optimal.*

Proof. Since it solves the relaxed problem, $y(t,v)$ is either 1 or 0 for any t and v. By assumption, $y(t,v)$ is non-decreasing in v for each t, so SOC_2 implies that there exists $k(t)$ for each t such that $y(t,v) = 0$ if $v \leq k(t)$ and $y(t,v) = 1$ if $v > k(t)$. By FOC_2, the transfer payments can be written as $x(t,v) = x_0(t)$ if $v \leq k(t)$ and $x(t,v) = x_1(t)$ if $v > k(t)$, with $k(t) = x_1(t) - x_0(t)$. By Lemma 3.1, all IC_2 constraints are satisfied.

The expected surplus of a type t consumer is

$$U(t) = -x_0(t) + \int_{k(t)}^{\overline{v}} (1 - G(v|t)) dv.$$

Taking derivatives and using FOC_1, we obtain

$$-\frac{dx_0(t)}{dt} - \frac{dk(t)}{dt}(1 - G(k(t)|t)) = 0.$$

[11] Matthews and Moore [1987] make the same observation in a multi-dimensional screening problem, and call such mechanisms "attribute-ordered."

The above condition gives a differential equation that can be used to find the function $x_0(t)$, with the boundary condition that $x_0(\underline{t})$ satisfies

$$U(\underline{t}) = -x_0(\underline{t}) + \int_{k(\underline{t})}^{\overline{v}} (1 - G(v|\underline{t}))dv = 0.$$

It remains to show that the sequential mechanism $\{k(t), x_0(t), x_1(t)\}$ defined above satisfies all IC_1 and IR constraints. The partial derivative of the expected utility $U(t', t)$ of type t' when he claims to be type t is given by

$$\frac{\partial U(t',t)}{\partial t} = -\frac{dx_0(t)}{dt} - \frac{dk(t)}{dt}(1 - G(k(t)|t')).$$

Since by assumption $y(t, v)$ is non-decreasing in t for all v, $\frac{dk(t)}{dt} \leq 0$. Suppose $t' < t$. Then,

$$\frac{\partial U(t',t)}{\partial t} \leq -\frac{dx_0(t)}{dt} - \frac{dk(t)}{dt}(1 - G(k(t)|t)) = 0.$$

By integration we have $U(t', t) \leq U(t')$. The same reasoning applies if $t < t'$. This shows that the sequential mechanism $\{k(t), x_0(t), x_1(t)\}$ satisfies all IC_1 constraints. As in Lemma 2.1, since all IC_1 constraints are satisfied and since IR holds for the lowest type $\underline{t}$, all IR constraints are satisfied. *Q.E.D.*

In the other case, when T is ordered by MPS with all distribution functions passing through a single point at z, the second term of $\phi(t, v)$ is positive for $v < z$ and negative for $v > z$. Depending on whether the cost c is low or high relative to z, the proof of Lemma 3.3 needs to be adapted. The statement of Lemma 3.3 holds for the case of MPS with an additional restriction on the solution to the relaxed problem, namely no under-production if $c < z$ and no over-production if $c > z$.

LEMMA 3.4. *Suppose that T is ordered by MPS with all distributions passing through a single point at z. If $c < z$ (resp. $c > z$) and $y(t, v)$ solves the relaxed problem with no under-production (over-production), and if $y(t, v)$ is non-increasing (non-decreasing) in t for all v and non-decreasing in v for all t, then there exists $x(t, v)$ such that $\{y(t, v), x(t, v)\}$ is optimal.*

PROOF. Define a sequential mechanism $\{k(t), x_0(t), x_1(t)\}$ as in the proof of Lemma 3.3. It suffices to show that all IC_1 constraints are satisfied. Suppose $c < z$; the case of $c > z$ is symmetric. We have

$$\frac{\partial U(t', t)}{\partial t} = -\frac{dx_0(t)}{dt} - \frac{dk(t)}{dt}(1 - G(k(t)|t')).$$

By assumption $y(t, v)$ is non-increasing in t for all v, so $\frac{dk(t)}{dt} \geq 0$. Moreover, since there is no under-production, $k(t) \leq c$ for all t. Then, if $t' < t$, MPS implies $G(k(t)|t') \leq G(k(t)|t)$, and

$$\frac{\partial U(t', t)}{\partial t} \leq -\frac{dx_0(t)}{dt} - \frac{dk(t)}{dt}(1 - G(k(t)|t)) = 0.$$

By integration we have $U(t', t) \leq U(t')$. The same reasoning applies if $t < t'$. *Q.E.D.*

Before we present a few parameterizations where optimal sequential mechanisms can be found by using Lemma 3.3 and Lemma 3.4, it is helpful to compare our model with the standard one-dimensional nonlinear pricing problem. The coefficient $\phi(t, v)$ is analogous to "virtual surplus" defined by Myerson [1981] in one-dimensional nonlinear pricing problems. As in nonlinear pricing problems, the first part of $\phi(t, v)$ corresponds to social surplus of type t with valuation v, and the second part represents the distortion. The difference is that in a nonlinear pricing problem, the second part contains only the "hazard rate" $\frac{1-F(t)}{f(t)}$, but in our sequential screening problem, it also contains an additional term $\frac{\partial G(v|t)/\partial t}{g(v|t)}$.[12] The hazard rate measures the distortion due to eliciting truthful type information from t, for any valuation v. Distortions are larger with a greater hazard rate, because whatever surplus conceded to type t must also be given to all higher types. The term $\frac{\partial G(v|t)/\partial t}{g(v|t)}$ has a straightforward interpretation of "informativeness measure" (Baron and Besanko [1984]), as it represents how informative the consumer's private type knowledge is about his valuation. It is zero if type and valuation are independently distributed, and is large if marginally different types have very different conditional distributions. Alternatively, holding $G(v|t)$ constant, we can think of v as a function of t, and the informativeness

[12] Whenever Lemma 3.3 and Lemma 3.4 apply, the sequential mechanism design problem is reduced to choosing refund (cutoff valuation) as a function of type. Virtual surplus can be instead defined as the expected total surplus for a given type and a given refund. This alternative definition of virtual surplus looks the same as in a standard one-dimensional nonlinear pricing problem. We choose to define virtual surplus for a pair of valuation and type, because it applies even when the conditions of Lemma 3.3 and Lemma 3.4 do not hold.

measure is equal to $-\frac{\partial v}{\partial t}$. The measure then represents how marginally differently types hit a fixed percentile $G(v|t)$ at different valuations. Distortions are larger with a greater informativeness measure, because more rent must be conceded in order for marginally different types not to claim to be type t with valuation v. Finally, in a nonlinear pricing problem, the usual second-order condition (analogous to SOC_2 in Lemma 3.1) implies downward distortion—consumers of every type except for the highest one are rationed. Here, the direction of distortion is not necessarily downward because the ratio $\frac{\partial G(v|t)/\partial t}{g(v|t)}$ can be either positive or negative. We will discuss the case of FSD and the case of MPS separately.

For the case of FSD, let's first consider the following "additive" structure of conditional distributions:

$$v = \theta t + (1-\theta)\epsilon_t,$$

where t is distributed over a positive range, $\theta \in (0,1)$, and ϵ_t is i.i.d. on the whole real line (this guarantees that the conditional distributions have the same support) with density $h(\cdot)$ and distribution $H(\cdot)$.[13] The distribution of v conditional on t is given by

$$G(v|t) = H\left(\frac{v-\theta t}{1-\theta}\right).$$

Note that $G(v|t)$ satisfies FSD. The additive specification has some nice properties that make it an interesting benchmark case of first-order stochastic dominance. Consumers face the same uncertainty regarding valuation but have private information about their expected valuation for the good. In this linear case, the informativeness measure becomes a global one—it equals θ for all types and valuations. A greater θ means that more information is learned early rather than late and the consumer's private type knowledge is more informative about valuation in that conditional distributions of valuations vary more with type.

With this additive specification, we have

$$\phi(t,v) = v - c - \frac{\theta(1-F(t))}{f(t)}.$$

[13] This formulation allows negative realized valuation. An example of negative valuation is a ticket-holder who is sick and must be paid to travel. Since the monopolist cannot force consumption (free disposal), in the profit maximization problem a distribution with a range of negative valuations is equivalent to one that has an atom at zero valuation with all the probability weights of the negative valuations. This does not change the results below.

Under the standard monotone hazard rate assumption that the hazard rate $\frac{1-F(t)}{f(t)}$ is non-increasing in t (see, e.g., Fudenberg and Tirole [1991]), we have that $\phi(t,v) \geq 0$ implies $\phi(t',v') \geq 0$ for any $t' \geq t$ and $v' \geq v$. The solution to the relaxed problem is monotone in v and t separately. By Lemma 3.3, it solves the original problem. The optimal delivery rule is therefore given by:

$$y(t,v) = \begin{cases} 1 & \text{if} \quad v > c + \frac{\theta(1-F(t))}{f(t)} \\ 0 & \text{otherwise.} \end{cases}$$

It is deterministic with a cutoff level for each type. Higher types have lower cutoffs. There is no production distortion for the highest type. We summarize the findings in the following proposition.

PROPOSITION 3.5. *Suppose that the conditional distribution functions have an additive structure. Then, under the monotone hazard rate assumption, the optimal sequential mechanism is deterministic with larger under-production distortions for lower types and no under-production for the highest type.*

Under-production distortion is larger when the consumer's private knowledge is more informative, because the monopolist prefers rationing the good to giving higher types a large informational rent. In the polar case where $\theta = 0$, type is completely uninformative of valuation, and the monopolist achieves perfect discrimination with a sequential mechanism. The monopolist sells the product in period one at the expected valuation, which is the same for all types, and allows the consumer to return the good for a refund equal to c. This refund policy guarantees social efficiency. In the other polar case where $\theta = 1$, the under-production distortion is the largest. Clearly, the optimal sequential mechanism coincides with usual monopoly pricing after the consumer learns his valuation.

A characteristic of the optimal menu of refund contract is that it is independent of the specification $H(\cdot)$ of the valuation shock ϵ_t. For fixed θ, increasing the variance of ϵ_t means that the consumer faces greater valuation uncertainty, yet this has no effect on the optimal menu of refund contracts. What matters is not how much the consumer knows about his valuation when he signs the contract, but how informative his private type knowledge is about his valuation from the monopolist's point of view. The shock ϵ_t may be interpreted as a demand shock. As long as it affects all consumer types independently so that it does not affect the informativeness measure, the size of this demand shock has no impact on the optimal sequential mechanism.

It is instructive at this point to compare the optimal sequential mechanism with *ex post* monopoly pricing. Whereas the sequential mechanism is deterministic with lower cutoff levels of valuations for higher types, *ex post* monopoly pricing can be thought of as a deterministic sequential mechanism with full refund and all types having the same cutoff level. In general, the optimal sequential mechanism yields greater profits than *ex post* optimal monopoly pricing; the gains from sequential screening tend to be greater when θ is close to 0. If sequential mechanisms involve greater implementation costs than *ex post* monopolist pricing, perhaps due to the cost of registering consumers in advance, then one is less likely to observe sequential screening in environments where conditional distributions of valuations vary substantially with type. Welfare comparison between sequential pricing and monopolist *ex post* pricing is ambiguous. In the monopolist pricing, expected downward distortions are smaller for higher types because they are more likely to reach above the same cutoff level. In the optimal menu of refund contracts, higher types have lower cutoff levels so their expected downward distortions are even smaller compared to lower types. However, since the optimal *ex post* monopoly price depends on both the distribution of types and the conditional distributions of valuations, aggregate downward distortion under *ex post* monopolist pricing can be either higher or lower than that under the optimal sequential mechanism.

We can also extend Proposition 3.5 to a multiplicative specification.[14] Suppose that $v = t^{\theta}\epsilon_t^{1-\theta}$ where t is distributed over a positive range, $\theta \in (0,1)$, and ϵ_t is i.i.d. on the whole positive real line (this guarantees that the conditional distributions have the same support) with density $h(\cdot)$ and distribution $H(\cdot)$. This specification is log-linear and can be useful in constructing empirically testable implications. We have,

$$\phi(t,v) = v - c - \frac{v\theta(1-F(t))}{tf(t)}.$$

As in the additive specification, the greater θ is, the more informative the type is as a signal of valuation, but informativeness measure is not uniform across types or across valuations. Define $\lambda(t) = \frac{1-F(t)}{tf(t)}$ and suppose that $\lambda(t) < 1$ for all t. This is satisfied as long as the

[14] A specification with a special distribution that works out similarly is $G(v|t) = 1 - \exp(-v/t)$.

range of t is sufficiently above zero, regardless of the value of θ. The solution to the relaxed problem is then

$$y(t,v) = \begin{cases} 1 & \text{if} \quad v > \frac{c}{1-\theta\lambda(t)} \\ 0 & \text{otherwise.} \end{cases}$$

If $\lambda(t)$ is non-increasing in t (monotone hazard rate is sufficient for this but clearly not necessary), the above solution has the monotone property required by Lemma 3.3 and therefore solves the original problem.

For the case of MPS, perhaps the most natural class of distributions is given by the same mean plus a multiplicative shock. Suppose that

$$v = z + t\epsilon_t,$$

where ϵ_t is i.i.d. on the whole real line (this guarantees that the conditional distributions have the same support) with zero mean, density $h(\cdot)$ and distribution $H(\cdot)$. Without loss of generality assume that $\underline{t} > 0$, so that greater t means greater dispersion. The distribution of v conditional on t is given by

$$G(v|t) = H\left(\frac{v-z}{t}\right).$$

It is easy to see that the distributions $G(v|t)$ satisfy MPS and pass through the same point at z, which is also the mean of the distributions. Consumers face the same expected valuation but have private information about the degree of valuation uncertainty. Consumers of higher types face greater valuation uncertainty. The informativeness measure is given by $(v-z)/t$. The private type knowledge of higher types is less uninformative about their valuation.

With the above specification, we have

$$\phi(t,v) = v - c - (v-z)\lambda(t),$$

where $\lambda(t)$ was defined in the previous example. Under the standard monotone hazard rate assumption (sufficient but not necessary), $\lambda(t)$ is non-increasing in t. Suppose that $\lambda(t) \leq 1$. This is satisfied if the range of t is sufficiently above zero, regardless of the distribution $F(t)$. Then, the solution to the relaxed problem is given by:

$$y(t,v) = \begin{cases} 1 & \text{if} \quad v > \frac{c-\lambda(t)z}{1-\lambda(t)} \\ 0 & \text{otherwise.} \end{cases}$$

The assumption of $\lambda(t) \leq 1$ guarantees that $y(t, v)$ is non-decreasing in v for any t. It is straightforward to show that the cutoff rule $y(t, v)$ defined above has the properties required by Lemma 3.4: if $c < z$, then $y(t, v)$ has no under-production and is non-increasing in t for all v; if $c > z$, then $y(t, v)$ has no over-production and is non-decreasing in t for all v. The following proposition follows.

PROPOSITION 3.6. *Suppose that the valuation distributions are given by the same mean z plus a multiplicative shock and $\frac{1-F(t)}{tf(t)}$ is non-increasing in t and less than 1. Then, if $c < z$ (resp. $c > z$), the optimal sequential mechanism is deterministic with greater over-production (under-production) distortions for types with smaller valuation uncertainty and no distortion for the highest type.*

Proposition 3.6 mirrors Proposition 2.3. When consumers have the same expected valuation but differ in the valuation uncertainty they face, the pattern of consumption distortion is determined by the comparison between the production cost and the expected valuation. In airplane ticket pricing, the cost of flying an additional passenger is typically small compared to the average willingness to pay when plane capacity is not binding. In this case, all travelers except for those with the greatest valuation uncertainty are subsidized and purchase advance tickets with refund lower than the cost of the ticket. Travelers with greater uncertainty about their plans pay more in advance for greater flexibility in terms of higher refund.

Note that as in Proposition 3.5 the optimal menu of refund contract does not depend on the specification $H(\cdot)$ of the valuation shock ϵ_t. The variance of ϵ_t can be great or small, but it has no effect on the optimal menu of refund contract as long as it is common to all types. Also, the multiplicative specification can be generalized by relaxing the assumption that consumer type enters linearly with ϵ_t. Finally, as in Proposition 3.5, consumption distortions (either downward or upward) are larger when the consumer's private type knowledge is more informative about his valuation at the time of contracting. To see this, recall that the informativeness measure is given by $(v - z)/t$. Greater informative type knowledge can be represented by a leftward shift of the type distribution. Since $\lambda(t)$ is decreasing in t, this amounts to an increase in $\lambda(t)$ for each t. The cutoff levels in the solution to the relaxed problem decrease if $c < z$ and increase if $c > z$. In either case, the consumption distortions are smaller.

4. Further Comments on Sequential Screening

The optimal sequential mechanisms characterized in previous sections are all deterministic. Deterministic sequential mechanisms are important because they are easy to implement in practice, through refund contracts, option contracts or cancellation fees. Deterministic sequential mechanisms are also related to two-part tariffs in nonlinear pricing models, as formally discussed in the mechanism design language by Laffont and Tirole [1986]. A typical optimal mechanism in the literature is a direct mechanism rarely seen in practice; instead, two-part tariffs are often used. Although any concave nonlinear tariff can be implemented through a two-part tariff, the sequential feature of the consumer's decision in two-part tariffs is purely artificial. One explanation for the popularity of two-part tariffs, given by Laffont and Tirole, is that they are robust against shocks. Our model of sequential screening, where the consumer self-selects twice, provides another explanation for the use of two-part tariffs. The shock mentioned by Laffont and Tirole can be viewed as the uncertainty faced by the consumer about his actual valuation at the time of contracting, which later becomes his additional private information.

Deterministic sequential mechanisms are also easy to characterize. Under either FSD or MPS, optimal deterministic mechanisms can be characterized with a local approach, because the design problem is reduced to a single-dimensional problem of choosing refund as a function of type. Unfortunately, neither FSD nor MPS is sufficient to imply that optimal sequential mechanisms are deterministic. Without restricting sequential mechanisms to menus of refund contracts, the local period-one constraints are generally insufficient to imply the global constraints. In our model of sequential screening, as in the standard one-dimensional price discrimination literature, localization of the period-two incentive compatibility constraints is guaranteed due to the standard "single crossing" condition (see, e.g., Cooper [1984]) that a consumer with a greater realized valuation is willing to pay more for an increase in delivery probability. This enables us to replace IC_2 constraints by the two local conditions in Lemma 3.1. However, the two local period-one incentive constraints in Lemma 3.2 are not sufficient to imply global period-one incentive constraints.[15] This

[15] In the multi-dimensional price discrimination literature, McAfee and McMillan [1988] have identified a "generalized single crossing" condition that guarantees that local incentive compatibility constraints imply global constraints. But their condition requires that the dimension of consumer's private information exceed the number of monopolist' price discrimination instruments. This dimensionality condition is not

insufficiency results from the multi-dimensional nature of the sequential mechanism design problem. To see this, note that from Lemma 3.3, under FSD a sufficient condition for the IC_1 constraints is that $y(t,v)$ is non-decreasing in t for all v, but SOC_1 states only that the delivery rule $y(t,v)$ is non-decreasing in t "on average," with the weights determined by local changes in the conditional distributions of valuations with respect to type. Without further restrictions on type space T besides stochastic dominance, the weights can change with type arbitrarily, and there is little hope that what holds locally extends globally.[16] When the delivery rule is not monotone increasing in type for each valuation, "bunching" occurs both across valuations and across types. In this case, random delivery rules allow fine tuning by the monopolist.[17] That the delivery rule need not be monotone in types is similar to the conclusion in the multi-product price discrimination literature that quantity or quality of each good need not be monotone when consumer demand characteristics are one-dimensional but the number of price discrimination instruments is greater than one. For example, Matthews and Moore [1987] show that if the monopolist in the model of Mussa and Rosen [1978] offers different levels of warranty as well as quality, more eager consumers need not buy higher quality or receive higher warranty.

To conclude the discussion of general sequential mechanisms, we use a two-type example of airplane ticket pricing to illustrate how random delivery rules can fine tune sequential screening in the absence of FSD or MPS. This example is closely related to the one in section 2. As in section 2, the two types are interpreted as business type and leisure type, with the business type facing greater valuation uncertainty in having a distribution of valuations that second-order stochastically dominates that of the leisure type. The difference here is that the business type has a smaller mean valuation than the leisure type. In particular, suppose that there are three possible valuations $v-2$, v, and $v+1$, where v is a number greater than 2. The leisure type has a deterministic valuation of v. The business type is

satisfied in our model: at the time of contracting, the consumer's private information is one-dimensional, but the monopolist has many instruments in contingent packages of delivery rule and payments.

[16] It is possible to make stronger assumptions on the type space to reduce the dimension of the design problem. For example, if in addition to satisfying FSD or MPS, the distribution functions $G(v|t)$ are separable in t and v in that there exist functions $\alpha(\cdot)$, $\beta(\cdot)$ and $\gamma(\cdot)$ such that $G(v|t) = \alpha(v) + \beta(t)\gamma(v)$ for all t and v, then one can use a variation of the standard local approach in nonlinear pricing problems to characterize the optimal sequential mechanism.

[17] Numerical examples of randomization in optimal sequential mechanism under FSD or MPS are available from the authors.

equally likely to draw valuation of $v-2$ and $v+1$, with a lower mean $v-\frac{1}{2}$. For simplicity we assume that the production cost is zero. We will see how a tension arises under refund contracts between exploiting the high mean valuation of the leisure type and the fat tail of the business type, and how randomization can help fine tune sequential discrimination.

First consider the following sequential mechanism with randomization. The seller offers one contract with an advance price of v and zero refund, intended for the leisure type, and another contract intended for the business type, with the same advance price but with an option of receiving a partial refund of $\frac{1}{2}v$ and then drawing a lottery with probability $\frac{1}{2}$ of getting the ticket. Note that if the business type purchases the partial refund contract, it is strictly better off not claiming the refund if it draws the higher valuation of $v+1$ (because $v+1 > \frac{1}{2}v + \frac{1}{2}(v+1)$), whereas if the leisure type buys the partial refund contract, it is always indifferent between claiming and not claiming the refund (because $v = \frac{1}{2}v + \frac{1}{2}v$). The leisure type gets zero surplus from either contract (constraints IR_L and $IC_{L,B}$ both hold with equality), and the business type gets zero surplus from the partial refund contract (IR_B binds) and negative surplus from the no-refund contract because it has a mean valuation smaller than the advance price of v ($IC_{B,L}$ does not bind). Since both participation constraints bind, the seller's expected profits are $\Pi = f_L v + f_B(\frac{1}{4}(v-2) + \frac{1}{2}(v+1))$, where f_L and f_B are the fractions of leisure type and business type respectively.

Next consider any deterministic sequential mechanism. Since in the sequential mechanism with randomization the seller extracts all surplus from the leisure type, for the profits from a deterministic sequential mechanism to exceed Π, the business type must be induced to use the ticket with probability one. But since the leisure type has a higher mean valuation, it will also use the ticket with probability one. Since the participation constraint of the business type must be satisfied, the seller is essentially restricted to offering the ticket at an advance price of $v-\frac{1}{2}$ with no refund, with expected profits of $v-\frac{1}{2}$. Therefore, if $v-\frac{1}{2} < \Pi$, or $f_L > \frac{v-2}{v}$, the optimal sequential mechanism must involve randomization.

The intuition behind this example can be understood as follows. It is straightforward to argue that for a sequential mechanism to be optimal, the seller must sell to the leisure type with probability one and the business type must be induced to use the ticket with probability one when it draws the high valuation $v+1$.[18] But the question is whether to

[18] Formally, any mechanism in which a type does not use the ticket with probability one when it draws

sell to the business type when it draws the low valuation. If a menu of refund contracts has to be used, the seller faces an all-or-nothing choice: she must either sell to the business type with probability one, or sell only when it draws the high valuation. In the first case, she is constrained to offering a single contract at an advance price of $v - \frac{1}{2}$ and no refund. Since the leisure type has a higher mean valuation of v, a rent equal to $\frac{1}{2}$ is left to the leisure type. This rent is large if f_L is large. In the second case, the seller can achieve separation by offering one contract at an advance price of v with no refund to the leisure type and another contract at an advance price of $v+1$ with full refund to the business type. But this means that she has to give up an expected surplus of $\frac{1}{2} f_B(v-2)$ from the business type when it draws the low valuation. Randomization provides an extra margin for the seller to operate at. By providing a partial refund of $\frac{1}{2}v$ and a lottery to the business type when it draws the higher valuation, the seller leaves no rent to the leisure type and extracts half of the surplus from the business type when it draws the low valuation.

5. Concluding Remarks

The closest work to the present study is by Miravete [1996], who also considers the monopolist's pricing problem when consumers face demand uncertainty. Miravete [1997] tests empirical implications of his model. In contrast with this paper, he assumes continuous demand functions. This allows him to compare *ex ante* two-part tariffs (where consumers choose a tariff based on their expected demand) and *ex post* two-part tariffs (after consumers learn their actual demand). He shows that expected profits are higher under an *ex post* tariff if the variance of the *ex ante* type distribution is large enough. The generality of his results is compromised by the restriction to two-part tariffs, because *ex ante* two-part tariffs are generally not optimal. In the present paper where unit demand is assumed, the optimal *ex post* tariff degenerates to standard monopolist pricing, and can be thought of as a uniform sequential contract with full refund for all *ex ante* types. As a result, in our model *ex post* mechanisms are dominated by sequential mechanisms. Furthermore, our results

the highest valuation is not optimal. To prove this by contradiction, suppose that the probability of getting the ticket is $p < 1$ for some type after drawing its highest possible valuation. Then, the seller can increase this probability to one and at the same time increase the corresponding payment from the type by $v(1-p)$. This will not affect the participation or the incentive constraint of any type, but will increase the profits by $v(1-p)$ when the buyer happens to be this type and happens to draw the highest valuation.

(Proposition 3.5 and Proposition 3.6) indicate that profit gains from using a sequential mechanism depend not only on the type distribution, but on how informative consumers' initial private knowledge is about their valuations: if different types of consumers have very different conditional distributions of valuations, then sequential mechanisms do not yield much greater expected profits than *ex post* monopolist pricing.

Our sequential mechanism design problem is related to the problems of dynamic price discrimination (e.g., Baron and Besanko [1984], Laffont and Tirole [1988, 1990]). An example of these problems is a monopolist facing a consumer making repeated purchases. Typically, consumers have only one piece of private information and it does not change over time. These problems focus on the implications of the monopolist's ability to commit. With no change in consumers' private information over time, the optimal dynamic mechanism under commitment is static: it simply replicates the optimal static screening contract in every period. In contrast, our sequential screening problem is driven by the demand uncertainty, and consumption decisions are made based on new information.

Throughout the paper we assume that the monopolist can commit to a sequential mechanism and examine how a sequential mechanism can be used by the monopolist to extract maximal surplus. If the monopolist cannot commit to a sequential mechanism, time-inconsistency problems as mentioned by Coase [1972] arise. In our problem of sequential screening, the monopolist may be tempted to renege both before and after consumer learn their valuation. Lack of ability to commit to a sequential mechanism reduces the monopolist's power to discriminate, which may explain why sometimes these types of mechanisms are not observed in practice. Incorporating the commitment issue to the design of sequential mechanisms is an interesting topic for further research.

Although menus of refund contracts arise naturally from price discrimination under consumer learning, they can be offered by producers for other reasons. One reason worth mentioning is that a menu of refund contract may allow a producer to learn about final consumer demand early on. This information can be valuable for production planning purposes. Consider the airplane ticket pricing example in the introduction. The fraction of business travelers may be unknown to the ticket seller, but it is revealed by travelers' choices of refund contracts. This information is valuable to the seller if a production capacity decision must be made before the final demand is realized. Although this paper has not explored the issue of capacity constraint in the presence of sequential price discrimination, it seems a promising line of research.

References

William Adams and Janet L. Yellen [1976]: "Commodity Bundling and the Burden of Monopoly." *Quarterly Journal of Economics* 90:475-98.

Mark Armstrong [1996]: "Multiproduct Nonlinear Pricing," *Econometrica* 64(1), 51-76.

David Baron and David Besanko [1984]:"Regulation and Information in a Continuing Relationship," *Information Economics and Policy* 1: 267-302.

Ronald Coase [1972]:"Durability and Monopoly," *Journal of Law and Economics* 15, 143-149.

Russell Cooper [1984]: "On Allocative Distortions in Problems of Self-Selection," *Rand Journal of Economics* 15, 569-577.

James Dana [1998]: "Advance-Purchase Discounts and Price Discrimination in Competitive Markets," *Journal of Political Economy* 106(2), 395-422.

Drew Fudenberg and Jean Tirole [1991]: *Game Theory*, MIT press.

Milton Harris and Robert Townsend [1981]: "Resource Allocation Under Asymmetric Information," *Econometrica* 49(1), 33-64.

Jean-Jacques Laffont and Jean Tirole [1986]: "Using Cost Observation to Regulate Firms," *Journal of Political Economy*, 94, 614-641.

Jean-Jacques Laffont and Jean Tirole [1988]: "The Dynamics of Incentive Contracts," *Econometrica* 56, 1153-1175.

Jean-Jacques Laffont and Jean Tirole [1990]: "Adverse Selection and Renegotiation in Procurement," *Review of Economic Studies*, 57, 597-625.

Tracy Lewis and David Sappington [1989]: "Countervailing Incentives in Agency Problems," *Journal of Economic Theory* 49(2), 294-313.

Steven Matthews and John Moore [1987]: "Monopoly Provision of Quality and Warranties: an Exploration in the Theory of Multidimensional Screening," *Econometrica*, 55, 441-467.

Eric Maskin and John Riley [1984]: "Monopoly with Incomplete Information," *Rand Journal of Economics* 15(2), 171-196.

R. Preston McAfee and John McMillan [1988]:"Multidimensional Incentive Compatibility and Mechanism Design," *Journal of Economic Theory* 46, 335-354.

Eugenio J. Miravete [1996]: "Screening Consumers through Alternative Pricing Mechanisms," *Journal of Regulatory Economics* 9: 111-132.

Eugenio J. Miravete [1997]: "Estimating Demand for Local Telephone Service with Asymmetric Information and Optimal Calling Plans" INSEAD Working Paper.

James Mirrlees [1971]: "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies* 83, 175-208.

Michael Mussa and Sherwin Rosen [1978]:"Monopoly and Product Quality," *Journal of Economic Theory* 18, 301-317.

Roger Myerson [1979]: "Incentive Compatibility and the Bargaining Problem," *Econometrica* 47(1), 61-73.

Roger Myerson [1981]: "Optimal Auction Design," *Mathematics of Operations Research* 6:58-73.

Jean-Charles Rochet and Philippe Chone [1998]: "Ironing, Sweeping and Multidimensional Screening," *Econometrica* 66(4), 783-826.

Michael Rothschild and Joseph E. Stiglitz [1970]: "Increasing Risk: I. A Definition," *Journal of Economic Theory* 2, 225-243.

Moshe Shaked and J. George Shanthikumar [1994]: *Stochastic Orders and Their Applications*. Academic Press.

Lars Stole [1996]: "Lectures on the Theory of Contracts and Organizations," manuscript, The University of Chicago.

Robert Wilson [1993]: *Nonlinear Pricing*, Oxford University Press.