Sequential Screening

Pascal Courty
Universitat Pompeu Fabra

Li, Hao
University of Hong Kong

Current version: May 18, 1998

Abstract: We present a model of price discrimination where a monopolist faces a consumer who is privately informed about the distribution of his valuation for an indivisible unit of good but has yet to learn privately the actual valuation. The monopolist sequentially screens the consumer with a menu of contracts: the consumer self-selects once by choosing a contract and then self-selects again when he learns the actual valuation. 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 sequential mechanisms may involve randomization. We characterize the optimal sequential mechanism when some consumer types are more eager in the sense of first-order stochastic dominance, and when some types face greater valuation uncertainty in the sense of mean-preserving-spread. We show that it can be optimal to subsidize consumer types with smaller valuation uncertainty (through low refund, as in airplane ticket pricing) in order to reduce the rent to those with greater uncertainty. The size of distortion depends both on the type distribution and on how informative the consumer's initial private knowledge is about his valuation, but not on how much he initially knows about the valuation per se.

Acknowledgements: We would like to thank Mark Armstrong, Michael Chwe, Jim Peck, Sherwin Rosen, Lars Stole, Wing Suen, the Associate Editor and the referees for helpful comments. The Associate Editor and the referees also suggested 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 they assume that consumers know their demand when they select a contract. This restriction on consumer demand excludes situations where consumers have only partial information at the time of contracting due to unforeseen contingencies and learn more about their demands as these contingencies are resolved. Consider the demand for plane tickets. Travelers typically do not know their valuations for tickets until just before departure, but they know 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 their distribution of 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. Suppose that half of all potential buyers of the ticket are leisure travelers who value it at \$600 for sure, and the other half are business travelers who have valuation \$1000 when they make the trip but are equally likely to have zero valuation. The cost of flying an additional traveler is small but not zero. If the seller waits until travelers have privately learned their valuations, she gets a profit of almost \$600 per traveler by charging the monopoly price \$600. Now suppose she offers two contracts before the travelers learn their valuations, one with an advance payment of \$600 and no refund and the other with an advance payment of \$1000 and full refund. Business travelers strictly prefer the contract with refund, and leisure travelers are indifferent between the two contracts so we assume that they choose the contract with no refund. The monopolist separates the two types

[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.

and earns an average profit of almost $800 per traveler, representing a one-third increase compared with charging the monopoly price after travelers have learned their valuations.[2] This menu of refund contracts remains optimal if there is a small probability that leisure travelers have zero valuation, showing that it can be optimal to subsidize travel by the leisure type.

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 incomplete private information is important.

Surprisingly, sequential screening has not received much direct attention in the screening literature.[3] It is true that sequential mechanisms share the characteristic with two-part tariffs that consumption decisions are made sequentially, and there is an abundant literature on the latter (see, e.g., Wilson [1993]). But the empirical importance of sequential mechanisms suggests that two-part tariffs are more than a simple way of implementing concave nonlinear tariffs, as suggested by the literature. 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 charged by the hotel had he known her itinerary when he reserved the room.

---

2 A similar example is presented in Courty [1996].

3 An exception is Miravete [1996]. Miravete [1997] tests empirical implications of his model. We thank Hugo Hopenhayn and an anonymous referee for drawing our attention to his works.

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 will consider only situations where consumers have unit demands such as the airplane ticket pricing problem. In these situations, optimal *ex post* pricing scheme (after consumers have complete private information about their demand) degenerates to the 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 (that the monopolist does not deliver the good for some valuations greater than the cost) are similar to the standard result in nonlinear pricing models (see, e.g., Mussa and Rosen [1978], or Maskin and Riley [1984]) that under-provision of quality or quantity is used to extract more surplus from more eager consumers. More surprisingly, it may be optimal to subsidize consumption by some consumers. Inefficient over-production never occurs in the single-product monopoly models but it does in multi-product price discrimination problems (see, e.g., Adam and Yellen [1976], or Rochet [1995]) when better separation is achieved by subsidizing some goods. 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 the case of two *ex ante* types of potential buyers. To continue with the airplane ticket pricing example, the business traveler type is either a more eager consumer more likely to draw greater valuations, in the sense of first-order stochastic dominance, or faces greater valuation uncertainty, in the sense of mean-preserving-spread (Rothschild and

Stiglitz [1970]).[4] In either case (and a combination of the two cases), 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 relatively 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 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, allowing for random delivery rules. This generalization enables us to discuss how type distribution and consumer learning affect the design of sequential mechanisms. Sequential mechanism design is similar to a static multi-product price discrimination problem. The monopolist can be thought of as selling multiple products, corresponding to the same product delivered for different valuations. 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 does not involve randomization, and can therefore be interpreted as a menu of refund contracts, or a two-part tariff. The size of distortion depends both on the type distribution and on how informative consumers' initial private knowledge is about their valuations, but not on how much consumers initially know about their valuations per se. Distortions are small if consumers' initial private knowledge is not informative about their valuations in that conditional distributions of valuations do not vary much across different types of consumers. 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

[4] After finishing an earlier version of the paper, we received some notes from Mark Armstrong, who had earlier considered the two-type case under first-order stochastic dominance. We are grateful for his comments.

and lower refund. This is similar to the quantity discount result in single-product nonlinear pricing models. 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 comments on the general problem of sequential screening. Neither first-order dominance nor mean-preserving-spread are sufficient in reducing the dimension of the design problem. Stronger conditions are necessary; we provide one condition under which the optimal sequential mechanism can be characterized. This condition imposes a linear structure on the type space and enables us to use a variation of the standard local approach in nonlinear pricing problems. An example in the appendix with three types and three valuations shows how "bunching" occurs across types and across valuations at the same time and results in a random delivery rule, although the extent of randomization is shown to be rather limited under the linear structure of the type space. Section 5 concludes with some discussion on related works and remarks on the monopolist's ability to commit to a sequential mechanism.

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

Consider a monopoly seller of airplane tickets facing two types of travelers, type H and L. Throughout this section, we will think of type H as the "business type," which values the ticket more or faces greater valuation uncertainty; type L is a "leisure" traveler. The proportion of the business and leisure types is f_H and f_L respectively. There are two periods. In the beginning of period one, a 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.

Greater mean valuation of the business type is captured by first-order stochastic dominance (FSD). The valuation distribution G_H of the business type H first-order stochastically dominates G_L of the leisure type L if $G_H(v) \leq G_L(v)$ 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_H dominates G_L by MPS if they have the same mean and $\int_{\underline{v}}^{v}(G_H(u) - G_L(u))du \geq 0$ for $v \in [\underline{v}, \overline{v}]$.

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 $\{a_H, k_H, a_L, k_L\}$. The profit maximization problem can be written as:

$$\max_{k_L, k_H, a_L, a_H} \sum_{t=L,H} f_t(a_t - G_t(k_t)k_t - (1 - G_t(k_t))c)$$

subject to

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

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

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 $IC_{H,L}$ imply IR_H.*

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

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

Then, $IC_{H,L}$ implies

$$-a_H + \int_{\underline{v}}^{\overline{v}} \max\{k_H, v\} dG_H(v) \geq -a_L + \int_{\underline{v}}^{\overline{v}} \max\{k_L, v\} dG_H(v).$$

Since $\max\{k_L, v\}$ is an increasing function of v, if G_H dominates G_L by first-order,

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

Since $\max\{k_L, v\}$ is a convex function of v, the above condition holds also if $G_H(v)$ 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_H that dominates G_L by MPS. Shifting the whole distribution G_H to the right gives a new distribution that has both greater valuation and greater uncertainty. It is easy to see that the above lemma continues to hold for this combination of FSD and MPS.[5]

The maximization problem can now be simplified. 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_H by the same amount would increase profits. Also, $IC_{H,L}$ binds in the optimal menu of refund contracts, otherwise profits could be increased by increasing a_H.

LEMMA 2.2. *In the optimal menu of refund contracts, $IC_{L,H}$ is satisfied if and only if*

$$\int_{k_H}^{k_L} (G_H(v) - G_L(v))dv \leq 0.$$

PROOF. Since $IR_{H,L}$ binds in the optimal menu of refund contracts,

$$a_L - a_H = \int_{k_H}^{k_L} G_H(v)dv.$$

We have

$$\begin{aligned} & -a_L + k_L G_L(k_L) + \int_{k_L}^{\overline{v}} v dG_L(v) \\ = & -a_H + k_H G_L(k_H) + \int_{k_H}^{\overline{v}} v dG_L(v) - \int_{k_H}^{k_L} (G_H(v) - G_L(v))dv. \end{aligned}$$

The lemma then follows from $IC_{L,H}$. *Q.E.D.*

[5] 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.

Under FSD, the condition in Lemma 2.2 is equivalent to $k_H \leq k_L$, a monotonicity condition (also called second-order condition) found in standard screening problems. In the case of MPS, however, there is no such implication of monotonicity. This is because the business type with greater uncertainty is not necessarily more eager to consume at the margin, even though by Lemma 2.1 the business type values any given contract more than the leisure type.

By Lemma 2.1 and Lemma 2.2, the profit maximization problem can be simplified as:

$$\max_{k_L, k_H} \int_{k_L}^{\overline{v}} (f_L(v-c)g_L(v) - f_H(G_H(v) - G_L(v)))dv + \int_{k_H}^{\overline{v}} (f_H(v-c)g_H(v))dv$$

subject to the second-order condition in Lemma 2.2. Let $S(k_L) = \int_{k_L}^{\overline{v}} (v-c)g_L(v)dv$ be the surplus from the leisure type, and $R(k_L) = \int_{k_L}^{\overline{v}} (G_L(v) - G_H(v))dv$ be the rent to the business type, both as function of the refund to the leisure type. Note that the rent function $R(k_L)$ behaves differently under FSD than under MPS: in the first case, $R(k_L)$ is decreasing for any k_L; in the second case, $R(k_L)$ is zero at both $\underline{v}$ and $\overline{v}$, and tends to be greater in the middle of the support. The different behavior of the rent function is what distinguishes the two cases.

Consider the relaxed problem by dropping the second-order condition in the simplified problem. The solution to the relaxed problem has $k_H = c$. The next lemma shows that the second-order condition in the above simplified problem never binds in the optimal menu of refund contracts.

LEMMA 2.3. *Suppose that $\{k_H, k_L\}$ solves the relaxed problem. Under either FSD or MPS, there exist a_L and a_H such that $\{a_H, k_H, a_L, k_L\}$ is the optimal menu of refund contracts.*

PROOF. Suppose that G_H first-order dominates G_L. Both the surplus $S(k_L)$ from type L and the rent $R(k_L)$ to type H are negative for any $k_L < c$. Thus, the solution to the relaxed problem has $k_L \geq c$. Since $k_H = c$, the solution to the relaxed problem satisfies the second-order condition, and therefore solves the simplified problem. The existence of a_H and a_L follows from IR_L and $IC_{H,L}$.

Suppose that G_H dominates G_L by MPS. If the solution also has $k_L = c$, the lemma follows immediately. Suppose that the solution has $k_L \neq c$, and that the second-order condition is violated:

$$\int_c^{k_L} (G_H(v) - G_L(v))dv > 0.$$

Consider an alternative menu where $k_L = k_H = c$. The surplus $S(k_L)$ from type L is greater in the alternative menu. The rent to type H is smaller in the alternative, because

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

This contradicts the assumption that $\{k_L, k_H\}$ solves the relaxed problem. Thus, the solution to the relaxed problem satisfies the second-order condition, and solves the simplified problem. The existence of a_H and a_L follows. *Q.E.D.*

The following characterization of the optimal menu of refund contracts follows immediately.

PROPOSITION 2.4. *Under either FSD or MPS, in the optimal menu of refund contracts,* $k_H = c$ *and* $k_L = \text{argmax}_k f_L S(k) - f_H R(k)$.

Thus, 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. This result is somewhat unintuitive in the case of MPS. One would think that the seller should ration the business type to extract more profits, because the valuation distribution of the business travelers is more dispersed and it is profitable for the seller to serve them only at the right tail of the distribution. The proof of Lemma 2.3 shows that this intuition is flawed. Rationing the business type is not incentive-compatible, because the leisure type would find it optimal to choose the refund contract with a higher refund (i.e., the second-order condition would be violated).

Under FSD, 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, this reasoning is invalid because the rent $R(k_L)$ is not a monotonically decreasing function of the refund k_L to the leisure type. 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_H 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 $G_H(v) > G_L(v)$ for all $v < z$ and $G_H(v) < G_L(v)$ for all $v > z$. An example of this is normal distributions with the same mean z and greater variance for type H.[6]

LEMMA 2.5. *Suppose $R(\cdot)$ is single-peaked at $z \neq c$. Then, $k_L \notin [\min\{c, z\}, \max\{c, z\}]$.*

PROOF. Suppose that $c < z$ and $k_L \in [c, z]$. By decreasing k_L toward c, one can increase the surplus $S(k_L)$ and decrease the rent $R(k_L)$. Note that setting $k_L = c$ cannot be optimal, since by decreasing k_L slightly below c, surplus from type L is not affected at the margin (because $S'(c) = 0$) but the rent to type H decreases. A similar argument holds when $c > z$. *Q.E.D.*

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. Lemma 2.5 suggests that it is optimal to subsidize (ration) consumption when the cost is low (high). The intuition 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 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.

[6] 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]. We thank Wing Suen for mentioning this result.

PROPOSITION 2.6. *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_H and g_L are symmetric around z; (ii) $f_H R(c) \leq f_L(S(c) - S(z))$.*

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

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

It follows that $S(\tilde{k}_L) > S(k_L)$. Since $R(\tilde{k}_L) = R(k_L)$, the alternative menu yields a greater profit, a contradiction. The argument is similar when $c > z > k_L$. The proposition then follows from Lemma 2.5.

(ii) Suppose that $c < z$. By Lemma 2.5, either $k_L < c$ or $k_L > z$. The profit of setting $k_L < c$ optimally is at least as great as $f_L S(c) - f_H R(c)$, 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 $f_H R(c) \leq f_L(S(c) - S(z))$, setting $k_L > z$ cannot be optimal. A similar argument holds when $c > z$. *Q.E.D.*

Thus, according to Proposition 2.6, the pattern of consumption distortion is determined by the comparison between the cost of the ticket and the peak of the distributions. Note that regardless of whether the cost is low or high, the type that generates more surplus on average ends up consuming more on average. Because they have a tighter distribution, when the cost is low leisure travelers generate more surplus on average and are optimally subsidized with a lower refund, and when the cost is high they generate less surplus on average are optimally rationed with a higher refund.

When neither condition in Proposition 2.6 holds, the pattern of consumption distortions for the leisure type can be different from the predictions of Propositions 2.6. 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. This is more likely, if subsidy is not effective in reducing rent for k_L slightly below the cost because the two distributions differ mostly at the low end of the range. Numerical examples are available from the authors.

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-type *ex ante* distributions 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 under the two-type assumption. Finally, by solving for optimal sequential mechanisms for a number of simple 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 $T = [\underline{t}, \bar{t}]$, with a density function $f(t)$ and cumulative function $F(t)$. Each type t is represented by a distribution of valuations over $[\underline{v}, \bar{v}]$, 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 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]

[7] The optimization problem does not get more complicated without the assumption of common support

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:

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

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

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

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

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

---

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.

[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. Clearly, 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.

$u(t,v) = vy(t,v) - x(t,v)$ 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. The proof 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*.*

Lemma 3.1 amounts to a "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.

Our sequential mechanism design problem is related to static multi-dimensional price discrimination models (e.g., Palfrey [1983]). 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 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

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

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,

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

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

[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. We thank an anonymous referee for pointing out this to us.

IC_2. The sufficiency of the two conditions in Lemma 3.2 requires additional assumptions on T; this will be the subject of section 4.

It is well-known that multi-product price discrimination problems are complex when consumers' private information is multi-dimensional (see Rochet [1995] and Armstrong [1996]). 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.

Under either FSD or MPS described above, the sequential design problem can be further simplified. In particular, in the case of FSD, since a consumer of a higher type has a greater probability of drawing greater valuations, and since by Lemma 3.1 the *ex post* surplus is greater when a consumer draws a greater valuation, a higher type has a greater expected surplus in any incentive compatible sequential mechanism. It can be seen from Lemma 3.2 that under the assumption that T is ordered by FSD, we can ignore the IR constraints for all types except for the lowest type $\underline{t}$. An implication is that in any optimal sequential mechanism, the lowest type earns zero expected surplus. By the argument in Lemma 2.1, the same implication holds for the case of MPS as well.

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 $\max_{y(t,v)} \int_{\underline{t}}^{\overline{t}} \int_{\underline{v}}^{\overline{v}} \phi(t,v)y(t,v)g(v|t)f(t)dvdt$ 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 monotonic 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 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. *Q.E.D.*

In the other case, when T is ordered by MPS with all distribution functions passing through a single point 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 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-increasing) 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 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.

[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.

In a nonlinear pricing problem, the usual second-order condition (analogous to 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. The greater θ is, the more informative the consumer's private type knowledge is about valuation in that conditional distributions of valuations vary more with type.

With the above additive specification, we have

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

Under the standard monotone hazard rate assumption (see, e.g., Fudenberg and Tirole [1991]), the hazard rate $\frac{1-F(t)}{f(t)}$ is non-increasing in t, we have that $\phi(t,v) \geq 0$ implies

[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.

$\phi(t', v') \geq 0$ for any $t' \geq t$ and $v' \geq v$. The solution to the relaxed problem is monotonic 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, a greater variance of ϵ_t means that the consumer faces greater valuation uncertainty, yet there is 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. The shock ϵ_t is common to all consumer types and may be interpreted as demand shocks. Proposition 3.5 implies that additive demand shocks do not affect the optimal sequential mechanism.[14]

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.

The additive specification can provide a tractable way of testing our model. One property already mentioned is that the optimal menu of refund contract does not depend on the specification $H(\cdot)$ of the valuation shock ϵ_t. This allows some freedom in specifying the conditional distributions. Also, we can relax the assumption that consumers differ linearly in their expected valuation. Suppose that valuation is linked to type through $v = \theta n(t) + (1-\theta)\epsilon_t$ where $n(t)$ is a nonlinear function. Assume that $n(t)$ is increasing so that first-order stochastic dominance is satisfied. In this generalized additive case, informativeness measure of type about valuation equals $\theta n'(t)$ and varies across types. It

[14] The log-linear specification and the multiplicative MPS specification discussed below show that under some mild assumptions on the type distribution, multiplicative demand shocks have no effects on the optimal sequential mechanism either. In general, demand shocks have no effects as long as conditions on the type distribution can be found to satisfy the assumptions in Lemma 3.3 and Lemma 3.4.

is straightforward to verify that a similar result to Proposition 3.5 holds if $n(t)$ is not too convex, or expected valuation does not increase too fast with type.

We can also extend Proposition 3.5 to a multiplicative specification.[15] 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. 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 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 monotonic 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. More precisely, suppose

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

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

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

[15] A specification with a special distribution that works out similarly is $G(v|t) = 1 - \exp(-v/t)$. We thank an anonymous referee for the suggestion.

It is easy to see that the distributions $G(v|t)$ satisfy MPS and pass through the same point 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 relatively 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 is a generalization of Proposition 2.6. 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, note that when valuation is a known mean plus a multiplicative shock, more informative type knowledge can be represented by a leftward shift of the distribution of types $F(t)$. Since $\lambda(t)$ is decreasing in t, this amounts to increasing the function $\lambda(t)$. 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. These include refund contracts, option contracts, and 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.

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 one 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. This insufficiency results from the multi-dimensional nature of the sequential mechanism design problem. To see this, note that from Lemma 3.3, 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.

Localization of incentive compatibility constraint is important. With rare exceptions,[16] the works in the price discrimination literature take a local approach. Localization is guaranteed in the standard one-dimensional price discrimination literature under the familiar "single crossing" condition that marginal rate of substitution is ordered by type (see, e.g., Cooper [1984]). This single crossing condition is satisfied for the period-two incentive compatibility constraints because a consumer with a greater valuation is willing to pay more for an increase in delivery probability. It enables us to replace IC_2 constraints by two local conditions in Lemma 3.1. In the multi-dimensional price discrimination literature, McAfee and McMillan [1988] have found a condition that guarantees that local incentive compatibility constraints imply global constraints (they call it "generalized single crossing" condition). But their condition requires that the dimension of consumer's private information exceed the number of monopolist' price discrimination instruments. This

[16] Matthews and Moore [1987] is one of them.

dimensionality condition is not 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.

A special case of FSD called "alignment" imposes a linear structure on the type space and guarantees that the two local conditions in Lemma 3.2 are sufficient for IC_1 constraints.[17] The meaning of alignment is best seen in the discrete formulation (the discrete notation is self-explanatory, for example, $G_{i,j}$ corresponds to $\sum_{k=1}^{j-1} g_{i,k}$). We say that three types t_i, $t_{i'}$ and $t_{i''}$ are aligned if they are ordered by FSD and

$$\frac{G_{i,j} - G_{i',j}}{G_{i,j'} - G_{i',j'}} = \frac{G_{i',j} - G_{i'',j}}{G_{i',j'} - G_{i'',j'}}$$

for all valuations v_j and $v_{j'}$ such that the above is defined. We say that the type space T is aligned if any three types are aligned. Alignment has a geometric interpretation. Each type can be identified as a point in a $J-1$ dimensional vector space, with coordinates $(G_{i,2}, G_{i,3}, \ldots, G_{i,J})$, where J is the number of possible valuations. Three types are aligned if the corresponding points form a line. Alignment in the continuous case can be defined in a similar way.

The following result shows how the assumption of alignment ensures that the two local conditions in Lemma 3.2 are sufficient for IC_2 constraints. The proof is in the appendix. The key is that under alignment the weights in the local period-one second order condition do not change with type. Thus, local IC_2 constraints extend globally.

LEMMA 4.1. *Suppose that the type space is aligned. Then a sequential mechanism satisfies period-one incentive compatibility constraints if and only if it satisfies the two local conditions in Lemma 3.2.*

The assumption of alignment is certainly restrictive, but it enables us to use a variation of the standard local approach and provides some insights about the optimal mechanism's properties. Moreover, one can show that in some sense alignment is the only case where the local approach applies: if T is not aligned, then there is a sequential mechanism

[17] Alignment for MPS (with the same restrictions as in Proposition 3.6) can be similarly defined.

which satisfies all the local period-one incentive constraints (and individual rationality and period-two constraints) but fails at least some global period-one constraints.

The next result states that under alignment over-production never occurs in an optimal sequential mechanism.[18] The proof relies on the observation that both over-production and under-production can be used for the purpose of price discrimination. Under alignment there is never any need to use over-production as an instrument of price discrimination because it is more costly to the monopolist than under-production. The assumption of alignment is crucial for the proof; without it, over-production can occur under stochastic dominance. The proof is in the appendix.

LEMMA 4.2. *Assume T is aligned. Then in any optimal sequential mechanism, there is no over-production for any type.*

Although the assumption of alignment allows a local approach, one cannot conclude that the delivery rule is monotone as in Lemma 3.3. That is, "bunching" may occur both across valuations and across types.[19] This makes it difficult to characterize the optimal mechanism. However, the assumption of alignment limits the extent of bunching.

PROPOSITION 4.3. *Assume that T is aligned. Then in any optimal mechanism, the contract for each type can be described by at most three different probabilities, 0, 1 and some number in between.*

An example in the appendix with three types and three valuations shows how bunching occurs across types and across valuations at the same time and results in a random delivery rule. Random delivery rules allow fine tuning by the monopolist. When the optimal mechanism involves randomization, the delivery rule may not be increasing in type for each valuation. Only the weighted average of the delivery probabilities increases with type. That the delivery rule need not be monotonic in types is similar to the conclusion

---

[18] The counterpart of this result for the case of MPS is that under alignment over-production does not occur when production cost is high and under-production does not occur when the cost is low.

[19] See, e.g., the appendix of Chapter 7 in Fudenberg and Tirole [1991] for an explanation of bunching techniques in standard nonlinear pricing problems.

in the multi-product price discrimination literature that quantity or quality of each good need not be monotonic when consumer demand characteristics are one-dimensional but the number of instruments of price discrimination 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.

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. 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. 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. While the restriction to *ex post* two-part tariffs can be shown to be without loss of generality, in general *ex ante* two-part tariffs are 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 (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.

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.

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.

Appendix

This appendix uses the discrete formulation to prove Lemma 4.1, Lemma 4.2, Proposition 4.3, and to present an example of randomization under alignment. Similar arguments apply to the continuous setting. Let $T = \{t_1, \ldots, t_I\}$ be the set of possible types, where f_i is the probability of $t_i \in T$, and $V = \{v_1, \ldots, v_J\}$ be the set of possible valuations (ordered from the smallest valuation to the greatest), where $g_{i,j}$ is the (strictly positive) probability of type t_i drawing valuation v_j (with $G_{i,j} = \sum_{k=1}^{j-1} g_{i,k}$). Denote $x_{i,j}$ as the payment to the monopolist when the consumer reports type t_i and then valuation v_j, and $y_{i,j}$ as the corresponding delivery probability. Let $u_{i,j} = v_j y_{i,j} - x_{i,j}$. A typical second-period incentive compatibility constraint $IC^i_{j,j'}$ is $v^j y_{i,j} - x_{i,j} \geq v^j y_{i,j'} - x_{i,j'}$. Let $\bar{j} = \max\{j | v_j < c\}$. To make the discrete exposition in this appendix self-contained, we first provide the proof of the discrete version of Lemma 3.1.

Lemma 3.1 *In any optimal mechanism $IC^i_{j+1,j}$ holds with equality for each i and each $j \geq \bar{j}+1$, and $IC^i_{j-1,j}$ holds with equality for each i and $j \leq \bar{j}$. All $IC^i_{j,j'}$ conditions are satisfied if and only if $y_{i,j+1} \geq y_{i,j}$ for each $j \neq \bar{j}$, and $IC^i_{\bar{j}+1,\bar{j}}$ and $IC^i_{\bar{j},\bar{j}+1}$ hold.*

Proof. Summing up $IC^i_{j,j'}$ and $IC^i_{j',j}$ and rearranging terms, we find that $\forall i$ and $j > j'$, $y_{i,j} \geq y_{i,j'}$, $x_{i,j} \geq x_{i,j'}$, and $u_{i,j} \geq u_{i,j'}$. Next, for any i and $j > j' > j''$, summing up $IC^i_{j,j'}$ and $IC^i_{j',j''}$ and noting $y_{i,j'} \geq y_{i,j''}$, we have $IC^i_{j,j'}$ and $IC^i_{j',j''}$ imply $IC^i_{j,j''}$. Similarly, $IC^i_{j'',j'}$ and $IC^i_{j',j}$ imply $IC^i_{j'',j}$. It follows that if $\bar{j} \leq J-1$, then in any optimal sequential mechanism, $y_{i,J} = 1$ for all i, and if $\bar{j} \geq 1$, then in any optimal sequential mechanism, $y_{i,1} = 0$ for all i.

Now we can show that $IC^i_{j-1,j}$ holds with equality for each i and $j \leq \bar{j}$. First, since $y_{i,1} = 0$, if $y_{i,j} = 0$, then for all $j' \leq j$, we have $y_{i,j'} = 0$ and $x_{i,j'} = x_{i,1}$, and therefore $IC^i_{j,j-1}$ holds with equality. Assume that $y_{i,j} > 0$. By way of contradiction, suppose that $IC^i_{j,j-1}$ holds with strict inequality. Consider an alternative mechanism that coincides with the original mechanism except $\tilde{y}_{i,j} = y_{i,j} - \epsilon$ and $\tilde{x}_{i,j} = x_{i,j} - v_j\epsilon$, where $\epsilon > 0$ is to be determined below. Note that it is possible to decrease $y_{i,j}$ because by assumption $y_{i,j} > 0$. Then $IC^i_{j-1,j}$ becomes

$$v_{j-1}y_{i,j-1} - x_{i,j-1} \geq v_{j-1}\tilde{y}_{i,j} - \tilde{x}_{i,j} = v_{j-1}y_{i,j} - x_{i,j} + (v_j - v_{j-1})\epsilon.$$

Since by assumption $IC^i_{j-1,j}$ holds with strict inequality, we can choose ϵ appropriately so that $IC^i_{j-1,j}$ holds with equality under the alternative mechanism. Since by construction $\tilde{u}_{i,j} = u_{i,j}$, $IC^i_{j,j+1}$ and $IC^i_{j,j-1}$ still hold, while $IC^i_{j+1,j}$ becomes

$$v_{j+1}y_{i,j+1} - x_{i,j+1} \geq v_{j+1}\tilde{y}_{i,j} - \tilde{x}_{i,j} = v_{j+1}y_{i,j} - x_{i,j} + (v_j - v_{j+1})\epsilon,$$

which is also satisfied under the alternative mechanism. Thus, all period-two incentive constraints are satisfied. Since by construction $\tilde{u}_{i,j} = u_{i,j}$, all period-one incentive constraints and participation constraints still hold. However, the expected profits of the monopolist under the alternative mechanism are greater than under the original mechanism by $f_i g_{i,j}(c - v_j)\epsilon$, a contradiction. A similar argument shows that in any optimal mechanism $IC^i_{j+1,j}$ holds with equality for each i and each $j \geq \bar{j} + 1$.

For the second statement of the lemma, it suffices to show that if $IC^i_{j+1,j}$ holds with equality, then $IC^i_{j,j+1}$ holds if and only if $y_{i,j+1} \geq y_{i,j}$. Since $IC^i_{j+1,j}$ holds with equality,

$$v_j y_{i,j} - x_{i,j} = v_j y_{i,j+1} - x_{i,j+1} + (v_{j+1} - v_j)(y_{i,j+1} - y_{i,j}).$$

Therefore, $IC^i_{j,j+1}$ is satisfied if and only if $y_{i,j+1} \geq y_{i,j}$. *Q.E.D.*

A.1. Proof of Lemma 4.1

Let $U_{i,i'} = \sum_{j=1}^{J} g_{i,j} u_{i',j}$. To simplify notation, for all i and j define

$$S_{i,j} = \begin{cases} (v_{j+1} - v_j) \sum_{k=j+1}^{J} g_{i,k} & \text{if} \quad J - 1 \geq j \geq \bar{j} + 1 \\ -(v_j - v_{j-1}) G_{i,j} & \text{if} \quad 2 \leq j \leq \bar{j}, \end{cases}$$

and $w_i = u_{i,\bar{j}+1} - u_{i,\bar{j}}$.

Lemma 4.1 *Assume T is aligned. Then in optimal sequential mechanism, $IC_{i,i-1}$ holds with equality for each $i \geq 2$. Moreover, all $IC_{i,i'}$ constraints are satisfied if and only if*

$$(M_{i,i-1}) \qquad -(G_{i,\bar{j}+1} - G_{i',\bar{j}+1})(w_i - w_{i'}) + \sum_{j=2}^{J-1} (S_{i,j} - S_{i',j})(y_{i,j} - y_{i',j}) \geq 0.$$

Proof. By applying Lemma 3.1, we can write

$$U_{i,i'} = u_{i',\bar{j}+1} - G_{i,\bar{j}+1} w_{i'} + \sum_{j=2}^{J-1} S_{i,j} y_{i',j}.$$

Then the profit maximization problem can be written as:

$$\max_{y_{i,j}, u_{i,\bar{j}+1}, w_i} \sum_{i=1}^{I} f_i \left(\sum_{j=1}^{J} (v_j - c) g_{i,j} y_{i,j} - U_{i,i} \right)$$

subject to constraints $w_i \geq (v_{\bar{j}+1} - v_{\bar{j}}) y_{i,\bar{j}}$ for all i ($IC^i_{\bar{j}+1,\bar{j}}$), $w_i \leq (v_{\bar{j}+1} - v_{\bar{j}}) y_{i,\bar{j}+1}$ for all i ($IC^i_{\bar{j},\bar{j}+1}$), $y_{i,j+1} \geq y_{i,j}$ for all i and $j \neq \bar{j}$ ($M^i_{j+1,j}$), $U_{i,i} \geq U_{i,i'}$ for all i and i' ($IC_{i,i'}$), $U_{i,i} \geq 0$ for all i (IR_i), and $0 \leq y_{i,j} \leq 1$ for all i and j ($R_{i,j}$).

First we show that if t_i, $t_{i'}$, and $t_{i''}$ are aligned and $i > i' > i''$, then $IC_{i,i'}$, $IC_{i',i''}$, and $IC_{i'',i'}$ imply $IC_{i,i''}$.

$$
\begin{aligned}
&U_{i,i} - U_{i,i''} \\
=&u_{i,\bar{j}+1} - u_{i',\bar{j}+1} + u_{i',\bar{j}+1} - u_{i'',\bar{j}+1} - G_{i,\bar{j}+1}(w_i - w_{i''}) + \sum_{j=2}^{J-1} S_{i,j}(y_{i,j} - y_{i'',j}) \\
\geq& -(G_{i,\bar{j}+1} - G_{i',\bar{j}+1})(w_{i'} - w_{i''}) + \sum_{j=2}^{J-1}(S_{i,j} - S_{i',j})(y_{i',j} - y_{i'',j}) \\
=&\frac{G_{i',2} - G_{i,2}}{G_{i'',2} - G_{i',2}} \left(-(G_{i',\bar{j}+1} - G_{i'',\bar{j}+1})(w_{i'} - w_{i''}) + \sum_{j=2}^{J-1}(S_{i',j} - S_{i'',j})(y_{i',j} - y_{i'',j}) \right),
\end{aligned}
$$

where the last line uses the assumption that i, i' and i'' are aligned, and the second-to-last uses $IC_{i,i'}$ and $IC_{i',i''}$. Adding up $IC_{i',i''}$ and $IC_{i'',i'}$ then implies that $U_{i,i} \geq U_{i,i''}$. Similarly, $IC_{i'',i'}$, $IC_{i',i}$, and $IC_{i,i'}$ imply $IC_{i'',i}$. The two statements of the lemma then follow similar arguments as in Lemma 3.1. *Q.E.D.*

A.2. Proof of Lemma 4.2

LEMMA 4.2. *Assume that T is aligned. Then in any optimal sequential mechanism, $y_{i,j} = 0$ for all i and $j \leq \bar{j}$, and $w_i = 0$ for all i.*

PROOF. Let $\delta_{i,j} = f_i(v_j - c)g_{i,j} - (S_{i+1,j} - S_{i,j}) \sum_{k=i+1}^{I} f_k$. By Lemma 4.1, the profit maximization problem can be written as

$$
\max_{y_{i,j}, w_i} \sum_{i=1}^{I} \sum_{j=2}^{J-1} \delta_{i,j} y_{i,j} + \sum_{i=1}^{I-1} (G_{i+1,\bar{j}+1} - G_{i,\bar{j}+1}) w_i \sum_{k=i+1}^{I} f_k
$$

subject to constraints $IC^i_{\bar{j}+1,\bar{j}}$, $IC^i_{\bar{j},\bar{j}+1}$, $M_{i,i-1}$, $M^i_{j+1,j}$, and $R_{i,j}$.

The proof is by induction. First, $y_{1,j} = 0$ for all $j \leq \bar{j}$. Moreover, we can show by contradiction $w_1 = 0$. Suppose not and consider decreasing it by ϵ. This would only affect constraints $M_{2,1}$, $IC^1_{\bar{j}+1,\bar{j}}$, and $IC^1_{\bar{j},\bar{j}+1}$, but since $y_{1,\bar{j}} = 0$, all three constraints remain satisfied. However, the monopolist would obtain greater expected profits. Thus, $w_1 = 0$ in any optimal mechanism.

Next, we assume that $y_{i,j} = 0$ for all $i \leq k - 1$ and $j \leq \bar{j}$, and $w_i = 0$ for all $i \leq k-1$. We want to show that $y_{k,j} = 0$ for all $j \leq \bar{j}$, and $w_k = 0$. Suppose $y_{k,j'} > 0$ for some $j' \leq \bar{j}$. Without loss of generality, we can assume that $v_{j'}$ is the smallest such valuation. Since $j' \leq \bar{j}$, we have $\delta_{k,j'} < 0$. If $M_{k,k-1}$ does not hold with equality, the monopolist could decrease $y_{k,j'}$ by ϵ so that $M_{k,k-1}$ is still satisfied. This would only affect constraints $M_{k+1,k}$ and $M^k_{j',j'-1}$, and possibly $IC^k_{\bar{j}+1,\bar{j}}$, but it is straightforward to check that all three constraints remain satisfied. However, the monopolist would obtain greater expected profits. Thus, $M_{k,k-1}$ holds with equality.

Since $y_{k-1,j} = 0$ for all $j \leq \bar{j}$ and $w_{k-1} = 0$ by the induction assumption, there exists $j'' \geq \bar{j} + 1$ such that $y_{k,j''} < 1$. Without loss of generality, we can assume that $v_{j''}$ is the largest such valuation. Consider decreasing $y_{k,j'}$ by ϵ' and increasing $y_{k,j''}$ by ϵ'' such that $y_{k,j'} - \epsilon' > 0$, $y_{k,j''} + \epsilon'' < 1$, and

$$\epsilon'(v_{j'} - v_{j'-1})(G_{k-1,j'} - G_{k,j'}) = \epsilon''(v_{j''+1} - v_{j''})(G_{k-1,j''+1} - G_{k,j''+1}).$$

We can check that $M_{k,k-1}$ and $M_{k+1,k}$ are unaffected by these changes. Moreover, since $j' \leq \bar{j}$ and $j'' \geq \bar{j} + 1$, constraints $IC^k_{\bar{j}+1,\bar{j}}$ and $IC^k_{\bar{j},\bar{j}+1}$ remain satisfied. The expected profits are changed by

$$\begin{aligned}
& - \delta_{k,j'}\epsilon' + \delta_{k,j''}\epsilon'' \\
=& f_k(\epsilon'(c - v_{j'})g_{k,j'} + \epsilon''(v_{j''} - c)g_{k,j''}) \\
& + (\epsilon'(v_{j'} - v_{j'-1})(G_{k,j'} - G_{k+1,j'}) - \epsilon''(v_{j''+1} - v_{j''})(G_{k,j''+1} - Q_{k+1,j''+1})) \sum_{k'=k+1}^{I} f_{k'}.
\end{aligned}$$

Using the definition of alignment, we can show that the above expression is equal to $f_k(\epsilon'(c - v_{j'})g_{k,j'} + \epsilon''(v_{j''} - c)g_{k,j''})$, which is positive. This contradicts assumed optimality. Therefore, $y_{k,j} = 0$ for all $j \leq \bar{j}$. A similar argument shows that $w_k = 0$. *Q.E.D.*

A.3. Proof of Proposition 4.3

PROPOSITION 4.3. *Assume that T is aligned. Then in any optimal mechanism, there is no i such that $y_{i,j} \neq y_{i,j'} \in (0, 1)$ for some j and j'.*

PROOF. Let $\{y_{i,j}\}$ be an optimal mechanism. For each i such that $2 \leq i \leq I - 1$, consider the following problem: $\max_{y'_{i,j}} \sum_{j=\bar{j}+1}^{J-1} \delta_{i,j} y'_{i,j}$ subject to two period-one monotonicity constraints $M_{i,i-1}$ and $M_{i+1,i}$, the period-two monotonicity constraints $M^i_{j+1,j}$, and the randomization constraints $R_{i,j}$. Since $\{y_{i,j}\}$ is an optimal mechanism, for any i such that $2 \leq i \leq I - 1$, $y_{i,j}$ is a solution to the above maximization problem. For $i = 1$, the counterpart of the problem is given by $\max_{y'_{1,j}} \sum_{j=\bar{j}+1}^{J-1} \delta_{i,j} y'_{1,j}$ subject to a single period-one monotonicity constraints $M_{2,1}$, the period-two monotonicity constraints $M^1_{j+1,j}$, and the randomization constraints $R_{1,j}$.

Let $z_{i,\bar{j}+1} = y_{i,\bar{j}+1}$ and define $z_{i,j} = y'_{i,j} - y'_{i,j-1}$ recursively for each j such that $\bar{j} + 2 \leq j \leq J - 1$. Then, $y'_{i,j} = \sum_{k=\bar{j}+1}^{j} z_{i,k}$ for any j such that $\bar{j} + 2 \leq j \leq J - 1$. For any i such that $2 \leq i \leq I - 1$, we can write the above maximization problem as a linear programing problem with only three constraints: two period-one monotonicity constraints and one randomization constraint $\sum_{k=\bar{j}+1}^{J-1} z_{i,k} \leq 1$. The period-two monotonicity constraints become the implicit non-negativity constraint on each $z_{i,j}$. Note that under the assumption of alignment, the two period-one monotonicity constraints are linearly dependent. Therefore, the solution to the linear programming problem has at most two non-zero variables, implying that at most two increments $z_{i,j}$ are non-zero. This establishes the proposition for any i such that $2 \leq i \leq I-1$. For $i=1$, the counterpart of the above linear programming problem has only two constraints, one corresponding to $M_{2,1}$ and the other corresponding to $\sum_{k=\bar{j}+1}^{J-1} z_{1,k} \leq 1$. Similar reasoning then establishes the proposition for $i=1$. *Q.E.D.*

A.4. An example of randomization under alignment

The simplest case of alignment that permits randomization is $I=J=3$. Suppose production cost is zero. ¿From the proof of Proposition 4.3, randomization occurs for type t_i only if one of the two local period-one second-order conditions $M_{i+1,i}$ and $M_{i,i-1}$ holds with equality. Since there is no under-production distortion for type t_3, randomization occurs if and only if $M_{2,1}$ holds with equality in an optimal sequential mechanism:

$$y_{1,1} + s y_{1,2} = y_{2,1} + s y_{2,2},$$

where for simplicity we denote

$$s = \frac{(v_3 - v_2)(G_{1,3} - G_{2,3})}{(v_2 - v_1)(G_{1,2} - G_{2,2})}.$$

In any optimal sequential mechanism, randomization cannot occur for type t_1 and t_2 at the same time. If there is randomization for both t_1 and t_2, there must be $j = 1, 2$ such that randomization occurs for valuation j for both t_1 and t_2, for otherwise $M_{2,1}$ cannot be satisfied with equality without violating local period-two second-order conditions. But then it would be possible to increase profits by either increasing or decreasing $y_{1,j}$ and $y_{2,j}$ appropriately at the same time so as to satisfy $M_{2,1}$.

Thus, two cases of randomization in optimal sequential mechanism are possible: i) $y_{1,1} = y_{1,2} \in (0,1)$, $y_{2,1} = 0$, and $y_{2,2} = 1$; and ii) $y_{2,1} = y_{2,2} \in (0,1)$, $y_{1,1} = 0$, and $y_{1,2} = 1$. We discuss the first case only; the second case is symmetric. The following conditions are necessary for randomization to be optimal; otherwise the delivery rule can be modified to increase profits without violating period-two incentive constraints or $M_{2,1}$.

1. $\delta_{1,1} \geq 0$. Otherwise, reduce $y_{1,1}$.
2. $\delta_{1,1} + \delta_{1,2} \geq 0$. Otherwise, reduce $y_{1,1}$ and $y_{1,2}$ by the same amount.
3. $s\delta_{1,1} \geq \delta_{1,2}$. Otherwise, increase $y_{1,2}$ by some small positive number ϵ and decrease $y_{1,1}$ by $s\delta_{1,1}$.
4. $\delta_{1,1} + \delta_{1,2} + (1+s)\delta_{2,1} \leq 0$. Otherwise, increase $y_{1,1}$ and $y_{1,2}$ by the same amount ϵ and increase $y_{2,1}$ by $(1+s)\epsilon$.
5. $\delta_{1,1} + \delta_{1,2} + \delta_{2,2}(1+s)/s \geq 0$. Otherwise, decrease $y_{1,1}$ and $y_{1,2}$ by the same amount ϵ and decrease $y_{2,2}$ by $\epsilon(1+s)/s$.
6. $\delta_{2,2} \geq s\delta_{2,1}$. Otherwise, decrease $y_{2,2}$ by ϵ and increase $y_{2,1}$ by $s\epsilon$.

Note that the last condition is implied by the fourth and fifth. When at least one of the first five conditions holds with strict inequality, these conditions are also sufficient for randomization to occur in optimal sequential mechanism. Numerical examples are available from the authors showing that parameters can be found to satisfy the five conditions. The optimal sequential mechanisms in these examples have random delivery rules for type t_1.

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.

Pascal Courty [1996]: "The Economics of Tickets," Ph. D. Dissertation, University of Chicago.

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.

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.

Thomas Palfrey [1983]: "Bundling Decisions by a Multiproduct Monopolist with Incomplete Information," *Econometrica* 51, 463-484.

Jean-Charles Rochet [1995]: "Ironing, Sweeping and Multidimensional Screening," IDEI, working paper.

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.