Pareto-Improving Segmentation of Multiproduct Markets

Nima Haghpanah

Pennsylvania State University

Ron Siegel

Pennsylvania State University

We investigate whether a market served by a multiproduct monopolistic seller can be segmented in a way that benefits all consumers and the seller. The seller can offer a different product menu in each market segment, combining second- and third-degree price discrimination. We show that markets for which profit maximization leads to inefficiency can, generically, be segmented into two market segments in a way that increases the surplus of all consumers weakly and of some consumers and the seller strictly. Our constructive proof is based on deriving implications of binding incentive compatibility constraints when profit maximization implies inefficiency.

I. Introduction

Market segmentation and price discrimination are common practices that benefit sellers but may harm consumers. In some cases, such as first-degree price discrimination, all consumers are harmed. In other cases, certain consumers are harmed, while other consumers benefit from lower prices

We thank the editor, Emir Kamenica, and three referees for detailed comments that significantly improved the paper. We thank Nageeb Ali, Dirk Bergemann, Wouter Dessein, Harry Di Pei, Piotr Dworzcak, Wioletta Dziuda, Teddy Kim, Stephen Morris, Alessandro Pavan, Eduardo Perez Richet, Andy Skrzypacz, Asher Wolinsky, and Jidong Zhou for very helpful comments and suggestions. We also thank the audiences in various seminars and conferences.

Electronically published April 17, 2023

Journal of Political Economy, volume 131, number 6, June 2023.

https://doi.org/10.1086/722932

or more suitable products. This paper investigates which markets can be segmented in a way that benefits all consumers and the seller when the seller can offer multiple products and maximizes profit in each market segment.

Our primary motivation in asking this question is to improve our understanding of the welfare effects of market segmentation and price discrimination in multiproduct environments. But the question we ask may also be relevant to regulatory discussions regarding consumer privacy and sellers' use of consumer data. A regulator interested in increasing consumer welfare may be able to control the data that sellers collect or access or the scope and type of targeted offers that sellers can make. Alternatively, consumers may be able to decide what data to provide to sellers.[1] As a report by the Federal Trade Commission (2012) puts it, "The Commission recognizes the need for flexibility to permit . . . uses of data that benefit consumers." Our analysis clarifies for which markets there exists some data that—if provided to or collected by a profit-maximizing seller—will be used by the seller to price discriminate in a way that benefits all consumers.

We consider a setting in which a multiproduct monopolistic seller faces a market of heterogeneous consumers with preferences over subsets of products. Consumer preferences are quasi-linear in money but are otherwise quite general.[2] If the seller can segment the market, she can offer a potentially different menu of products and product bundles in each market segment, thereby combining second- and third-degree price discrimination. Otherwise, the seller offers the same menu to all consumers.

Our focus is on *Pareto-improving* segmentations, in which the surplus every consumer obtains when choosing from the menu that the seller offers in his segment is no lower than the surplus the consumer would obtain in the unsegmented market and is strictly higher for some consumers and the seller. If a Pareto-improving segmentation exists for a market, we say that the market is *Pareto improvable.* Because any Pareto-improving segmentation increases total surplus, any Pareto improvable market is necessarily inefficient in that its profit-maximizing menu leads to some inefficiency.

Our main result is that, generically, inefficient markets are Pareto improvable by a segmentation with two market segments.[3] In other words, whenever total surplus is not maximized, market segmentation can benefit all consumers and the seller. This suggests that properly regulated data collection and usage can have unambiguously positive welfare effects in a wide range of market settings.

A natural approach to proving our result is to consider, for each market and each segmentation, each consumer's surplus in the seller's

[1] In the single-agent interpretation of our model, discussed in app. A2, the agent can commit to an information disclosure policy prior to learning his type, as in Ichihashi (2020).

[2] Importantly, however, we assume that the number of products and consumer types is finite. Section VI discusses the implications of increasing the number of products.

[3] We formalize our notion of genericity in sec. IV.

profit-maximizing menu. However, no characterization of profit-maximizing menus exists in our multiproduct environment. We therefore develop a different approach. This novel approach is based on understanding the interaction between binding incentive constraints and what drives market inefficiency: the only reason a seller serves some consumers inefficiently is to reduce the information rents of other consumers.

This simple observation has far-reaching implications. We show that for every inefficient market there is an efficient *Pareto-dominating* market: every consumer in this market weakly prefers—and some strictly prefer—the menu that maximizes the seller's profit in this market to the profit-maximizing menu in the inefficient market. The seller's profit from the consumers in the efficient market also increases. Our constructive proof shows that the Pareto-dominating market may have to include numerous consumer types.

The Pareto-dominating market forms one of the two segments in the Pareto-improving segmentation. We then show that for every market with a unique optimal payment rule, induced by the profit-maximizing menu, a small perturbation of the market does not change the profit-maximizing menu and that the set of such markets is generic. For the generic set of inefficient markets with a unique optimal payment rule, therefore, the two-market segmentation that consists of a small fraction of the efficient Pareto-dominating market and the large fraction of the remaining consumers is Pareto improving.

The rest of the paper is organized as follows. Section II discusses the related literature. Section III describes the model and provides an example that illustrates the various concepts. Section IV derives the main result. Section V presents some special cases and applications. Section VI discusses several aspects of our model and analysis, including the magnitude of the improvements, and concludes. The appendix includes proofs not given in the main text, a single-agent interpretation of the model, an additional application, and an example with a large number of products.

II. Related Literature

Our work connects second- and third-degree price discrimination. The literature that studies third-degree price discrimination and its effects on producer and consumer surplus is broad. Pigou (1920) provides examples in which a segmentation may decrease total and hence consumer surplus. Follow-up work provides conditions for a segmentation to increase or decrease total surplus or consumer surplus (Robinson 1969; Schmalensee 1981; Varian 1985; Aguirre, Cowan, and Vickers 2010; Cowan 2016). Our work differs from this literature in three significant ways. First, with third-degree price discrimination, the seller offers a single product to all consumers in a market, whereas the seller in our setting may offer a menu of products. Second, instead of considering expected consumer surplus, we use the Pareto criterion. Third, most of the literature assumes that the segmentation is exogenously fixed.

A recent literature on third-degree price discrimination studies surplus across all possible segmentations of a given market. Bergemann, Brooks, and Morris (2015) identify the set of producer and consumer surplus pairs that result from all segmentations of a given market. It follows from their analysis that in environments with a single product, any inefficient market can be segmented in a way that is Pareto improving (see proposition 1). Glode, Opp, and Zhang (2018) study optimal disclosure by an informed agent in a bilateral trade setting and show that the optimal disclosure policy leads to socially efficient trade, even though information is revealed only partially. Ichihashi (2020) and Hidir and Vellodi (2021) consider maximum consumer surplus when a multiproduct seller offers a single product in each market segment. Ichihashi (2020) considers a finite number of products and compares two regimes, one in which the seller may offer the same product at different prices to different segments and another one in which the seller fixes the price in advance. Hidir and Vellodi (2021) characterize optimal segmentations with a continuum of products. Braghieri (2019) studies market segmentation with a continuum of firms, each producing a single differentiated product. In contrast to these papers, the seller in our setting may offer multiple products in each market segment. Pram (2021) and Haghpanah and Siegel (2022) also allow the seller to offer multiple products in each market segment. Pram (2021) shows that under a single-crossing assumption, a market in which is it profitable to exclude some consumers is Pareto improvable. This insight is also present in the single-product setting of Ali, Lewis, and Vasserman (2022). Haghpanah and Siegel (2022) identify markets for which the entire surplus triangle of Bergemann, Brooks, and Morris (2015) is achievable and markets for which the highest consumer surplus in the surplus triangle is achievable. Finally, Bergemann, Brooks, and Morris (2015) provide a parametric example with two types and nonlinear valuations in which the seller sometimes offers more than one product in a segment.

Our model can also be cast in a Bayesian persuasion framework (Kamenica and Gentzkow 2011) with a single consumer, the sender, who faces the seller, the receiver (app. A2 provides details). However, techniques from that literature, such as concavification and the duality approach of Dworczak and Martini (2019), are not applicable to our setting for two reasons. First, whereas the usual persuasion settings consider the agent's expected utility, we consider the agent's ex post utility. Second, and more importantly, these techniques require a specification of the sender's utility for inducing any given posterior. In our setting, the consumer's utility depends on the seller's optimal menu, for which no characterization exists when there are multiple products.

III. Setup

A monopolistic seller faces a continuum of consumers (app. A2 discusses the interpretation of a single consumer). The *environment* includes a set T of n consumer types and a finite set A of alternatives, where alternative $0 \in A$ is consumers' outside option. We will refer to $k = |A| - 1$ as the number of alternatives (excluding the outside option). Each consumer type specifies a valuation for every alternative: type t's valuation for alternative a is $v(t, a)$. Type t's valuation for a random alternative $x \in \Delta(A)$ is $v(t, x) = E_{a \sim x}[v(t, a)]$. Type t's surplus from a random alternative x and payment p to the seller is $v(t, x) - p$. The valuation for the outside option is zero for all types, that is, $v(t, 0) = 0$. The seller's cost of producing each alternative is normalized to zero without loss of generality.[4] We assume that each type t has a unique efficient alternative $\bar{a}(t) \neq 0$ that maximizes the type's valuation over all alternatives.[5] Different consumer types may rank the alternatives differently, and consumers' valuations need not be ordered by their types or satisfy a condition like increasing differences.

Each alternative $a \neq 0$ corresponds to a product or a set of products. This captures horizontal and vertical differentiation, allows for multiunit demand, and accommodates bundling. To illustrate this, suppose that the seller can produce two products, 1 and 2, and product 2 has a low-quality version L and a high-quality version H. Suppose that consumers may want to buy one or both products but not both versions of product 2. This setting can modeled by an environment with six alternatives, which correspond to the relevant subsets of $\{1, L, H\}$: $0, \{1\}, \{L\}, \{H\}, \{1, L\}, \{1, H\}$. Alternatively, we could specify an alternative for every subset of $\{1, L, H\}$ and reflect in consumers' valuations the fact that consumers do not want to buy both versions of product 2. If instead some consumers demand multiple units of a single product, that would be captured by additional alternatives.

An *allocation rule* $x : T \to \Delta(A)$ is a mapping from types to random alternatives, where $x(t)$ is the allocation of type t. A (direct) *mechanism* $M = (x, p)$ consists of an allocation rule x and a *payment rule* $p : T \to \mathbb{R}_+$.[6] A mechanism is incentive compatible (IC) if no type benefits from misreporting, that is,

$$v(t, x(t)) - p(t) \geq v(t, x(t')) - p(t')$$

[4] A nonzero cost $c(a)$ for alternative $a \neq 0$ can be accommodated by redefining valuations as $\tilde{v}(t, a) = v(t, a) - c(a)$ without changing the analysis or results. Notice that $\tilde{v}(t, a)$ may be negative even if all valuations $v(t, a)$ are nonnegative. Thus, throughout the paper, we allow for negative valuations.

[5] The assumption that $\bar{a}(t) \neq 0$ is without loss of generality because we can remove from the environment any type whose unique efficient alternative is 0.

[6] The restriction to nonnegative payments is without loss of generality since negative payments are never optimal for the seller.

for all types t and t'. A mechanism is individually rational (IR) if every type obtains at least zero surplus by reporting truthfully, that is,

$$v(t, x(t)) - p(t) \geq 0$$

for all types t. Any mechanism we will refer to will be IC-IR unless otherwise stated. Every mechanism can be represented by a menu of random alternative and price pairs such that each type chooses a pair that maximizes his surplus. If a type is indifferent between two or more pairs, he chooses the one with a higher price and chooses any one of the pairs if the prices are the same.

A *market* $f \in \Delta(T)$ is a distribution over types, where $f(t)$ is the fraction of consumers of type t. An IC-IR mechanism is *optimal* for a market f if it maximizes the seller's expected revenue among all IC-IR mechanisms. A market may have multiple optimal mechanisms, and a type's surplus may vary across these mechanisms. Thus, to compare consumer surplus across different markets, we fix a selection rule that specifies an optimal mechanism for each market. The selection rule should satisfy a mild consistency requirement but is otherwise arbitrary. The requirement is that if two markets have the same set of optimal mechanisms, then the same mechanism is selected for both markets. We henceforth fix such a selection rule and refer to the selected optimal mechanism for a market as the optimal mechanism for that market. Type t's surplus $\mathrm{CS}(t, f)$ in market f is the type's surplus from the optimal mechanism. A market is efficient if the allocation in the optimal mechanism of every type in the market is efficient. In this case, we say that the optimal mechanism is efficient. Otherwise, the optimal mechanism and the market are inefficient.

A *segmentation* of market f is a distribution $\mu \in \Delta(\Delta(T))$ over a finite set of markets that averages to f, that is, $E_{f' \sim \mu}[f'] = f$.[7] We refer to a market in the support of a segmentation as a (market) segment. A segmentation is nontrivial if not all segments are identical to the original market. Given a segmentation, the seller offers in each market segment the optimal mechanism for that segment.

A. Pareto Improvements

Our goal is to understand, for each environment, which markets can be segmented in a way that benefits all consumers and the seller. To formalize this, we say that market f' *weakly Pareto dominates* market f if the set of types in market f' is a subset of the set of types in market f and every type t in market f' weakly prefers the optimal mechanism for market f' to the one for market f, that is, $\mathrm{CS}(t, f') \geq \mathrm{CS}(t, f)$. If, in addition, the preference

[7] The restriction to a finite set of markets is without loss of generality for all our results and examples because the number of types is finite.

is strict for some type in market f', then f' *Pareto dominates* f. A segmentation μ of market f is *Pareto improving* if every segment weakly Pareto dominates f, some segment Pareto dominates f, and the segmentation strictly increases the seller's revenue.[8] A market is *Pareto improvable* if it has a Pareto-improving segmentation. Such a segmentation increases the value of any monotone function of all consumers' surplus. Because the seller's revenue strictly increases, she has a strict incentive to carry out the segmentation. The rest of the paper focuses on identifying, for every environment, which markets are Pareto improvable and constructing a Pareto-improving segmentation for these markets.

We begin by observing that if a market is efficient, then it is not Pareto improvable. This is because a Pareto-improving segmentation strictly increases total surplus, and the total surplus that any segmentation generates is at most the surplus generated by an efficient mechanism.

Observation 1. Any Pareto improvable market is inefficient.

Which inefficient markets are Pareto improvable? The following proposition provides the answer for environments with a single alternative.

Proposition 1. In any environment with a single alternative, all inefficient markets are Pareto improvable.

Proposition 1 follows from the proof of theorem 1 in Bergemann, Brooks, and Morris (2015). Their result implies that any inefficient market with a single alternative can be segmented in a way that increases both the seller's revenue and average consumer surplus. But their proof in fact shows that Pareto-improving segmentations exist.[9] However, that proof relies heavily on there being a single alternative and does not generalize to multiple alternatives. The following example illustrates which markets are Pareto improvable in a particular environment with two types and two alternatives.

B. *Pareto Improvements in a Two-Type, Two-Alternative Example*

Consider an environment with a low type t_L, a high type t_H, and two alternatives a_L and a_H (in addition to the outside option). Both types prefer a_H to a_L. The low type's valuations for the two alternatives are 0.75 and 1, and the high type's valuations for the two alternatives are 1 and 2. Denote a market by the fraction q of high-type consumers. We will show that all

[8] Any segmentation weakly increases the seller's revenue since the seller can use the optimal mechanism for the original market in all segments. The increase is strict if the optimal mechanism for the original market is not optimal for some segment.

[9] A technical point is that their proof and our proposition 1 require selecting the efficient mechanism if it is optimal for a market, whereas our definition of the optimal mechanism allows for any selection rule when there are multiple optimal mechanisms. Proposition 1 does not hold for any selection rule, but our results in the rest of the paper do, and they apply in particular to markets with a single alternative.

the inefficient markets except for market $q = 0.75$ are Pareto improvable by a two-market segmentation.

The set of markets can be divided into three intervals: an efficient low interval $[0, 0.25]$ in which the optimal mechanism assigns alternative a_H to both types at price 1, an inefficient intermediate interval $(0.25, 0.75]$ in which the optimal mechanism assigns alternative a_L to the low type at price 0.75 and alternative a_H to the high type at price 1.75, and a high interval $(0.75, 1]$ in which the optimal mechanism assigns the outside option to the low type at price 0 and alternative a_H to the high type at price 2. The markets in the high interval are inefficient except for market $q = 1$, which includes only high-type consumers.

The surplus of the low type is zero in any market. The surplus of the high type across markets is depicted in figure 1. This surplus is constant on each interval because the surplus depends only on the optimal mechanism. The surplus is lower on higher intervals.

We now determine which inefficient markets are Pareto improvable. Consider first any market q in the interior of the intermediate interval, $(0.25, 0.75)$. This market is inefficient and can be segmented into two market segments, q' and q'', such that $q' > 0$ is in the low interval and $q'' > q$ is in the intermediate interval. The surplus of the high type in market q' increases relative to his surplus in q, and the surplus of the high type

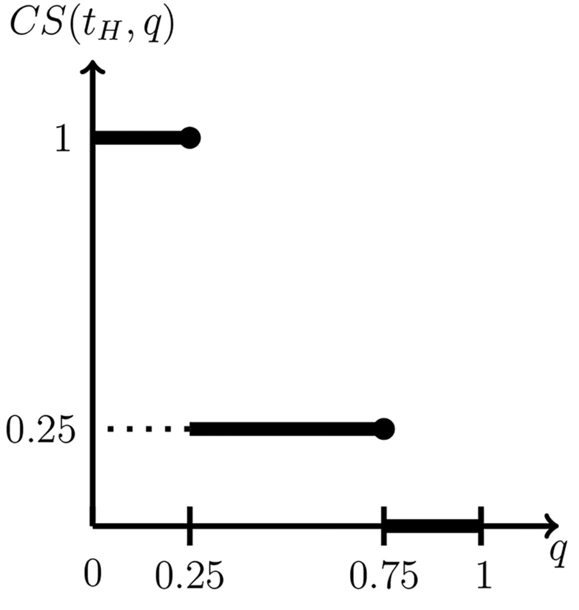

FIG. 1.—Surplus $CS(t_H, q)$ of high type in any market q.

in market q'' is the same as his surplus in q. The surplus of the seller also increases because the optimal mechanism for market q is not optimal for market q'. This shows that the segmentation is Pareto improving.[10] A similar argument shows that any market in the interior of the high interval, (0.75, 1), is Pareto improvable by a two-market segmentation. Thus, it remains to determine only whether market 0.75 is Pareto improvable.

Market 0.75 is inefficient but not Pareto improvable. This is because any nontrivial segmentation of this market contains some segment strictly larger than 0.75, and the surplus of the high type in this segment is zero, which is lower than his surplus of 0.25 in market 0.75.[11]

IV. The Main Result

Our main result is that, generically, inefficient markets are Pareto improvable by a two-market segmentation. We formalize the result after defining our notion of genericity.

Definition 1. A set F of markets is nongeneric in a set G of markets if, for some $l > 0$, $F \cap G$ is contained in a finite union of hyperplanes of dimension $l - 1$ and G contains a ball of dimension l.[12] A set of markets F is generic in G if G is empty or the complement of F is nongeneric in G.

Theorem 1. For any environment, the set of markets that are Pareto improvable by a two-market segmentation is generic in the set of inefficient markets.

Theorem 1 shows that if the set of inefficient markets is not empty, then the subset of inefficient markets that are not Pareto improvable lies in a lower dimensional space (comprised of a finite union of lower-dimensional hyperplanes).

To illustrate definition 1 and theorem 1, consider the two-type example from section III.B. The set of inefficient markets is (0.25, 1), which contains a ball of dimension $l = 1$ (an interval), and the only inefficient market that is not Pareto improvable is 0.75, which is contained in a hyperplane of dimension $l - 1 = 0$ (a point). Thus, the set of markets that are not Pareto improvable is nongeneric in the set of inefficient markets, so the set of markets that are Pareto improvable is generic in the set of inefficient markets.

[10] Recall that the surplus of the low type is zero in all markets.

[11] In fact, every nontrivial segmentation of market 0.75 also lowers the average consumer surplus. This can be seen by concavifying the function that maps any market to the average consumer surplus in the optimal mechanism for that market. Because in any market q the surplus of the low type is zero, the average consumer surplus is $q \cdot \mathrm{CS}(t_H, q)$, where $\mathrm{CS}(t_H, q)$ is the surplus of the high type in market q, depicted in fig. 1.

[12] This implies measure theoretic and topological notions of nongenericity. Indeed, if $l > 0$ is the largest integer such that G contains a ball of dimension l, then $F \cap G$, for a nongeneric F, has Lebesgue measure 0 and is nowhere dense (in $\mathbb{R}^l$).

The proof of theorem 1 shows that inefficient markets with a unique optimal payment rule are Pareto improvable by a two-market segmentation, and the set of markets with a unique optimal payment rule is generic in the set of inefficient markets.[13] The idea behind the genericity is that the set of payment rules is a convex polytope, and the optimality of multiple payment rules translates into one or more linear equalities the market must satisfy. Since the polytope has a finite number of vertices, the set of markets for which multiple payment rules are optimal lies in a finite union of hyperplanes of lower dimension.

Showing that inefficient markets with a unique optimal payment rule are Pareto improvable by a two-market segmentation relies on a new two-step approach. The first step is to construct, for any inefficient market f, an efficient Pareto-dominating market f'. This is achieved by understanding what makes inefficient mechanisms optimal and is the key to theorem 1. The second step shows that slightly perturbing a market with a unique optimal payment rule does not change the set of optimal mechanisms. Combining the two steps leads to theorem 1: segment market f by assigning probability ε to the Pareto-dominating market f' and probability $1 - \varepsilon$ to the remaining market f'' so that $f = \varepsilon f' + (1 - \varepsilon)f''$. If ε is small, then f'' is a small perturbation of f; thus, as long as market f belongs to the generic set of markets with a unique optimal payment rule, market f'' has the same optimal mechanism as f and hence weakly Pareto dominates market f.[14] The segmentation also strictly increases the seller's revenue because the optimal mechanism for f' is different from the optimal mechanism for f. Therefore, the segmentation of market f into f' and f'' is Pareto improving. We now describe the approach in greater detail.

A. *Step 1: Constructing a Pareto-Dominating Market*

The first step is formalized as follows.

Proposition 2. For any market f, there exists an efficient Pareto-dominating market with a unique optimal mechanism if and only if f is inefficient.

In the two-type example from section III.B, the set of inefficient markets is (0.25, 1), and every inefficient market is Pareto dominated by all the markets in (0, 0.25], which are efficient. The challenge in proving proposition 2 is that in some environments with more than two types, a Pareto-dominating market may necessarily include more than two types.

[13] A market has a unique optimal payment rule if $p(t) = p'(t)$ for any type t in the market and any two optimal mechanisms (x, p) and (x', p') for the market.

[14] Recall that if two markets have the same set of optimal mechanisms, the selection rule selects an arbitrary but identical optimal mechanism for both.

We first provide an example that illustrates this and then prove proposition 2.

1. Pareto-Dominating Market with Necessarily More than Two Types

Consider the environment with three types (t_1, t_2, and t_3) and two alternatives (a_1 and a_2), illustrated in figure 2*A*. Each type is described by the circle with the type's label to its left. The horizontal axis shows the valuation for alternative a_1, and the vertical axis shows the valuation for alternative a_2.

Figure 2*B* depicts a mechanism in which the two alternatives are offered at prices $p(a_1) = v(t_1, a_1)$ and $p(a_2) = v(t_3, a_2) - v(t_3, a_1) + v(t_1, a_1)$. At these prices, the light gray region contains the set of types that prefer alternative a_1, the dark gray region contains the set of types that prefer alternative a_2, and the unshaded region contains the set of types that prefer alternative 0 (the outside option). In particular, type t_1 is indifferent between alternative a_1 and the outside option (and chooses a_1), and type t_3 is indifferent between a_1 and a_2 (and chooses a_2). Type t_2 strictly prefers (and chooses) alternative a_1. This mechanism gives zero surplus to type t_1 and strictly positive surplus to the other types.

There exists a market f with full support for which this mechanism is optimal. In this market, the fraction of type t_1 is large enough that it is optimal to assign him his efficient alternative a_1 for a price that is equal to his valuation; among the remaining consumers, the fraction of type t_3 is large enough that is it optimal to assign him his efficient alternative a_2 for the maximal price that maintains IC. This market is inefficient because type t_2 is assigned alternative a_1.

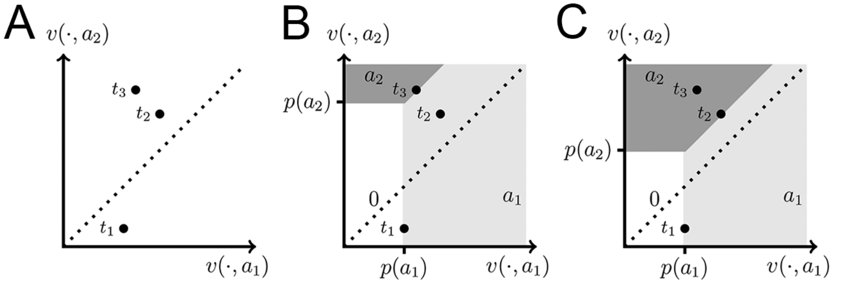

FIG. 2.—*A*, Environment with three types and two alternatives. *B*, Optimal prices $p(a_1)$ and $p(a_2)$ for market f. *C*, Optimal prices for Pareto-dominating market f'. In each panel, types in the light gray region prefer alternative a_1, types in the dark gray region prefer alternative a_2, and types in the unshaded region prefer alternative 0.

There exists an efficient market f' with full support that Pareto dominates f. In this market, the fraction of type t_1 is large enough that it is optimal to assign him his efficient alternative a_1 for a price that is equal to his valuation; among the remaining consumers, the fraction of type t_2 is large enough that is it optimal to assign him his efficient alternative a_2 for the maximal price that maintains IC. Type t_3 is also assigned alternative a_2 for this price. This mechanism is illustrated in figure 2*C*. Since the price of alternative a_2 is lower than in the optimal mechanism for market f, market f' Pareto dominates market f.

There is, however, no two-type market that Pareto dominates f: in any market without type t_1, the surplus of one of the other types is zero (any optimal mechanism gives surplus zero to some type); in any market without type t_2, either the allocation and surpluses of the other types is unchanged or the surplus of type t_3 is strictly lower than in market f.[15] Similarly, in any market without type t_3, either the allocation and surpluses of the other types is unchanged or the surplus of type t_2 is strictly lower than in market f.

The reason that all three types are needed to form a Pareto-dominating market is that in order to increase the surplus of type t_3 (who is already assigned his efficient alternative a_2 in market f), type t_2 must be present in sufficient proportion to make it optimal for the seller to lower the price of alternative a_2 in order to extract more surplus from type t_2. But type t_2's surplus in market f is positive; in order to maintain this surplus in the Pareto-dominating market, type t_1 must be present in sufficient proportion to make it optimal for the seller to assign alternative a_1 to type t_1, thereby providing information rents to type t_2.

2. Proof of Proposition 2

In an inefficient market f, some type t is assigned an inefficient alternative. This inefficiency allows the seller to lower the surplus (information rents) of some other type t'. (In the example from sec. IV.A.1, $t = t_2$ and $t' = t_3$.) In a new market that includes only type t and t' and in which the proportion of type t is sufficiently high, it is optimal to assign type t his efficient alternative; this increases the surplus that type t' obtains from being able to mimic type t. But the surplus of type t may decrease; to prevent this, we identify an information rents path in market f that begins

[15] If type t_1 is assigned the outside option, then the surplus of type t_3 is zero. Otherwise, type t_1 must be assigned alternative a_2 with probability zero, since replacing alternative a_2 with alternative a_1 in his allocation allows the seller to charge type t_1 and type t_3 higher prices. And among the allocations that assign alternative a_2 with probability zero to type t_1, the one that assigns type t_1 alternative a_1 with certainty gives type t_3 the highest surplus, which is equal to his surplus in f. This is achieved by assigning type t_3 alternative a_2.

with type t and ends with some type t'' that has surplus zero, and we add to the new market all the types in the path and type t' in the appropriate proportions. (In the example, $t'' = t_1$.) This generates a market that Pareto dominates market f. We now describe this procedure in more detail.

Take an inefficient market f that (without loss of generality) has full support, and let t be some type that is assigned an inefficient alternative in the optimal mechanism M for market f. We inductively construct a set of types S that contains t such that for every type t' in S, there is a directed path of types in S from type t to type t' such that the IC constraint, given mechanism M, from each type t_j to the next type t_{j+1} in the path binds (i.e., type t_j is indifferent between reporting truthfully and misreporting that he is type t_{j+1}). The construction of S stops when a type that has zero surplus is added to S. If type t has zero surplus, we are done. Otherwise, given the set S so far constructed, there is a type t' not in S such that the IC constraint from some type in S to type t' binds. Otherwise, the revenue in market f can be increased by increasing the payments of all types in S by the same small amount. This concludes the construction of S.

Consider the set of types $\bar{S} \subset S$ in the binding IC path that begins with type t and ends with the type that has zero surplus. Without loss of generality, type t is the only type in $\bar{S}$ that is assigned an inefficient alternative (otherwise, denote by t the last type in the path that is assigned an inefficient alternative and remove from $\bar{S}$ all types that preceded t in the path). Notice that the payments of the types weakly decrease along the path (otherwise, the revenue in market f can be increased by replacing some type's assigned alternative and payment with those of the next type in the path without violating IC and IR).

Now, modify the optimal mechanism M for market f by assigning type t his efficient alternative and increasing his payment to leave his surplus unchanged. The modified mechanism M^1 violates IC; otherwise, mechanism M^1 would generate more revenue than mechanism M in market f. Therefore, when faced with mechanism M^1, some type $t' \neq t$ strictly prefers to misreport that he is type t. This type t' is not in $\bar{S}$, since in M (and therefore in M^1), every type in $\bar{S}$ other than type t is assigned his efficient alternative and pays less than type t does in M (since payments weakly decrease along the path). Modify mechanism M^1 by replacing the assigned alternative and payment of type t' with those of t. This modified mechanism M^2 satisfies IC and IR for the set of types $\bar{S} \cup \{t'\}$, and type t' has strictly higher surplus than in mechanism M. Finally, if t' is not assigned his efficient alternative in mechanism M^2, modify M^2 by assigning type t' his efficient alternative and increasing his payment to leave his surplus unchanged. Denote the resulting mechanism by M^*. Notice that in mechanism M^*, every type in $\bar{S}$ is assigned his efficient alternative and pays at most what t' does, so no type different from t' benefits from misreporting that he is type t'. Consider the restricted environment with types $\bar{S} \cup \{t'\}$.

Mechanism M^* is efficient and IC-IR in this environment. Moreover, the surplus of every type in $\bar{S} \cup \{t'\}$ is weakly higher than in mechanism M, and the surplus of type t' is strictly higher.

It remains to show that M^* is the unique optimal mechanism for some full-support market in the restricted environment. This can be done because mechanism M^* is efficient.[16] We provide the intuition here and defer the formal proof to the appendix. Such a market can be constructed iteratively. Take the path that defined $\bar{S}$ and add type t' to its beginning (so type t follows type t'). Begin with a large enough fraction, smaller than 1, of the last type in the path so that it is strictly optimal for the seller to assign this type his efficient alternative for a price that is equal to his valuation. Add a large enough fraction of the second to last type in the path so that it is strictly optimal for the seller to assign this type his efficient alternative for the maximal price that maintains IC, and so on. The resulting mechanism is M^*, which is the unique optimal mechanism for the resulting market.[17] This completes the proof of the "if" direction in proposition 2. The proof of the "only if" direction is in the appendix.[18]

3. Revisiting the Example from Section IV.A.1

Consider the optimal mechanism for market f shown in figure 2B. Figure 3A illustrates the binding IC and IR constraints and the allocation of each type in the optimal mechanism. Types t_1 and t_2 are assigned alternative a_1, and type t_3 is assigned alternative a_2. The arrows indicate the binding constraints: type t_1's IR constraint binds, types t_1 and t_2 are indifferent between reporting truthfully and mimicking each other (since they both obtain alternative a_1), and type t_3 is indifferent between reporting truthfully and mimicking types t_1 and t_2.

To construct the set S, we begin with type $t = t_2$, whose allocation is inefficient. Figure 3A shows that this type's IR constraint does not bind, so his surplus is positive. We therefore add to $S = \{t_2\}$ a type t' not in S such that the IC constraint from type t_2 to type t' binds. This must be type $t' = t_1$, as the binding IC constraints in figure 3A illustrate. This concludes the construction of $S = \{t_1, t_2\}$ since the IR constraint of type t_1 binds. The

[16] This is why we modified mechanism M^2 to obtain mechanism M^*.

[17] This construction is where we use the assumption that each type has a unique efficient alternative. If types have multiple efficient alternatives, it is still true that there exists a market for which it is optimal to assign each type an efficient alternative. However, that alternative may be different from the efficient alternative prescribed by mechanism M^*. If there are multiple efficient alternatives, a more complicated construction may be required to address the issue of which efficient alternative is selected.

[18] That proof shows the stronger result that for an efficient market, there is no Pareto-dominating market, and not just no efficient Pareto-dominating market with a unique optimal mechanism.

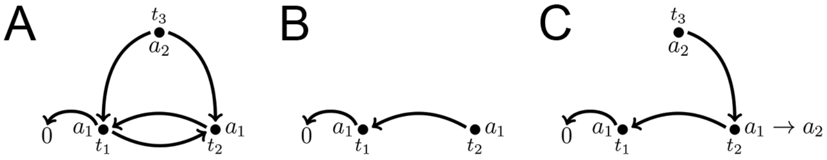

FIG. 3.—Execution of procedure that constructs a Pareto-dominating market. *A*, Binding IC and IR constraints for market *f*. *B*, Path of binding IC constraints that starts with type t_2, whose allocation is inefficient, and ends with type t_1, whose IR constraint binds. *C*, Appended path with type t_3, which strictly benefits in Pareto-dominating market.

binding IC path in S that begins with type t_2 and ends with type t_1, whose surplus is zero, is depicted in figure 3*B*. Assigning type $t = t_2$ his efficient alternative and increasing his payment to leave his surplus unchanged violates the IC constraint from type t_3 to type t_2, which binds in figure 3*A*. The appended path with type $t' = t_3$ at its beginning and the modifications to the optimal mechanism for the original market are illustrated in figure 3*C*. The allocation in the resulting mechanism is the one in figure 2*C*. The payments in the resulting mechanism can be obtained from the allocation and the binding IR and IC constraints.

B. *Step 2: Perturbing the Market*

We now show that for a market with a unique optimal payment rule, perturbing the market leaves the set of optimal mechanisms unchanged. We then show that the set of markets with a unique optimal payment rule is generic in the set of inefficient markets.

We first describe these results informally and illustrate them in the context of the example from section III.B.[19] Consider the set P of all payment rules $p : T \to \mathbb{R}_+$ that are part of IC-IR mechanisms. We show that P is a polytope. The set P in the example from section III.B is the polytope depicted in figure 4*A*, which has three vertices other than the origin. One vertex is payment rule $p^1 = (p^1(t_L), p^1(t_H)) = (1, 1)$, which is part of a mechanism that assigns alternative a_H to both types at price 1. Another vertex is payment rule $p^2 = (p^2(t_L), p^2(t_H)) = (0.75, 1.75)$, which is part of a mechanism that assigns alternative a_L to type t_L at price 0.75 and alternative a_H to type t_H at price 1.75. The third vertex is payment rule $p^3 = (p^2(t_L), p^2(t_H)) = (0, 2)$, which is part of a mechanism that assigns the outside option to type t_L at price 0 and alternative a_H to type t_H at price 2.

A mechanism is optimal for a market if and only if its payment rule is maximal in P in the direction specified by the market. If the market is

[19] Recall that in that example, there are two types, t_L and t_H, and two alternatives, a_L and a_H, with $v(t_L, a_L) = 0.75$, $v(t_L, a_H) = v(t_H, a_L) = 1$, and $v(t_H, a_H) = 2$.

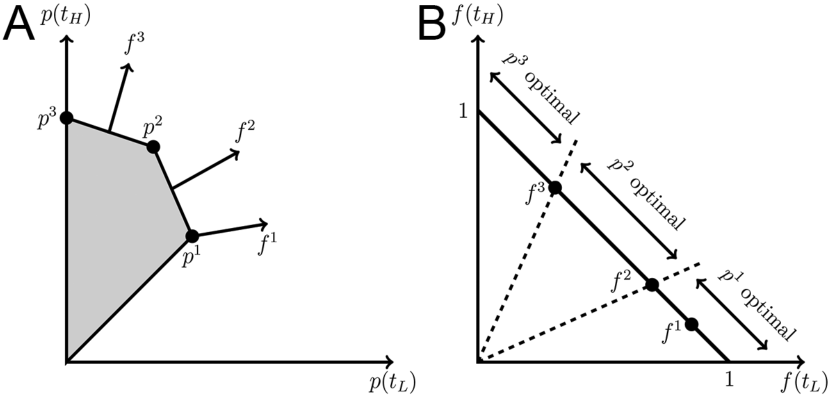

FIG. 4.—*A*, Shaded polytope is set P of all payment rules in example from section III.B. Market f^1 is not orthogonal to a face of P. Markets f^2 and f^3 are orthogonal to a face of P. *B*, Solid line is set of all markets in this environment. A perturbation of f^1 does not change the optimal mechanism, but a perturbation of f^2 and f^3 may do so.

not orthogonal to a face of P, then the market has a unique optimal payment rule, which is a vertex of P. This is the case for market $f^1 = (f^1(t_L), f^1(t_H)) = (0.9, 0.1)$ and payment rule p^1 in figure 4*A*. The same payment rule remains uniquely optimal for small enough perturbations of such a market. In contrast, if the market is orthogonal to a face of P, then the market has multiple optimal payment rules and a small perturbation of the market may change the optimal payment rule and therefore the optimal mechanism. This is the case for markets $f^2 = (f^2(t_L), f^2(t_H)) = (0.75, 0.25)$ and $f^3 = (f^3(t_L), f^3(t_H)) = (0.25, 0.75)$ in figure 4*A*. For market f^2, the optimal payment rules are p^1, p^2, and their convex combinations (corresponding to the face of P that connects p^1 and p^2). For market f^3, the optimal payment rules are p^2, p^3, and their convex combinations.

To see why the effects of small perturbations depend on whether the market is orthogonal to a face of P and why such markets are nongeneric, consider the set of all markets and their optimal payment rules among the vertices of P. Figure 4*B* depicts this set and the optimal payment rules for the example. Any market lies on the hyperplane $f(t_L) + f(t_H) = 1$, which is the solid line in figure 4*B*. A market with multiple optimal payment rules additionally lies on a hyperplane of vectors that have the same inner product with all these payment rules.

Market f^1 has a unique optimal payment rule, p^1, which is also uniquely optimal for nearby markets. Market f^2 has two optimal payment rules that are vertices of P, p^1 and p^2. This market lies on the hyperplane $f(t_L) - 3f(t_H) = 0$, which is the dashed line that goes through f^2 in figure 4*B*.[20] For markets on one side of this hyperplane, payment rule p^1 has a higher revenue than payment rule p^2, so a small perturbation of f^2 in that direction makes p^1 uniquely optimal. A small perturbation of f^2 in the other direction makes p^2 uniquely optimal. Market f^3 has two optimal payment rules that are vertices of P, p^2 and p^3. This market lies on the hyperplane $3f(t_L) - f(t_H) = 0$, which is the dashed line that goes through f^3 in figure 4*B*. Both markets f^2 and f^3 are not Pareto improvable, but f^2 is the only inefficient market that is not Pareto improvable.

Before presenting the formal perturbation and genericity arguments, we point out that the argument is not completely self-evident; we allow for random alternatives, of which there is a continuum, and appendix A4 shows that in some environments with a continuum of alternatives, the perturbation argument fails. The finite number of types and (nonrandom) alternatives leads to the set of payment rules being a polytope, which facilitates the perturbation argument.

C. *Formalizing the Perturbation Argument*

To formalize the argument, we observe that the set of IC-IR mechanisms is a polytope in $\mathbb{R}_+^{(k+2)n}$, where k is the number of alternatives and n is the number of types. Indeed, a mechanism is a point in $\mathbb{R}_+^{(k+2)n}$ (for each of the n types, it specifies the nonnegative payment and the probability of being assigned each one of the k alternatives and the outside option), each of the finite number of IC and IR constraints corresponds to a half space, and the (linear) probability constraints together with the IR constraints guarantee that the set is bounded. The set P of payment rules that are part of IC-IR mechanisms is a projection of the set of IC-IR mechanisms and is therefore a polytope in $\mathbb{R}_+^n$. Consequently, set P has a finite set of vertices.

LEMMA 1. There exists a finite set $P_V \subseteq \mathbb{R}_+^n$ such that P is the convex hull of P_V.

We say that markets f and f' are ε-close if $|f(t) - f'(t)| \leq \varepsilon$ for all types t. We say that perturbing market f leaves the set of optimal mechanisms unchanged if for some $\varepsilon > 0$, the set of optimal mechanisms for f is equal to the set of optimal mechanisms for any market f' that is ε-close to f. The consistency requirement from the selection of optimal mechanisms guarantees that the optimal mechanism selected for f is the same as the one selected for f'.

LEMMA 2. If a market has a unique optimal payment rule, then perturbing the market leaves the set of optimal mechanisms unchanged.

[20] For a market $(f(t_L), f(t_H))$, the revenues from payment rules $p^1 = (1, 1)$ and $p^2 = (0.75, 1.75)$ are $f(t_L) + f(t_H)$ and $0.75 \cdot f(t_L) + 1.75 \cdot f(t_H)$, which are equal if and only if $f(t_L) - 3f(t_H) = 0$.

Proof. Consider a market f with a unique optimal payment rule p. Since the set P_V is finite (lemma 1) and p is the unique optimal payment rule, there exists $\delta > 0$ such that $E_{t\sim f}[p(t)] > E_{t\sim f}[p'(t)] + \delta$ for all $p' \in P_V\backslash\{p\}$. By continuity of the expected revenue in f, there exists $\varepsilon > 0$ such that $E_{t\sim f'}[p(t)] > E_{t\sim f'}[p'(t)]$ for all $p' \in P_V\backslash\{p\}$ and all f' that are ε-close to f. Since all payment rules are convex combinations of the payment rules in P_V, we have $E_{t\sim f'}[p(t)] > E_{t\sim f'}[p'(t)]$ for all payment rules $p' \in P\backslash\{p\}$. That is, the payment rule p is also the unique optimal payment rule for all f' that are ε-close to f. Therefore f and any such f' have the same set of optimal mechanisms, since any two mechanisms with the same payment rule generate the same revenue. QED

We complete step 2 by showing that the set of markets that have a unique optimal payment rule—and can therefore be perturbed without changing the set of optimal mechanisms—is generic in the set of inefficient markets.

LEMMA 3. The set of markets with a unique optimal payment rule is generic in the set of inefficient markets.

Proof. By definition 1, if all markets are efficient, we are done. Suppose that not all markets are efficient.[21] We will show that the set of inefficient markets contains a ball of dimension $n - 1$ and the set of markets with multiple optimal payment rules is contained in a finite union of hyperplanes of dimension $n - 2$ (so the same is true for the set of inefficient markets with multiple optimal payment rules).

Since not all markets are efficient, the first-best mechanism—which assigns to each type his efficient alternative at a price equal to his valuation—is not IC. Denote by t and t' two types such that in the first-best mechanism, the IC constraint from t' to t is violated (so type t' prefers to misreport that he is type t). We will construct a market with full support that has a unique optimal payment rule p and for which any optimal mechanism is inefficient because the allocation of type t is inefficient. We construct the payment rule p and the market together inductively as follows. First, set the payment $p(t')$ of type t' equal to his valuation for his efficient alternative and put a large enough fraction of type t' in the market that any optimal payment rule in P specifies for type t' payment $p(t')$.[22] Now take a type t'' that has not yet been added to the market and consider the maximal payment of type t'' across all payment rules in P that coincide with the payment rule p so far constructed. Set $p(t'')$ equal to this maximal payment and add to the market a large enough fraction of type t'' that any optimal payment

[21] In particular, $n > 1$.

[22] Recall that the polytope P is the set of all payment rules that are part of some IC-IR mechanism. Since P has a finite number of vertices, there is a finite number of vertices in which the payment of type t' is less than $p(t')$. Thus, in all those vertices, the payment of type t' is bounded away from $p(t')$, so none of these vertices can be an optimal payment rule for any market with a large enough fraction of type t'.

rule in P specifies for type t'' payment $p(t'')$. Repeat this process until the market includes all n types.

By construction, p is the unique optimal payment rule for the resulting market. And any optimal mechanism with this payment rule is inefficient because the payment of type t' is equal to his valuation for his efficient alternative, so the allocation of type t must be inefficient (because in the first-best mechanism, the IC constraint from t' to t is violated). Therefore, by lemma 2, perturbing the market leaves the set of optimal mechanisms unchanged, so these optimal mechanisms are all inefficient. This shows that the set of inefficient markets contains a ball of dimension $n - 1$.

We now turn to the set of markets with multiple optimal payment rules. Consider a market f for which more than one payment rule maximizes revenue. Since P is the convex hull of P_V, there exist two payment rules $p \neq p'$ in P_V that are both optimal for f. Thus, f is contained in the hyperplane of dimension $n - 2$ defined by the equations $\Sigma_t f(t)(p(t) - p'(t)) = 0$ and $\Sigma_t f(t) = 1$. Since P_V is finite, the set of markets with multiple optimal payment rules is contained in a finite union of such hyperplanes, one for each pair of payment rules in P_V. QED

V. Special Cases and Applications

The key to our construction of Pareto-improving segmentations is identifying Pareto-dominating markets. These Pareto-dominating markets can vary widely across environments and markets: they may necessarily include a large number of types or they may include only two types, and these types can vary across dominated markets with the same set of types. Many settings, however, satisfy properties that may allow us to say more about the Pareto-dominating markets. We briefly describe two examples of such properties and their implications and relegate the details to appendix A3.

The first property of a market is the presence of a lowest type, whose valuation for any alternative (other than the outside option) is strictly lower than the valuation of any other type in the market for that alternative. We show that any inefficient market with a lowest type has a Pareto-dominating market that includes the lowest type. In the example from section IV.A.1, type t_1 is the lowest type. The second property of a market is that a best alternative exists, which is the efficient alternative for all the types in the market. We show that any inefficient market with a best alternative has a Pareto-dominating market that includes only two types. In the example from section IV.A.1, not all types have the same efficient alternative, and there is no Pareto-dominating market that includes only two types.

In specific environments, even more can be said about the structure of Pareto-dominating markets. We illustrate this in the context of two applications. The first application is a version of Mussa and Rosen (1978) with linear valuations and ranked types. We show that in this setting, for every inefficient market there exists a Pareto-dominating market that consists of a type and all lower types in the market. In addition, Pareto-dominating markets may necessarily include more than two types. The second application, whose details are in appendix A3, is a bundling setting with multiple products, additive valuations, and zero production costs. The grand bundle of all products is the best alternative in every market, so for every inefficient market there exist Pareto-dominating segments with only two types. But we show that these types may differ across inefficient markets with the same set of types.

We now describe the application with linear valuations in greater detail. There are k alternatives $0 < a_1 \cdots < a_k$ (in addition to the outside option) and n types $0 < t_1 \cdots < t_n$. The valuation of type t for alternative a is $v(t, a) = ta$. The cost of producing an alternative a is $c(a) \geq 0$, where $c(0) = 0$ and c is increasing and convex.[23] Recall our assumption that each type t has a unique efficient alternative $\bar{a}(t)$; that is, $\bar{a}(t)$ is the unique alternative that maximizes $ta - c(a)$.

We apply the procedure for constructing a Pareto-dominating segment as described in section IV.A. Given an inefficient market f with full support (for notational simplicity), we start by choosing some type t_j that is assigned an inefficient alternative in the optimal mechanism for f. We then identify a binding IC path of types in which each type is indifferent between his own allocation and payment and those of the next type in the path, and the last type in the path has zero surplus. To identify such a path, we observe, by standard analysis of optimal mechanisms, that in the optimal mechanism for f, the IC constraint from type t_i to the next lowest type t_{i-1} binds for every i. In addition, the IR constraint of at least one type binds. Thus, there is a binding IC path $(t_j, t_{j-1}, \ldots, t_{j'})$ that begins with type t_j and ends with type $t_{j'}$ whose IR constraint binds.

As in section IV.A, we assume without loss of generality that t_j is the only type in this path that is assigned an inefficient alternative in the optimal mechanism for f. The path contains only type t_j if his IR constraint binds in the optimal mechanism for f and otherwise contains additional types.[24]

Now consider the restricted environment that contains type t_{j+1} and the types in the binding IC path.[25] The procedure constructs an efficient

[23] Even though costs can be normalized to zero, as discussed in sec. III, it is convenient in this application to write them explicitly.

[24] For an example of a path with necessarily more than one type, consider an environment with three types, $t_1 = 1$, $t_2 = 2$, and $t_3 = 3$; two alternatives, $a_1 = 1$ and $a_2 = 2$; and costs $c(a_1) = 0.25$ and $c(a_2) = 1.75$. The efficient alternatives are $(\bar{a}(t_1), \bar{a}(t_2), \bar{a}(t_3)) = (a_1, a_2, a_2)$. For market $f = (f(t_1), f(t_2), f(t_3)) = (2/3, 1/6, 1/6)$, the allocation in the optimal mechanism is $a = (a(1), a(2), a(3)) = (a_1, a_1, a_2)$. The only type assigned an inefficient alternative is type t_2, but the only type whose IR binds is type t_1, so the binding IC path consists of type t_2 followed by type t_1.

[25] Type t_{j+1} exists, i.e., $j < n$, because in any market in a linear environment, the optimal allocation of the highest type is efficient (no distortion at the top).

market that includes all the types in the restricted environment. For any market f' in this environment, consider the virtual value (Myerson 1981; see Vohra 2011 for the formulation used here with a finite number of types) of each type t_i: $\phi_i = t_i - (t_{i+1} - t_i)(\Sigma_{i'>i} f'(t_{i'})/f'(t_i))$, where the summation is over types in the restricted environment. If $\Sigma_{i'>i} f'(t_{i'})$ is small enough relative to $f'(t_i)$ for each i, then the virtual value of each type t_i approaches t_i.[26] Consequently, for such a market f', any optimal mechanism assigns to each type his efficient alternative. By standard arguments, the surplus of each type is pinned down by the allocation of all lower types (and the surplus of the lowest type) and strictly increases in the allocation of each of those lower types. Thus, all types in the restricted environment other than type t_{j+1} obtain the same surplus in f and in f', and the surplus of type t_{j+1} is strictly higher in f'. This shows that market f' Pareto dominates market f.

We now discuss the second step in the proof of theorem 1. Because market f' Pareto dominates market f, market f is Pareto improvable if perturbing it does not change the optimal mechanism. This is the case if for each type t_i, there is a unique alternative that maximizes the virtual surplus $\phi_i a - c(a)$ over all a, because then that same alternative continues to be optimal if the virtual value ϕ_i is perturbed slightly. We now show that uniqueness of the virtual surplus maximizers is a generic property. Fix a type t_i. As we vary the market, the virtual value ϕ_i can take any value not greater than t_i. But there is only a finite number of values for which multiple alternatives maximize the virtual surplus: by convexity of the cost function, this happens if and only if $\phi_i a_j - c(a_j) = \phi_i a_{j+1} - c(a_{j+1})$ for some j.

In this application with linear valuations, the procedure results in a Pareto-dominating market f' that has a special structure: it consists of some interval of types from t'_j to t_{j+1}. If the allocation of multiple types in the optimal mechanism for market f is inefficient, then the Pareto-dominating market we obtain may depend on which of those types we start with.[27] Despite this multiplicity, for any inefficient market there exists a binding IC path that ends in t_1, so the corresponding Pareto-dominating market contains a prefix of all types $t_1, \ldots, t_j, t_{j+1}$. To see this, start the process from type t_j who is the lowest type whose allocation in the optimal mechanism for the market is inefficient. Then the only type

[26] In particular, virtual values strictly increase in type, so no ironing is needed.

[27] For example, suppose that there are three types, t_1, t_2, and t_3, and consider a market such that in the optimal mechanism, type t_1 is assigned the outside option, type t_2 is assigned an inefficient alternative different from the outside option, and type t_3 is assigned his efficient alternative. Then the IR constraints of types t_1 and t_2 bind, and the same two types have inefficient allocations. If the procedure starts by choosing type t_1, then the binding IC path consists of only type t_1, and the Pareto-dominating market contains types t_1 and t_2. If the procedure starts by choosing type t_2, then the binding IC path consists of only type t_2, and the Pareto-dominating market contains types t_2 and t_3.

weakly lower than t_j whose IR constraint may bind in the efficient mechanism for the market is type t_1, because an efficient allocation for type t_1 necessarily means that the surplus of all other types is strictly positive, so their IR cannot bind. As a result, the binding IC path must end in t_1.

VI. Discussion

This paper studies the existence of Pareto-improving segmentations, which weakly increase the surplus of all consumers and strictly increase the surplus of some consumers and the seller relative to the unsegmented market. We show that, generically, inefficient markets can be segmented in a way that is Pareto improving. This finding may contribute to discussions about regulating sellers' use of information and ability to price discriminate and consumers' privacy and control of their data. The result implies that, generically, consumers in inefficient markets can provide information to the seller (or allow this information to be collected by the seller) in a way that benefits all consumers and the seller.

To obtain this result, we developed a novel methodology that avoids the difficult problem of characterizing optimal menus with multiple products. The methodology relies on implications of binding incentive compatibility constraints when the seller optimally serves some types inefficiently. This methodology could potentially be useful in addressing other mechanism design questions in multiproduct settings that so far have proven intractable.

Our notion of Pareto improvements does not speak to the magnitude of the improvements. A natural question is how this magnitude changes with the number of alternatives and whether the improvements become small as the number of alternatives increases. Adding alternatives that are worse than the outside option clearly has no effect, but even adding alternatives that are valuable may have little or no impact. Consider, for example, adding an alternative that is equivalent to a convex combination of two existing alternatives.[28] The addition of this alternative has no welfare implications because it does not change the set of random alternatives. Thus, we can add an infinite number of such alternatives without changing the magnitude of any Pareto improvements.

But in many cases, adding valuable alternatives affects the scope and magnitude of the possible Pareto improvements. The improvements may also become arbitrarily small as the number of alternatives increases. This, however, may depend on how the improvements are measured, since Pareto improvements can increase the surplus of different consumers by different amounts. One possibility is to consider the average improvement

[28] That is, for some α in $[0, 1]$ and alternatives a_1 and a_2, every type t's valuation for the new alternative is $\alpha v(t, a_1) + (1 - \alpha)v(t, a_2)$.

over all consumers. In appendix A4, we construct a sequence of environments with two types and an increasing number of alternatives. We show that for half the markets, which are inefficient, the increase in the seller's revenue and in the average consumer surplus from any Pareto-improving segmentation approaches zero along the sequence. But even for these markets, the largest increase across all consumers does not approach zero. In the limit with a continuum of alternatives, these markets are not Pareto improvable, so the set of Pareto improvable markets is not generic in the set of inefficient markets.

Our main result shows that inefficient markets are generically Pareto improvable with two-market segmentations. This raises the question of whether some nongeneric markets are Pareto improvable but not by a two-market segmentation. Exploring this possibility is beyond the scope of this paper and likely requires a deeper understanding of how the surplus of different types varies across different markets. We point out, however, that in the example from section III.B, any Pareto improvable market (the inefficient markets other than market 0.75) is Pareto improvable by a two-market segmentation. We leave for future work a characterization of environments in which Pareto improvability is equivalent to Pareto improvability using two segments.

Finally, an important aspect of our analysis is that we do not impose any constraints on the set of segmentations we consider. In reality, the type of information that can be collected and the seller's ability to price discriminate on the basis of the available information may be restricted because of technological and other limitations, and these limitations may vary across settings. Our results establish that in general there is scope for Pareto improvements via segmentation. One direction for future research is to investigate specific settings by identifying and incorporating the limitations they imply.

Appendix

A1. Completing the Proof of Proposition 2

We complete the proof of proposition 2 in two steps. First, we show that mechanism M^* is optimal in the restricted environment, completing the proof of the "if" direction. Second, we prove the "only if" direction.

LEMMA 4. Consider an environment with a set of types $\{t_1, \dots, t_n\}$ and an efficient mechanism M^* such that the IC constraint from each type t_j to the next type t_{j+1} and the IR constraint for type t_n bind. There exists a market with full support over $\{t_1, \dots, t_n\}$ for which mechanism M^* is the unique optimal mechanism.

Proof. Consider any IC-IR mechanism (x, p) and any market f. Using the IC and IR constraints, we can write the expected revenue of the mechanism. For each type t_j, let $F(t_j) = \Sigma_{i \leq j} f(t_i)$ be the cumulative fraction of types t_1 to t_j. Now define $p(t_{n+1}) = 0$, $x(t_{n+1}) = 0$ (i.e., $x(t_{n+1})$ is a deterministic assignment of the outside option), and $F(t_0) = 0$ and write

$$\begin{aligned}\sum_j p(t_j)f(t_j) &= \sum_j (p(t_j) - p(t_{j+1}))F(t_j) \\ &\leq \sum_j \big(v(t_j, x(t_j)) - v(t_j, x(t_{j+1}))\big)F(t_j) \\ &= \sum_j \big(v(t_j, x(t_j))F(t_j) - v(t_{j-1}, x(t_j))F(t_{j-1})\big).\end{aligned} \tag{1}$$

Therefore, for any market f, the revenue of any IC-IR mechanism is at most the maximum of expression 1 over all allocation rules x.

By definition, the efficient alternative $\bar{a}(t_j)$ of type t_j satisfies $v(t_j, \bar{a}(t_j)) > v(t_j, a)$ for all alternatives $a \neq \bar{a}(t_j)$. Thus, if $F(t_{j-1})$ is small enough relative to $F(t_j)$, that is, $F(t_{j-1}) \leq \delta_j F(t_j)$ for some $\delta_j > 0$, then

$$v(t_j, \bar{a}(t_j))F(t_j) - v(t_{j-1}, \bar{a}(t_j))F(t_{j-1}) > v(t_j, a)F(t_j) - v(t_{j-1}, a)F(t_{j-1})$$

for all $a \neq \bar{a}(j)$. As a result, for such a market, the unique maximizer of $v(t_j, x)F(t_j) - v(t_{j-1}, x)F(t_{j-1})$ over all distributions x over alternatives is a distribution that assigns probability 1 to alternative $\bar{a}(t_j)$.

Now consider any market f with full support over the set of types $\{t_1, \ldots, t_n\}$ such that $F(t_{j-1}) \leq \delta_j F(t_j)$ for all j. By the above discussion, the allocation rule of the mechanism M^* is the unique maximizer of expression 1 over all allocation rules. In addition, since the IC constraint from each type t_j to the next type t_{j+1} and the IR constraint for type t_n bind, then the revenue of the mechanism M^* is equal to the maximum of expression 1 over all allocation rules. Thus, mechanism M^* is the unique optimal mechanism for market f. QED

Proof of proposition 2, the "only if" direction. Consider an efficient market f with an efficient optimal mechanism $M = (x, p)$, and suppose that some market f' with an optimal mechanism $M' = (x', p')$ Pareto dominates f. By Pareto dominance, for every type t in f', the surplus $v(t, x'(t)) - p'(t)$ of type t in M' is weakly higher than the surplus $v(t, \bar{a}(t)) - p(t)$ of type t in M and is strictly higher for some type. Since $v(t, \bar{a}(t)) \geq v(t, x'(t))$, we have that $p(t) \geq p'(t)$ for every type t in f', with a strict inequality for some type, so in market f', mechanism M generates a strictly higher revenue than mechanism M', a contradiction. QED

A2. *Single-Agent Interpretation*

Consider an agent whose type is drawn from the set T according to a prior distribution f. Before learning his type, the agent commits to an information disclosure policy, which maps every type in T to a distribution over signals. The seller observes the policy and the realized signal and forms a posterior f' over the agent' type. The seller then selects a mechanism to maximize revenue, and the agent responds by reporting his type optimally.[29] For which prior distributions f does there exist an information disclosure policy that, for each signal, increases the

[29] One motivating example is an online purchase setting in which the seller may be better than the consumer at determining which products are most appropriate for the consumer on the basis of personal data the consumer discloses (see Ichihashi 2020 for a discussion).

agent' ex post utility relative to a policy that discloses no information? This model and question are equivalent to those described earlier. Following Aumann, Maschler, and Stearns (1995) and Kamenica and Gentzkow (2011), we can describe the process as the agent choosing a distribution μ over posteriors f' that averages to f, that is, $E_{f'\sim\mu}[f'] = f$.

The single-agent model corresponds to a Bayesian persuasion setting (Kamenica and Gentzkow 2011) in which the agent is the sender and the seller is the receiver. The state is the sender's type, the receiver's set of actions is the set of IC-IR mechanisms, and the sender's state-dependent utility from the receiver's chosen mechanism (action) is the sender's utility from responding optimally to the mechanism. Existing results and techniques in the Bayesian persuasion literature concern the sender's expected utility, whereas our focus is on ex post utility.[30] In addition, no analytical description exists of the sender's state-dependent utility as a function of the receiver's action because there is no characterization of optimal mechanisms in our environment.

A3. *Appendix for Special Cases and Applications*

We formally define a lowest type and a best alternative and show that their existence implies the existence of Pareto-dominating markets with certain properties. We then discuss a bundling setting with multiple products, additive valuations, and zero production costs.

A3.1. Lowest Type

Type t in the support of a market is the lowest type in the market if $v(t, a) < v(t', a)$ for any type $t' \neq t$ in the support of the market and any alternative $a \neq 0$.

LEMMA 5. For any inefficient market f with a lowest type t, there exists a Pareto-dominating market that includes t. If t is not assigned the outside option in the optimal mechanism for f, then any Pareto-dominating market includes t.

Proof. If t is not assigned the outside option in the optimal mechanism for f, then because every other type can mimic t, the surplus of every type other than t is strictly positive. Thus, every market that Pareto dominates f contains t, because in any optimal mechanism some type has surplus zero. If t is assigned the outside option in the optimal mechanism for f, then the allocation of type t is inefficient and his surplus is zero. The proof of proposition 2 then shows that there exists a Pareto-dominating market that includes only type t and one other type. QED

A3.2. Best Alternative

Alternative a is the best alternative in a market if it is the efficient alternative for all types in the market.

[30] This ex post criterion may be relevant, e.g., when we would like to find improvements that work for all possible social welfare functions, which assign possibly different weights to different types.

LEMMA 6. For any inefficient market f with a best alternative a, there exists a Pareto-dominating market that includes only two types.

Proof. The result follows from the following observation: if an inefficient market has a best alternative, then in the optimal mechanism the IR constraint of some type t with an inefficient allocation binds. This observation implies the result because the proof of proposition 2 shows that some market that Pareto dominates the original market includes only type t and some type t' that is indifferent between mimicking type t and reporting truthfully.

To show the observation and complete the proof, choose an inefficient market with a best alternative, and suppose for contradiction that in the optimal mechanism the surplus of any type not assigned the best alternative is strictly positive. Consider such a type and the binding IC path that starts with this type and ends with a type whose IR constraint binds (as in the proof of proposition 2). By assumption, this latter type is assigned the best alternative. Thus, along the path there are consecutive types t and t' such that type t is not assigned the best alternative and type t' is assigned the best alternative. We argue that this contradicts the optimality of the mechanism for the market. Indeed, because type t is not assigned the best alternative, his valuation for his allocation is strictly lower than his valuation for the allocation of type t', which is the best alternative. The binding IC constraint from type t to t' then implies that the payment of type t is strictly lower than that of type t', that is, $p(t) < p(t')$. But then we can increase the mechanism's revenue by assigning type t the best alternative and charging him $p(t')$. Because no type's surplus is changed and the allocation and the payment of type t in this new mechanism is the same as the allocation and payment of type t' in the original mechanism, the new mechanism is IR and IC. QED

A3.3. Product Bundling

We now discuss an application with m products that may be bundled together. There are 2^m alternatives, which we call bundles, each corresponding to a subset of the set of products $\{1, \dots, m\}$. We refer to the alternative $\{1, \dots, m\}$ as the grand bundle.' Valuations are additive; that is, the valuation $v(t, b)$ of type t for a bundle b is the sum $\Sigma_{g \in b} v(t, \{g\})$ of his valuations for the individual products in that bundle. The cost of producing a bundle is the sum of the costs of producing the products in that bundle, and we assume that all these costs are zero (this may be the case for digital goods). We assume that there are at least two types and all types' valuations for each of the products are strictly positive. Therefore, the efficient alternative for each type is the grand bundle.

Regardless of the number of types, lemma 6 shows that any inefficient market has a Pareto-dominating market that contains only two types because the grand bundle is the best alternative.' But unlike the application with linear valuations, the binding IC paths do not have a prefix or interval structure. That is, different inefficient markets with the same support may have dominating markets in which the types with inefficient allocations whose IR constraint bind differ. We show this in an environment with three types and two products, which is illustrated in figure A1*A*. Each of the three types t_1, t_2, and t_3 is described by the circle with the type's label to its left. The horizontal axis shows the valuation for product 1, and the vertical axis shows the valuation for product 2.

Figure A1*B* depicts a mechanism in which product 1 is offered at price $v(t_2, \{1\})$, and the grand bundle is offered at price $v(t_3, \{2\}) + v(t_2, \{1\})$. At these prices, the light gray region contains the set of types that prefer product 1, the dark gray region contains the set of types that prefer the grand bundle, and the unshaded region contains the set of types that prefer the outside option. In particular, type t_2 is indifferent between product 1 and the outside option (and chooses product 1), and type t_3 is indifferent between product 1 and the grand bundle (and chooses the grand bundle). Type t_1 strictly prefers (and chooses) product 1. Consider a market for which this mechanism is optimal.[31] This market is inefficient because type t_2 is not assigned the grand bundle. To find a Pareto-dominating market, notice that because type t_2 is the only type whose IR constraint binds, any binding IC path necessarily ends in t_2. For example, a Pareto-dominating market may consist of types t_2 and t_3, where the fraction of type t_2 is large enough that it is optimal to assign both types the grand bundle for a price equal to the valuation of t_2. This mechanism is depicted in figure A1*C*. In this mechanism, the surplus of type t_3 is $(v(t_3, \{1\}) + v(t_3, \{2\})) - (v(t_2, \{1\}) + v(t_2, \{2\}))$, which is strictly higher than his surplus of $v(t_3, \{1\}) - v(t_2, \{1\})$ in the original market, because $v(t_3, \{2\}) > v(t_2, \{2\})$.

Now consider a market with all three types in which the fraction of type t_2 is large enough that the mechanism depicted in figure A1*C* is optimal. The IR constraints of both types t_1 and t_2 bind, and t_1 is the only type whose allocation is inefficient. Thus, unlike in the market discussed above, where any binding IC path ends in t_2, any binding IC path in this market necessarily ends in t_1. For this market, there are Pareto-dominating markets with two types, type t_1 and either type t_2 or type t_3, where the fraction of type t_1 is large enough that it is optimal to assign both types in the market the grand bundle for a price equal to the valuation of type t_1. This strictly increases the surplus of the other type in the market.

A4. Increasing the Number of Alternatives

We construct a sequence of environments with two types and linear valuations, as in section V. Along the sequence, the number of alternatives increases and for half the markets the gains from Pareto-improving segmentations become arbitrarily small.

The sequence of environments is $\mathcal{E}_1, \mathcal{E}_2, \ldots$, and the set of alternatives in environment $\mathcal{E}_k$ is $\{0, 1/k, 2/k, \ldots, k/k\}$. The cost of producing alternative a is $c(a) = a^2/2$. The two types, t_1 and t_2, have valuations a and $2a$ for alternative a, respectively. A market is identified by the proportion q of type t_2. In any market, the surplus of type t_1 in the optimal mechanism is zero. The surplus $\mathrm{CS}^k(t_2, q)$ of type t_2 in the optimal mechanism is depicted in figure A2*A*. The unit interval of all markets is partitioned into $k + 1$ intervals $[0, \tau_0], (\tau_0, \tau_1], \ldots, (\tau_{k-1}, \tau_k = 1]$ such that the surplus of type t_2 is constant within each interval, is lower on higher intervals, and is zero in the last interval. The surplus $\mathrm{CS}(t_2, q)$ of type t_2 in the limiting case with a continuum of alternatives $[0, 1]$ is depicted in figure A2*B*. The

[31] This mechanism is optimal, e.g., when the values are $(v(t_1, \{1\}), v(t_1, \{2\})) = (2.5, 1)$, $(v(t_2, \{1\}), v(t_2, \{2\})) = (2, 2)$, and $(v(t_3, \{1\}), v(t_3, \{2\})) = (3, 4)$ and the market is $f(t_1) = \varepsilon$, $f(t_2) = (1 - \varepsilon)(2/5)$, and $f(t_3) = (1 - \varepsilon)(3/5)$ for some small $\varepsilon > 0$.

surplus is strictly decreasing on the interval [0, 0.5] and is identically zero on the interval [0.5, 1].[32]

We start by considering the largest gain in average consumer surplus across all markets $q < 0.5$, where for each market we consider the Pareto-improving segmentation with the largest gain. We show that the largest gain across these markets approaches zero as the number k of alternatives increases. To see this, consider an environment $\mathcal{E}_k$ and a market q in some interval $(\tau_{k'-1}, \tau_{k'}]$ with $k' < k$. Any Pareto-improving segmentation of q may involve segments only in $[0, \tau_{k'}]$, since the surplus of type t_2 in any other segment is lower than his surplus in q. Since the expectation of the segments in any segmentation is q, to find the largest average gain across Pareto-improving segmentations of q, we concavify the average consumer surplus function $q \cdot CS^k(t_2, q)$ over the interval $[0, \tau_{k'}]$. The average consumer surplus is depicted by solid lines in figure A2*C*, and the average consumer surplus $q \cdot \mathrm{CS}(t_2, q)$ in the limiting case with a continuum of alternatives [0, 1] is depicted by a solid curve in figure A2*D*. It is easy to see that this surplus $q \cdot CS^k(t_2, q)$ is concave over the domain $\{0, \tau_1, ..., \tau_{k'}\}$, so the concavification linearly connects the average surplus function at consecutive markets in $\{0, \tau_1, \ldots, \tau_{k'}\}$. This is depicted by the dashed lines in figure A2*C*. In particular, the largest gain for market q is achieved by segmenting it into markets $\tau_{k'-1}$ and $\tau_{k'}$, which is Pareto improving for $q \neq \tau_{k'}$. As the number of alternatives goes to infinity, both the average consumer surplus function and its concavification converge uniformly (across markets) to the limit shown in figure A2*D*, so the largest average gain for all markets converges to zero.

However, not every consumer's gain converges to zero. To see this, consider an environment $\mathcal{E}_k$ and, for $k' \geq 2$, a market q in $(\tau_{k'-1}, \tau_{k'})$. This market can be segmented in a Pareto improving way into markets $\tau_1 - \varepsilon_k$ for small $\varepsilon_k > 0$ and $\tau_{k'}$ (we segment into $\tau_1 - \varepsilon_k$ and not τ_1 because our notion of Pareto improvement requires that the seller's revenue also increases). For large k, the surplus of the type t_2 consumers in segment $\tau_1 - \varepsilon_k$ increases from approximately $\mathrm{CS}(t_2, q)$ to approximately $\mathrm{CS}(t_2, 0)$, where CS is the surplus function in the limit with the continuum of alternatives [0, 1]. Therefore, for each market in $\cup_{k'\geq 2}(\tau_{k'-1}, \tau_{k'})$, which is generic in [0, 0.5], there is a positive measure set of consumers for whom the gain from some Pareto-improving segmentation is bounded away from zero. Therefore, for some $\gamma > 0$ and all large enough $k > 0$, each market q in a generic set of markets in [0, 0.5] has the property that a positive measure of type t_2 consumers gain at least γ from some Pareto improvement. Of course, this measure approaches zero as k grows large, since the weight of any market substantially lower than q in any Pareto-improving segmentation of q approaches zero. (This also shows that the seller's revenue increase from any Pareto-improving segmentation approaches zero.) In the limit with a continuum of alternatives [0, 1], inefficient markets in (0,0.5) are not Pareto improvable. Because this set is generic in the set of all inefficient markets, theorem 1 fails in the limit. The reason for this is that the perturbation argument in section IV.B no longer applies with infinitely many alternatives: perturbing a market in (0,0.5) necessarily changes the optimal mechanism.

[32] If $q \leq 0.5$, then type t_1 is assigned alternative $(1 - 2q)/(1 - q)$ at a price of $(1 - 2q)/(1 - q)$, type t_2 is assigned his efficient alternative, and the surplus of type t_2 is $(1 - 2q)/(1 - q)$. If $q > 0.5$, then type t_1 is assigned the outside option and type t_2 is assigned his efficient alternative at a price equal to his valuation.

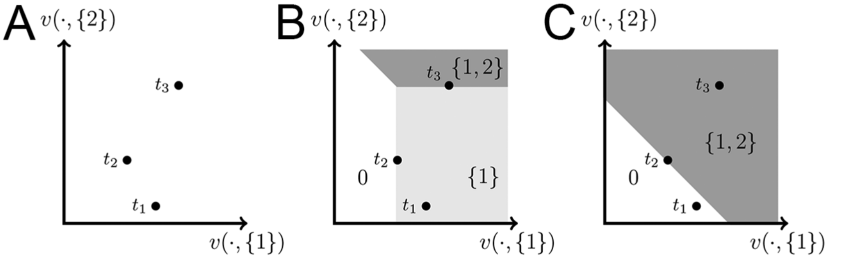

FIG. A1.—*A*, Environment with three types and two products. *B*, Mechanism in which product 1 is offered at price $v(t_2, \{1\})$ and grand bundle is offered at price $v(t_2, \{1\}) + v(t_3, \{2\})$. *C*, Mechanism in which grand bundle is offered at price $v(t_2, \{1\}) + v(t_2, \{2\})$. In each panel, types in the light gray region prefer product 1, types in the dark gray region prefer the grand bundle, and types in the unshaded region prefer alternative 0.

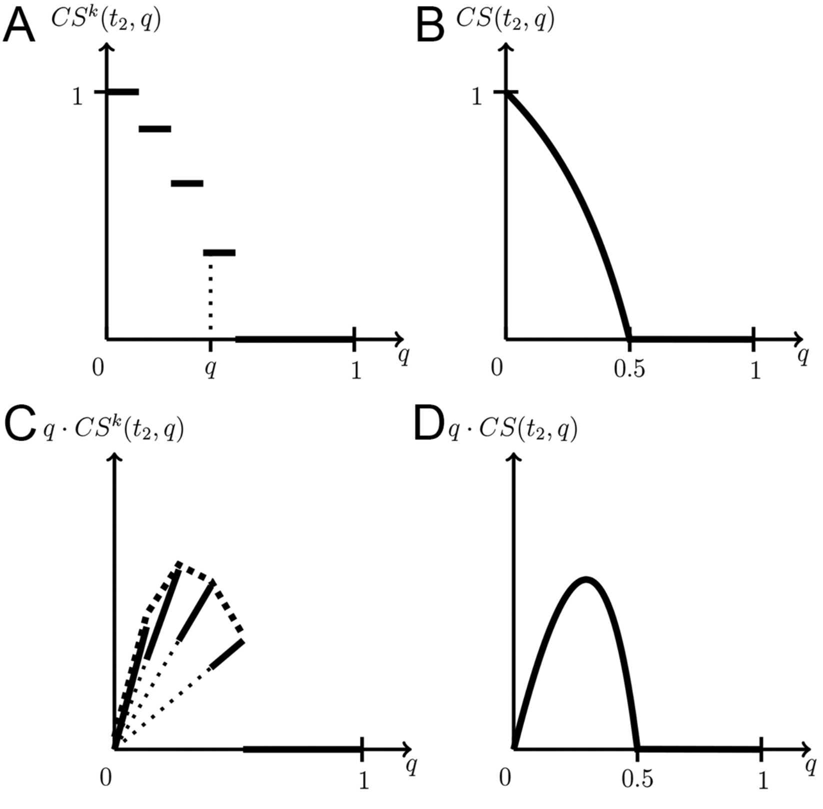

FIG. A2.—Surplus of type t_2 as a function of fraction $q = f(t_2)$ of consumers of type t_2 is shown for an environment $\mathcal{E}_k$ (*A*) and for the limiting environment with a continuum of alternatives [0, 1] (*B*). The average consumer surplus is shown for an environment $\mathcal{E}_k$ (*C*) and for the limiting environment (*D*).

References

Aguirre, I., S. Cowan, and J. Vickers. 2010. “Monopoly Price Discrimination and Demand Curvature.” *A.E.R.* 100 (4): 1601–15.

Ali, S. N., G. Lewis, and S. Vasserman. 2022. “Voluntary Disclosure and Personalized Pricing.” *Rev. Econ. Studies* 90 (2): 538–71.

Aumann, R. J., M. Maschler, and R. E. Stearns. 1995. *Repeated Games with Incomplete Information.* Cambridge, MA: MIT Press.

Bergemann, D., B. Brooks, and S. Morris. 2015. “The Limits of Price Discrimination.” *A.E.R.* 105 (3): 921–57.

Braghieri, L. 2019. “Targeted Advertising and Price Discrimination in Intermediated Online Markets.” Working paper.

Cowan, S. 2016. “Welfare-Increasing Third-Degree Price Discrimination.” *RAND J. Econ.* 47 (2): 326–40.

Dworczak, P., and G. Martini. 2019. “The Simple Economics of Optimal Persuasion.” *J.P.E.* 127 (5): 1993–2048.

Federal Trade Commission. 2012. “Protecting Consumer Privacy in an Era of Rapid Change: Recommendations for Businesses and Policymakers.” Washington, DC: Federal Trade Commission.

Glode, V., C. C. Opp, and X. Zhang. 2018. “Voluntary Disclosure in Bilateral Transactions.” *J. Econ. Theory* 175:652–88.

Haghpanah, N., and R. Siegel. 2022. “The Limits of Multi-Product Price Discrimination.” *A.E.R.* 4 (4): 443–58.

Hidir, S., and N. Vellodi. 2021. “Privacy, Personalization, and Price Discrimination.” *J. European Econ. Assoc.* 19 (2): 1342–63.

Ichihashi, S. 2020. “Online Privacy and Information Disclosure by Consumers.” *A.E.R.* 110 (2): 569–95.

Kamenica, E., and M. Gentzkow. 2011. “Bayesian Persuasion.” *A.E.R.* 101 (6): 2590–615.

Mussa, M., and S. Rosen. 1978. “Monopoly and Product Quality.” *J. Econ. Theory* 18 (2): 301–17.

Myerson, R. 1981. “Optimal Auction Design.” *Mathematics of Operations Research* 6 (1): 58–73.

Pigou, A. 1920. *The Economics of Welfare.* London: Macmillan.

Pram, K. 2021. “Disclosure, Welfare and Adverse Selection.” *J. Econ. Theory* 197: 105327.

Robinson, J. 1969. *The Economics of Imperfect Competition.* London: Palgrave Macmillan.

Schmalensee, R. 1981. “Output and Welfare Implications of Monopolistic Third-Degree Price Discrimination.” *A.E.R.* 71 (1): 242–47.

Varian, H. R. 1985. “Price Discrimination and Social Welfare.” *A.E.R.* 75 (4): 870–75.

Vohra, R. V. 2011. *Mechanism Design: A Linear Programming Approach.* Econometric Soc. Monographs, vol. 47. Cambridge: Cambridge Univ. Press.