Price Discrimination and Consumer Surplus in Multi-product Markets

Nima Haghpanah† Ron Siegel‡

August 25, 2020

Abstract

Consumer surplus in a market is affected by how the market is segmented. We study the maximum consumer surplus across all possible segmentations of a given market served by a multi-product seller. We characterize markets for which the maximum consumer surplus equals a first best benchmark of maximum total surplus minus minimum profit for the seller. We show that the first best benchmark is always achievable if and only if the seller never finds it profitable to screen types by offering multiple products or product bundles. The same condition also characterizes when the entire “surplus triangle” of Bergemann et al. (2015) can be achieved. These results highlight a novel impact of screening in settings that combine second and third degree price discrimination. They also help clarify to what extent the result of Bergemann et al. (2015) that the entire surplus triangle is always achievable in single-product markets relies on there being a single product.

*We thank Dirk Bergemann, Wouter Dessein, Harry Di Pei, Piotr Dworczak, 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 thank seminar participants at Hebrew University, Johns Hopkins University, Northwestern University, the Pennsylvania Economic Theory Conference, Summer Microeconomics Seminars 2019 (Hong Kong), and the University of Toronto.

†Department of Economics, Pennsylvania State University, haghpanah@psu.edu.

‡Department of Economics, Pennsylvania State University, ron.siegel@gmail.com.
1 Introduction

Consumer data, such as geographic location, age, or purchase history, is often used to identify different subsets of consumers and treat them as separate markets. The welfare effects of such market segmentation have been studied by a large and growing literature (e.g., Pigou 1920, Varian 1985, Bergemann et al. 2015). This literature typically assumes that the seller conducts third degree price discrimination, that is, sells a single product in each market segment. But a seller may be able to offer a variety of products, multiple quality or quantity levels of a given product, or bundles of heterogeneous products, and market segmentation allows such a seller to offer a possibly different menu of products and product bundles in each market segment. We study the welfare effects of market segmentation in such environments.

Consider a multi-product seller, for example an online retailer such as Amazon. The seller may have access to certain consumer data, and may be able to segment the market based on this data and offer each market segment a potentially different menu of products and product bundles to maximize profit. For instance, the seller may offer bundle discounts to consumers in certain locations, or offer products exclusively to consumers in certain age groups. The resulting consumer surplus (aggregated across consumers and across market segments) and producer surplus (the seller’s profit) depend on how the market is segmented, which in turn depends on the data the seller can access. But regardless of the segmentation, the resulting consumer-producer surplus pair must satisfy the following three constraints. First, producer surplus is at the least as high as in the unsegmented market, since the seller can offer the same menu in all market segments; second, consumer surplus is non-negative, since each consumer can opt not to purchase anything; and third, the sum of consumer and producer surplus is at most the total surplus generated by the efficient allocation.

The set of consumer-producer surplus pairs that satisfy these three constraints generates a “surplus triangle,” a term coined by Bergemann et al. (2015) (see Figure 1). This triangle is the convex combination of its three vertices. The top vertex corresponds to first-degree price discrimination: selling each consumer his preferred product at a price equal to the consumer’s willingness to pay for the product. The lower right vertex corresponds to the selling each consumer his preferred product at a price that does not increase producer surplus beyond that in the unsegmented market. We refer to the resulting consumer surplus as “first best consumer surplus,” since this is the highest consumer surplus in the surplus triangle. The lower left vertex
Figure 1: The surplus triangle is the convex combination of three vertices.

corresponds to consumer surplus of zero and producer surplus equal to that in the unsegmented market. The top vertex can be achieved by giving the seller detailed data on each consumer’s preferences and allowing him to make personalized offers. It is far less clear whether there are market segmentations that achieve the lower right or the lower left vertices. The difficulty is that, given a segmentation, in each market segment the seller offers a menu that maximizes his profit, without any regard to consumer surplus.

Remarkably, however, Bergemann et al. (2015) showed that any market with a single product can be segmented in ways that achieve the lower right vertex and the lower left vertex. Since all three vertices are achievable, the entire surplus triangle can be achieved by considering all possible segmentations of the market. In particular, for every market with a single product there exists some information about consumer preferences that, if provided to the seller who then uses it to segment the market and set the profit-maximizing price in each market segment, leads to first best consumer surplus, that is, consumers appropriate the entire increase in surplus from the efficient allocation. As we show in Section 4, this is no longer true in markets with two or more products.

Our goal is to understand when and to what extent first best consumer surplus and the (entire) surplus triangle are achievable in settings with multiple products. We provide economically meaningful conditions under which these achievability results hold and under which they fail. Our analysis also clarifies what drives the achievability results in settings with a single product.

In our setting, there are multiple types of consumers, where each type specifies a valuation for every product and product bundle. There is a “best product” or “grand bundle” of all products that all consumers agree is better than any other product or product bundle. Consumers’
valuations may be non-linear, and as a result the seller may find it profitable to screen consumers in a given market segment by offering a menu of products or product bundles, each targeted at a different subset of consumers. The possible profitability of screening consumers is a key distinguishing feature of our setting compared to most of the related literature on third degree price discrimination. As our results and an example in Section 2 demonstrate, this feature leads to first best consumer surplus and the surplus triangle not being achievable in general, unlike in settings with a single product.

First best consumer surplus is achievable for a given market if this surplus coincides with the maximum consumer surplus for that market, which is the highest (average) consumer surplus across all possible segmentations of the market. The maximum consumer surplus can in principle be identified using a concavification approach from Aumann et al. (1995) and Kamenica and Gentzkow (2011). In our setting, however, the concavification approach is of limited use. First, to apply concavification one must identify optimal (profit-maximizing) mechanisms for each market in order to identify consumer surplus in that market. But no general characterization of optimal mechanisms is known in settings with multiple products. Second, even if consumer surplus can be identified for all markets, it is not clear how to usefully compare its concavification to first best consumer surplus in each market. We obtain sharp characterizations of the achievability of first best consumer surplus by developing new techniques that do not rely on either a characterization of optimal mechanisms or concavification.

Our key findings relate the achievability of first best consumer surplus to the profitability of screening. For a given market, we say that screening is profitable if the optimal menu for that market includes at least two bundles. We provide four results which, roughly speaking, show that profitability of screening interferes with the achievability of first best consumer surplus and the surplus triangle. Our first result states that if screening is profitable for a market, then first

\[\text{More precisely, no characterization is known when valuations are non-linear in quantity. An alternative modeling approach would be to assume that valuations are linear but costs are non-linear, as in Mussa and Rosen (1978). In such a setting, ironing can be used to identify optimal mechanisms. Nevertheless, the difficulty with concavification persists since it requires identification of consumer surplus for all markets. A closed form expression for consumer surplus for all markets cannot be obtained since ironing needs to be performed differently for different markets.}\]

\[\text{Thus, the problem remains difficult even with a single product. Bergmann et al. (2015) develop methodologies to address achievability of first best consumer surplus for a single product.}\]

\[\text{More precisely, we say that screening is profitable if any menu that offers only one bundle is not optimal.}\]
best consumer surplus is unachievable for that market. If screening is profitable for a market, then, because not all consumers obtain the best product (or the grand bundle) in the optimal mechanism for that market, the resulting allocation is inefficient. Our result shows that it is impossible to restore efficiency via segmentation without the seller appropriating some of the gains.

Our other three results characterize when first best consumer surplus and the surplus triangle are achievable for either every market or no market with a given set of types. These results relate the achievability of first best consumer surplus for a market to the profitability of screening in other markets with the same set of types. Our second result shows that first best consumer surplus and the surplus triangle are achievable for all markets with a given set of types if and only if screening is not profitable for any market with that set of types. Our third result shows that first best consumer surplus is unachievable for any market with a given set of types that is not already efficient if and only if for any market with that set of types there is at most one price at which selling only the grand bundle is optimal for the seller. Our fourth result shows that the surplus triangle is unachievable for any market with a given set of types if and only if for any market with that set of types there is at most one price at which selling only the grand bundle is optimal for the seller. Combining these four results also provides a characterization of when first best consumer surplus and the surplus triangle are achievable for some but not all markets with a given set of types.

These characterizations have clear economic interpretations in terms of screening, but may be difficult to apply in specific settings since they rely on identifying optimal mechanisms. Building on the results of Haghpanah and Hartline (2020), we provide equivalent characterizations in terms of properties of the set of types, which are a primitive of the setting. First best consumer surplus and the surplus triangle are achievable for all markets with a given set of types if and only if for any pair of types and any bundle of products, the ratio of the valuations of that bundle to the grand bundle of all products is larger for the type with the higher valuation for the grand bundle. Roughly speaking, this happens when lower types consider the products to be more complementary, in that lower types have a higher relative value than higher types for adding any set of products to its complement. Our results imply that in such settings, if consumers control the information that the seller can access, they can obtain first best consumer surplus. And a social planner who controls the information can achieve any consumer-producer surplus pair in the surplus triangle. On the other hand, first best consumer surplus and the surplus
triangle are unachievable for all markets with a given set of types if and only if for every pair of types, there exists a bundle of products such that the ratio of valuations of that bundle to the grand bundle of all products is smaller for types with a higher valuation for the grand bundle. We elaborate on the economic interpretation of this condition in Section 6.

Related Work. Our work connects the literature on 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 where a segmentation may decrease total and hence consumer surplus. Followup work provides conditions for a segmentation to increase or decrease total surplus or consumer surplus (Robinson 1969; Schmalensee 1981; Varian 1985; Aguirre et al. 2010; Cowan 2016). Our work differs from this literature in two 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 screen consumers in each market by offering a menu of multiple bundles. Second, with the exceptions we now discuss, the literature assumes that the segmentation is exogenously fixed.

A growing part of the literature on third degree price discrimination studies surplus across all possible segmentations of a given market for a single product. The most closely related paper to ours is Bergemann et al. (2015), who identify the set of all producer and consumer surplus pairs that can result from some segmentation of a given market. As we mentioned earlier, they show that in every market for a single product the set of these pairs coincides with the surplus triangle. Our results show that the set of surplus pairs may not coincide with the surplus triangle when there are multiple products. Glode et al. (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 (2020) consider maximum consumer surplus when a multi-product seller offers a single product to each market. 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 (2020) characterize optimal segmentation with a continuum of products. Braghieri (2017) studies market segmentation with a continuum of firms, each producing a single differentiated product. In contrast, the seller in our setting may offer multiple products in a market in order
to screen consumers.\footnote{The only instance of this we are aware of is \cite{Bergemann2015}'s parametric example with two types and non-linear valuations.} The literature on multi-product bundling goes back to \cite{Stigler1963} and \cite{Adams1976}, who study bundling as an instrument to engage in second degree price discrimination. Theoretical findings on welfare effects of bundling are inconclusive.\footnote{The main hurdles are the difficulty with identifying optimal mechanisms and their complexity. \cite{Thanassoulis2004} and \cite{Daskalakis2017} show that optimal mechanisms may offer randomized bundles. \cite{Vincent2007} and \cite{Hart2013} show that the optimal menu may offer infinitely many bundles. \cite{Daskalakis2014} and \cite{Chen2015} show that the problem of finding optimal mechanisms is computationally intractable. A more recent literature empirically estimates the welfare effects of bundling \cite{Ho2012, Crawford2012}.} The main hurdles are the difficulty with identifying optimal mechanisms and their complexity. \cite{Thanassoulis2004} and \cite{Daskalakis2017} show that optimal mechanisms may offer randomized bundles. \cite{Vincent2007} and \cite{Hart2013} show that the optimal menu may offer infinitely many bundles. \cite{Daskalakis2014} and \cite{Chen2015} show that the problem of finding optimal mechanisms is computationally intractable. A more recent literature empirically estimates the welfare effects of bundling \cite{Ho2012, Crawford2012}.

Parts of our analysis use results from the literature on information design and multi-dimensional screening. \cite{Haghpanah2020} identify conditions under which selling only the grand bundle of products is optimal. Under these conditions we can use a segmentation from \cite{Bergemann2015} to achieve the surplus triangle. To show that these conditions are necessary, however, and to identify markets for which first best consumer surplus is unachievable, we develop a novel approach that does not rely on a characterization of optimal mechanisms.

All proofs not given in the text are in the appendix.

## 2 An Example

In this section we discuss a parametric example to highlight our results. We directly calculate the closed form expression for the maximum consumer surplus and compare it to first best consumer surplus. Even though the calculations are straightforward, they are not easily extendable beyond this example. A reader who is only interested in our general treatment can skip ahead to the next section.

\footnote{We reiterate that we use the term screening to mean that there are at least two bundles in the seller’s menu. A mechanism that offers a single product at a high price and therefore excludes certain consumers is not a screening mechanism.}

\footnote{\cite{Adams1976} show that bundling may be inefficient as it leads to oversupply or undersupply of certain goods. \cite{Salinger1995} argues that bundling may result in lower or higher prices and therefore may increase or decrease consumer surplus.}
Consider a seller who produces a single product at zero cost and a population of consumers who each demand at most two units of the product. There are two types of consumers. A type 1 consumer has valuation \( v \in (0, 1) \) for one unit and valuation 1 for two units. A type 2 consumer has valuation 1 for one unit and valuation 2 for two units. The two types are illustrated in Figure 2, in which case (a) corresponds to \( v \leq 0.5 \) and case (b) corresponds to \( v \geq 0.5 \). A market \( q \) consists of a fraction \( 1 - q \) of type 1 consumers and a fraction \( q \) of type 2 consumers.

To identify maximum consumer surplus in different markets, it is useful to first identify the optimal mechanism in each market. Consider the following three mechanisms and their revenue in a market \( q \). Mechanism \( N_1 \) offers two units of the product (as a bundle) at price 1. Mechanism \( N_2 \) offers two units (as a bundle) at price 2. Mechanism \( S \) screens; it offers each consumer a choice between buying one unit at price \( v \) or two units at price \( v + 1 \). It can be shown that for any market \( q \), one of these three mechanisms is optimal, as illustrated in Figure 3. If \( v \leq 0.5 \), then mechanisms \( N_1 \) is optimal for markets in \([0, 0.5] \) and mechanism \( N_2 \) is optimal for markets in \([0.5, 1] \). If \( v \geq 0.5 \), then mechanism \( N_1 \) is optimal for markets in \([0, 1 - v] \), mechanism \( S \) is optimal for markets in \([1 - v, v] \), and mechanism \( N_2 \) is optimal for markets in \([v, 1] \).

---

\(^6\)An equivalent interpretation is a seller who can produce a product at two quality levels (e.g., a smartphone with low or high memory) for which consumers have unit demand.
Next, we compute the (average) consumer surplus in each market $q$ generated by the optimal mechanism for that market. Type 1 does not receive any information rents in any optimal mechanism. Thus, consumer surplus $CS(q)$ in market $q$ is $q$ times the utility of type 2 in the optimal mechanism for that market. Consumer surplus $CS(q)$ is illustrated in Figure 4.

A segmentation of market $q$ is a distribution $\mu$ over markets $[0,1]$ such that $E_{q' \sim \mu}[q'] = q$. The maximum consumer surplus is $MCS(q) = \max_\mu E_{q' \sim \mu}[CS(q')]$, that is, the highest consumer surplus across all segmentations $\mu$. The maximum consumer surplus is obtained by concavifying the function $CS$. That is, $MCS(q) = \overline{CS}(q)$, where $\overline{CS}$ is the lowest concave function that is point-wise at least as high as $CS$.

The maximum consumer surplus $MCS(q)$ is at least $CS(q)$ and at most first best consumer surplus $FBCS(q)$, which is the surplus from the efficient allocation (that is, two units for each type) minus the seller’s revenue in market $q$. If the optimal mechanism for market $q$ implements the efficient allocation, then the two bounds are equal, that is, $CS(q) = FBCS(q)$, so $CS(q) = MCS(q) = FBCS(q)$. This is the case for a market $q$ for which mechanism $N^1$ is optimal and for market $q = 1$ which contains only type 2 consumers and for which mechanism $N^2$ is optimal. We refer to such markets as efficient, and otherwise as inefficient. If a market is efficient, then there is no scope for market segmentation to increase consumer surplus.

We can now address the possibility of achieving first best consumer surplus for all markets $q \in [0,1]$. The relationship between maximum consumer surplus, $MCS$, and first best consumer surplus, $FBCS$, is illustrated in Figure 5, and depends on the value of $v$. If $v$ is in $(0, 0.5)$, as in Figure 5 (b) and (c), then first best consumer surplus is not achievable for any inefficient market. The only difference between cases (b) and (c) in Figure 5 is that in the former, $MCS(q)$ strictly exceeds $CS(q)$ for every inefficient market $q$ whereas in the latter $MCS(q) = CS(q)$ for market

\footnote{If there is more than one optimal mechanism we choose the one with higher consumer surplus.}
If $v \in (0, 0.5]$, as in Figure 5(a), then first best consumer surplus is achievable for all markets. Equivalently, first best consumer surplus is achievable for all markets if and only if for every market either mechanism $N^1$ or $N^2$ is optimal, that is, the seller does not find it profitable to screen consumers.

What is the economic significance of $v$ being greater than or smaller than 0.5? For type 2 consumers, two units of the products are twice as valuable as one unit. For type 1 consumers, whether $v$ is greater than or smaller than 0.5 determines whether two units of the product are more than or less than twice as valuable than one unit. In other words, when $v \leq 0.5$ the second unit of the product is relatively more complementary to the first unit of the product for type 1 consumers than for type 2 consumers, and vice versa when $v > 0.5$. In the former case, the results above mean that for any inefficient market there is some information that consumers can provide to the seller that will induce him to set prices that give consumers the highest possible surplus. In the latter case, there is no such information, even for inefficient markets $q$ for which the seller does not find it optimal to screen consumers.

Turning to the surplus triangle, it is trivially achievable for markets with a single type of consumer ($q = 0$ and $q = 1$). For all other markets, the same conditions that characterize achievability of first best consumer surplus also characterize when the surplus triangle is achievable for every market or no market (efficient or inefficient). Indeed, whenever first best consumer surplus is not achievable, the surplus triangle is clearly not achievable. And an observation based on the results of Bergemann et al. (2015) shows that when the seller does not find it optimal to screen, that is, in every market $q$ only offers two units as a bundle, the entire surplus triangle is

---

8See Haghpanah and Siegel (2020) for a detailed investigation of when MCS strictly exceeds CS.

9Notice that type 1 consumers will, of course, still obtain a surplus of 0.
achievable.

We will show that all of these results extend to any setting with two types (even with multiple products and multi unit demand), although the proofs require developing new techniques to characterize optimal mechanisms. We then generalize the results to settings with more than two types, in which optimal mechanisms cannot in general be characterized. One complication is that with more than two types it is not always the case that first best consumer surplus is either achievable for all markets or not achievable for all inefficient markets: It may be that it is achievable for some but not all markets, and similarly for the surplus triangle. We characterize precisely when this occurs both in terms of the optimality of screening and in terms of the ranking of types according to the relative complementarity between individual products and the grand bundle of all products.

3 Model

There is a mass 1 of consumers and a set $T = 1, \ldots, n$ of consumer types. There is a set $A = 0, 1, \ldots, k$ of alternatives, where $k \geq 1$ and alternative 0 is the outside option. An alternative can correspond to a particular quantity or quality of a product or to a bundle of products. For example, if there are two products that can be bundled, then there are four alternatives: the outside option, product 1, product 2, and a bundle that includes both products. The cost of production is zero. The valuation of type $i \in T$ for an alternative $a \in A$ is $v_i(a) \geq 0$, with $v_i(0) = 0$. Assume without loss of generality that there are no redundant alternatives, that is, for each pair of alternatives $a \neq a'$, there exists a type $i$ such that $v_i(a) \neq v_i(a')$. We assume that some alternative $\bar{a} \in A$ is all types’ most preferred alternative, that is, $v_i(\bar{a}) > v_i(a)$ for all types $i$ and alternatives $a \neq \bar{a}$. In a bundling application, it is natural to think of $\bar{a}$ as the grand bundle of all products. We do not assume that alternatives are otherwise ranked. We assume that a higher type has a higher valuation for any alternative, that is, $v_1^i < v_2^i < \ldots < v_n^i$ for all alternatives $a \neq 0$. Importantly, we do not require a single-crossing property.

An allocation $x \in X = \Delta(A)$ is a distribution over alternatives, where $x_a$ denotes the probability of alternative $a$. The (expected) utility of a type $i$ from an allocation $x$ and a payment $p$ is $v_i \cdot x - p = (\sum_a v_i^a x_a) - p$. We say that an allocation $x$ is empty if $x_0 = 1$, and is

\[10\] The finiteness assumption simplifies some of the notation and proofs. All of our results continue to hold when the sets of types and alternatives are infinite but compact.
non-empty otherwise. Note that for each type the efficient allocation $x$ satisfies $x_a = 1$.

A mechanism consists of an allocation function $x : T \rightarrow X$ and a payment function $p : T \rightarrow R$. Mechanism $M = (x, p)$ is incentive compatible (IC) if for all types $i$ and $j$,

$$v^i \cdot x(i) - p(i) \geq v^i \cdot x(j) - p(j).$$

That is, type $i$ does not benefit from mimicking type $j$. Mechanism $M$ is individually rational (IR) if for all types $i$,

$$v^i \cdot x(i) - p(i) \geq 0.$$

Henceforth, by “mechanism” will refer to an IC and IR mechanism, unless otherwise stated.

A market $f \in \Delta(T)$ is a distribution over types, where $f_i$ denotes the fraction of consumers with type $i$. The expected utility of consumers in a market $f$ and mechanism $M = (x, p)$ is $EU(f, M) = E_i[f \cdot v^i \cdot x(i) - p(i)]$. A mechanism $(x, p)$ is optimal for market $f$ if it maximizes revenue

$$E_i[f \cdot p(i)]$$

across all mechanisms. For a market $f$, let $ER(f)$ be the maximum expected revenue, $\mathcal{M}(f)$ be the set of optimal mechanisms, and $CS(f)$ be the highest consumer surplus (expected utility) across all optimal mechanisms,

$$CS(f) = \max_{M \in \mathcal{M}(f)} EU(f, M).$$

A segmentation $\mu \in \Delta(\Delta(T))$ of a market $f$ is a distribution over markets that average to $f$, that is, $E_{f' \sim \mu}[f'] = f$. We refer to a market $f'$ in the support of the segmentation $\mu$ as a market segment (or simply a segment). Let $SEG(f)$ denote the set of segmentations of $f$. Abusing notation, let $CS(\mu)$ be the consumer surplus in the segmentation $\mu$,

$$CS(\mu) = E_{f' \sim \mu}[CS(f)].$$

When discussing segmentations of a given market $f$, we refer to $f$ as the unsegmented market.

We often represent a mechanism indirectly by a menu of allocation-price pairs, where each type chooses a pair that maximizes its utility. If a type is indifferent between two allocation-price

\footnote{An optimal mechanism exists: IR implies that the revenue of any mechanism is at most $E_{i,\sim f}[v_a^i]$, and the set of mechanisms is closed.}

11
pairs, it chooses the one with a higher price. If, further, the prices are identical, then the tie breaking can be arbitrary since it does not affect consumer surplus or revenue. Assume without loss of generality that for each allocation-price pair, there exists a type that chooses that pair. Unless stated otherwise, every menu includes the outside option at price zero.

We distinguish between screening and non-screening mechanisms. A mechanism \((x, p)\) is a non-screening mechanism if it can be represented by a menu with a single allocation-price pair, in addition to the outside option at price 0. Of particular interest is the set of non-screening mechanisms \(\{N^i\}_{i \in T}\), where mechanism \(N^i\) offers the most preferred alternative \(\bar{a}\) at price \(v^i_{\bar{a}}\) (in addition to the outside option at price 0). In particular, the allocation and payment functions \((x, p)\) of mechanism \(N^i\) are as follows: \(x_0(j) = 1\) and \(p(j) = 0\) for all \(j < i\), and \(x_{\bar{a}}(j) = 1\) and \(p(j) = v^i_{\bar{a}}\) for all \(j \geq i\). Among all non-screening mechanisms, \(N^i\) is optimal for some \(i^{12}\). A mechanism is a screening mechanism if it is not a non-screening mechanism, that is, every menu that represents it must include at least two allocation-price pairs (in addition to the outside option at price 0). In such a mechanism, there exist two types with different and non-empty allocations. We say that non-screening is optimal for a market \(f\) if for some \(i\), mechanism \(N^i\) is optimal. Otherwise, that is, if \(N^i\) is not optimal for any \(i\), we say that screening is optimal for market \(f\). In this case, any optimal mechanism for market \(f\) is a screening mechanism.

3.1 Upper Bound On The Maximum Consumer Surplus

For a given market \(f\), we denote by \(\text{MCS}(f) = \max_{\mu \in \text{SEG}(f)} \text{CS}(\mu)\) the maximum consumer surplus across all segmentations of \(f\), and refer to a segmentation that achieves the maximum as a consumer-optimal segmentation. Clearly, \(\text{MCS}(f)\) is at least \(\text{CS}(f)\), since the distribution \(\mu\) over markets that puts probability 1 on market \(f\) is a segmentation of \(f\). Also \(\text{MCS}(f)\) is at most the expected surplus of an efficient allocation, \(E_{i \sim f} [v^i_{\bar{a}}]\), minus the maximum expected revenue, \(\text{ER}(f)\). The reason is that for any segmentation, (1) the sum of the expected revenue and the consumer surplus is at most \(E_{i \sim f} [v^i_{\bar{a}}]\), and (2) the expected revenue is at least \(\text{ER}(f)\), since the seller can choose a mechanism in \(\mathcal{M}(f)\) for all markets segments. Thus, consumer

\(^{12}\)Consider any non-screening mechanism \(M\) that offers a single allocation \(x\) for price \(p\). The mechanism that offers \(\bar{a}\) at price \(p\) obtains at least as much revenue. Further, a price \(p\) such that \(v^i_{\bar{a}}^{-1} \leq p < v^i_{\bar{a}}\) for some \(i \in T\) cannot be optimal, since offering \(\bar{a}\) at price \(v^i_{\bar{a}}\) generate more revenue. Thus, it is optimal to offer \(\bar{a}\) at price \(v^i_{\bar{a}}\) for some \(i\) among all non-screening mechanisms.
surplus is at most $E_{i \sim f}[v_i^a] - ER(f)$, which we refer to as first best consumer surplus. The following lemma formalizes this discussion.

**Lemma 1** For any market $f$, $CS(f) \leq MCS(f) \leq E_{i \sim f}[v_i^a] - ER(f)$.

We study when the upper bound is tight.

**Definition 1**
1. First best consumer surplus is achievable for a market $f$ if $MCS(f) = E_{i \sim f}[v_i^a] - ER(f)$.
2. A segmentation $\mu$ of market $f$ achieves first best consumer surplus if $CS(\mu) = E_{i \sim f}[v_i^a] - ER(f)$.

If a market $f$ has an optimal mechanism with an efficient allocation, then first best consumer surplus is clearly achievable for $f$: the surplus generated is $E_{i \sim f}[v_i^a]$ and the seller’s profit is $ER(f)$. Thus, $CS(f) = E_{i \sim f}[v_i^a] - ER(f)$. We refer to such markets as efficient.

**Definition 2** A market $f$ is efficient if $N_i(f)$ is an optimal mechanism for the market, where $i(f)$ is the lowest type in the support of $f$. Otherwise, the market is inefficient.

### 3.2 The Surplus Triangle

Given a market $f$, denote by $\Gamma(f)$ the set of consumer-producer surplus pairs resulting from all possible segmentations of $f$. Let $ER(\mu) = E_{f \sim \mu}[ER(f)]$ be the producer surplus resulting from a segmentation $\mu$, and consider a consumer-producer surplus pair $(CS(\mu), ER(\mu))$. Since $ER(\mu) \geq ER(f)$, $CS(\mu) \geq 0$, and $CS(\mu) + ER(\mu) \leq E_{i \sim f}[v_i^a]$, the set $\Gamma(f)$ is a subset of the “surplus triangle”

$$\Delta(f) = \{(a, b) : b \geq ER(f), a \geq 0, a + b \leq E_{i \sim f}[v_i^a]\},$$

which is illustrated in Figure 1. If every pair in $\Delta(f)$ results from some segmentation of $f$, we say that the surplus triangle is achievable.

**Definition 3** The surplus triangle is achievable for a market $f$ if $\Gamma(f) = \Delta(f)$.

Bergemann et al. (2015) coined the term “surplus triangle,” and showed that it is achievable for any market $f$ with a single alternative. The surplus triangle is also obviously achievable for
any “singleton market,” which consists only of consumers of some single type \( i \). In this case, the surplus triangle consists of the single pair \((0, v^i_a)\). For other markets, however, Section 2 above shows the surplus triangle is not always achievable when there are multiple alternatives. To proceed, observe that the surplus triangle is the convex hull of its vertices, and a convex combination of segmentations is a segmentation whose consumer-producer surplus pair is the same convex combination of the consumer-producer surpluses of the segmentations. Thus, to determine whether the surplus triangle is achievable it is enough to determine whether each of the three vertices of the surplus triangle is generated by some segmentation. The top vertex, \((0, E_i \sim f[v^i_a])\), is generated by segmenting consumers according to their type, and corresponds to first-degree price discrimination. The lower right vertex, \((ER(f), E_i \sim f[v^i_a] - ER(f))\), is generated by segmentations that achieve first best consumer surplus. The lower left vertex, \((0, ER(f))\), generates the lowest possible total surplus of \( ER(f) \).

**Definition 4** A segmentation \( \mu \) of market \( f \) achieves the lowest possible total surplus if the resulting consumer-producer pair is \((0, ER(f))\). If such a segmentation exists then the lowest possible total surplus is achievable for market \( f \).

The above discussion shows the following.

**Lemma 2** The surplus triangle is achievable for a market if and only if first best consumer surplus and the lowest possible total surplus are achievable for the market.

The rest of the paper investigates when first best consumer surplus is achievable for all markets, some markets, or no inefficient markets and when the surplus triangle is achievable for all markets, some markets, or no non-singleton markets (efficient or inefficient).

### 3.3 Preliminary Observations

We start with two useful observations.

The first observation specifies two conditions, which are together necessary and sufficient for a segmentation to achieve first best consumer surplus. First, total surplus must be maximized. That is, every segment must be efficient [Definition 2]. Second, the seller should not benefit from the segmentation. The seller benefits from the segmentation if an optimal mechanism for the
unsegmented market is no longer optimal for some segment. Thus, every optimal mechanism for the unsegmented market must also be optimal for every segment. This is summarized by the following lemma.

**Lemma 3** For any segmentation $\mu$ of a market $f$, the following are equivalent:

1. $\mu$ achieves first best consumer surplus.

2. for some optimal mechanism $M$ of $f$ and every segment $f'$ of $\mu$, $f'$ is efficient and has an optimal mechanism $M$.

3. for every optimal mechanism $M$ of $f$ and every segment $f'$ of $\mu$, $f'$ is efficient and has an optimal mechanism $M$.

The second observation uses Lemma 3 to obtain sufficient conditions for achieving first best consumer surplus. We describe a class of segmentations used by Bergemann et al. (2015), starting with the following definition.

**Definition 5** A market $f$ is an equal-revenue market if for every type $i$ in the support of $f$, the mechanism $N^i$ is optimal.

Consider a market $f$ for which the mechanism $N^i$ is optimal for some $i$, and a segmentation $\mu$ of $f$. Suppose that every segment $f'$ is an equal-revenue market and includes $i$ in its support. Notice that $f'$ is efficient since any equal-revenue market is efficient. Additionally, by definition, $N^i$ is optimal for $f'$. Thus, $\mu$ achieves first best consumer surplus by Lemma 3. We therefore have the following corollary.

**Corollary 1** Consider a market $f$ for which $N^i$ is optimal for some $i$. A segmentation $\mu$ of $f$ achieves first best consumer surplus if for every segment $f'$ of $\mu$, $f'$ is an equal-revenue market and includes $i$ in its support.

---

13This is because the seller can offer in this segment an optimal mechanism for this segment, and in all other segments an optimal mechanism for the unsegmented market.
4 Two Types

We begin by considering markets with only two types of consumers, 1 and 2, and any number $k \geq 1$ of alternatives. Subsequent sections consider markets with more than two types. With two types, optimal mechanisms can be characterized in closed form. We defer the characterization to Appendix B.1 and invoke here only the properties that are needed for the analysis.

We identify each market by its fraction $q \in [0, 1]$ of type 2 consumers. The following lemma shows that the set of markets $[0, 1]$ can be qualitatively divided into at most three regions. The first region consists of markets for which mechanism $N_1$ (which sells only the best alternative $\bar{a}$ at price $v_1$) is optimal. These are markets in which the fraction of type 1 consumers is high. The second region consists of markets for which mechanism $N_2$ (which sells $\bar{a}$ at price $v_2$) is optimal. These are markets in which the fraction of type 2 consumers is high. If these two regions do not cover the whole interval $[0, 1]$, then the third region consists of the remaining, intermediate markets. For these markets, neither $N_1$ nor $N_2$ is optimal. Rather, screening is optimal, that is, the allocations of the two types are different and non-empty. Moreover, the optimal mechanisms may vary across markets in this region. To formalize this, denote by $F(M)$ the (possibly empty) set of markets for which a particular mechanism $M$ is optimal.

**Lemma 4** There exist thresholds $q_1$ and $q_2$, $0 < q_1 \leq q_2 < 1$, such that $F(N_1) = [0, q_1]$, $F(N_2) = [q_2, 1]$, and $F(M) \subseteq [q_1, q_2]$ for any mechanism $M \neq N_1, N_2$.

Given the above lemma, we can distinguish two cases. Either $q_1 = q_2$, in which case for any market either $N_1$ or $N_2$ is optimal, or $q_1 < q_2$, in which case neither $N_1$ nor $N_2$ is optimal for markets in the interval $(q_1, q_2)$ (see Figure 6). The proposition below shows that if $q_1 = q_2$ then first best consumer surplus is achievable for all markets, and if $q_1 < q_2$ then first best consumer surplus is unachievable for any inefficient market (the markets in $(q_1, 1)$).

**Proposition 1** For any inefficient market $q$, first best consumer surplus is achievable if and only if $q_1 = q_2$.

14If $q \in [0, q_1]$, then mechanism $N_1$ is optimal. If $q = 1$, then mechanism $N_2$ is optimal and specifies an efficient allocation for all types in the market (only type 2 is in the market).
Figure 6: (a) $q_1 = q_2$. For any market, either $N^1$ or $N^2$ is optimal. (b) $q_1 < q_2$. Neither $N^1$ nor $N^2$ is optimal for markets in the interval $(q_1, q_2)$.

Proof. Suppose first that $q_1 = q_2$. We show that for any market $q$, the exists a segmentation that achieves first best consumer surplus. The segmentation is identical to that of Bergemann et al. (2015). First consider $q \in [0, q_1]$, which implies that the mechanism $N^1$ is optimal for $q$. Consider a segmentation of $q$ into two segments $q' = 0$ and $q'' = q_1 = q_2$\footnote{Segmentation $\mu$ assigns probability $\alpha$ to segment $q''$ and probability $1 - \alpha$ to segment $q'$, where $\alpha = \frac{q}{q_2}$. Since $q = (1 - \alpha) \cdot 0 + \alpha \cdot q_2$, $\mu$ is a segmentation of $q$.} Notice that both $q'$ and $q''$ are equal-revenue markets \footnote{Segmentation $\mu$ assigns probability $\alpha$ to segment $q'$, and probability $1 - \alpha$ to segment $q''$, where $\alpha = \frac{q - q_2}{q_2 - q}$. Since $q = \alpha \cdot 1 + (1 - \alpha) \cdot q_2$, $\mu$ is a segmentation of $q$.} and include type 1 in their support. Thus the segmentation achieves first best consumer surplus by Corollary 1. Now consider $q \in [q_2, 1]$, which implies that the mechanism $N^2$ is optimal for $q$. Consider a segmentation of $q$ into two segments $q' = 1$ and $q'' = q_1 = q_2$\footnote{Segmentation $\mu$ assigns probability $\alpha$ to segment $q''$ and probability $1 - \alpha$ to segment $q'$, where $\alpha = \frac{q}{q_2}$. Since $q = (1 - \alpha) \cdot 0 + \alpha \cdot q_2$, $\mu$ is a segmentation of $q$.} Notice that both $q'$ and $q''$ are equal-revenue markets and include type 2 in their support. Thus the segmentation achieves first best consumer surplus by Corollary 1.

Now suppose that $q_1 < q_2$, and suppose that some segmentation $\mu$ of a market $q$ achieves first best consumer surplus. We show that $q$ must be efficient, that is, $q$ must be in $[0, q_1] \cup \{1\}$. By Lemma 3, $\mu$ must satisfy two conditions. First, every segment in $\mu$ must be efficient, that is, it must be in $[0, q_1] \cup \{1\}$. Second, any optimal mechanism for $q$ must be also optimal for every segment. By Lemma 4, the only optimal mechanism for market 1 is mechanism $N^2$. But since $q_1 < q_2$ and $\mathcal{F}(N^2) = [q_2, 1]$, $N^2$ is not optimal for any market in $[0, q_1]$. Therefore, either every segment of $\mu$ must be equal to 1, in which case $q = 1$, or every segment must be in $[0, q_1]$, in which case $q \in [0, q_1]$. Therefore, $q$ is efficient. ■

We now turn to the achievability of the surplus triangle. Proposition 2 will show that if $q_1 = q_2$ then the surplus triangle is achievable for all markets, and if $q_1 < q_2$ then the surplus
triangle is not achievable for any non-singleton market, that is, for any market in \((0, 1)\). Recall from Section 3 that the surplus triangle is achievable for all singleton markets. We thus obtain a characterization of the achievability of the surplus triangle.

**Proposition 2** For any non-singleton market \(q\), the surplus triangle is achievable if and only if \(q_1 = q_2\).

Proposition 1 and Proposition 2 provide a complete characterization of when first best consumer surplus and the surplus triangle are achievable when there are only two types. The characterization is in terms of the regions for which different mechanisms are optimal. Haghanpanah and Hartline (2020) characterize the two cases, \(q_1 = q_2\) or \(q_1 < q_2\), in terms of the valuations of the two types, which are a primitive of the model. The characterization shows that \(q_1 = q_2\) if and only if for any alternative \(a\), type 2 has a higher ratio of valuations of alternative \(a\) to \(\bar{a}\), that is, \(r_a^1 \leq r_a^2\), where \(r_a^i = v_a^i/v_{\bar{a}}^i\). Figure 2 illustrates this inequality for the case of two alternatives. The intuition for the characterization uses the logic of screening. Making the allocation of type 1 inefficient, that is, screening, reduces the revenue of the seller from type 1 but also reduces the information rents of type 2. For this trade off to be profitable it must be, roughly, that the decrease in value for type 1 from getting a less efficient alternative is lower than the resulting decrease in type 2’s information rents. This leads to the ratio condition. We discuss this characterization, which generalizes to any number of types, and provide an economic interpretation in Section 6.

5 More Than Two Types

In this section, we generalize Proposition 1 and Proposition 2 to any number of types and any number \(k \geq 1\) of alternatives. We saw that with two types, non-screening is optimal for every market if and only if first best consumer surplus and the surplus triangle are achievable for every market. We will see that this is true with any number of types. We also saw that with two types, screening is optimal for some markets if and only if first best consumer surplus is not achievable for any inefficient market and the surplus triangle is not achievable for any non-singleton market. We will see that this is not generally true with more than two types. For first best consumer surplus to be unachievable for every inefficient market and for the surplus triangle to be unachievable for every non-singleton market, it is not enough that screening is optimal for
some markets. Instead, the set of markets for which screening is optimal must “separate” the set of markets for which non-screening is optimal, in the sense that for every market at most one non-screening mechanism is optimal. With only two types this happens whenever the set of markets for which screening is optimal is not empty.

We start with the following key result, which states that if first best consumer surplus is achievable for a given market, then non-screening must be optimal for that market. In other words, if screening is optimal for a market, then first best consumer surplus is unachievable for that market. Of course, a screening mechanism is inefficient. What the result shows is that if a market is inefficient because screening is optimal for the market, then it is impossible to achieve efficiency via segmentation without the seller appropriating some of the gains.

**Proposition 3** If first best consumer surplus is achievable for market \( f \), then there exists a type \( j \) such that mechanism \( N^j \) is optimal for \( f \).

For the proof of Proposition 3, consider a market \( f \) with an optimal mechanism \( M \), and suppose that first best consumer surplus is achievable for \( f \). We show that mechanism \( N^j \) must also be optimal for \( f \), where \( j \) is the lowest type that is not excluded in \( M \) (a type is excluded if it gets an empty allocation). Indeed, notice that any type strictly higher than \( j \) gets strictly positive utility in mechanism \( M \) since it can mimic type \( j \) and has a strictly higher value than type \( j \) for any alternative \( a \neq 0 \). Now consider a segmentation of market \( f \) that achieves first best consumer surplus, and take any segment \( f' \). By Lemma 3, \( M \) must be optimal for \( f' \). Consider the lowest type \( i(f') \) in the support of \( f' \). We must have \( i(f') \leq j \), otherwise every type in the support of \( f' \) gets strictly positive utility in \( M \), which is a contradiction to the optimality of \( M \) for market \( f' \). By Lemma 3 both \( M \) and \( N^i(f') \) must be optimal mechanisms for \( f' \). The following lemma, which is the key to Proposition 3, shows that in this case, \( N^j \) must also be optimal for \( f' \).

**Lemma 5** Consider a market \( f \) and an optimal mechanism \( M = (x, p) \), and let \( j \) be the lowest type that gets a non-empty allocation in \( M \), that is, \( j = \min \{ j' : x_0(j') < 1 \} \). Suppose that for some \( i \leq j \), \( N^i \) is also optimal for \( f \). Then, \( N^j \) is also optimal for \( f \).

**Proof.** Assume without loss of generality that \( f \) has full support on types 1 to \( n \). Assume for contradiction that mechanisms \( M \) and \( N^i \) are optimal for \( f \) but mechanism \( N^j \) is not.

---

\(^{17}\)Consider the support \( \{i_1, \ldots, i_{\ell} \} \) of \( f \). Relabel the types such that type 1 refers to \( i_1 \), type 2 refers to \( i_2 \), and so on.
construct a mechanism $M'$ that has a higher revenue than $M$. In $M'$, types below $i$ get an empty allocation (as they do in $M$). Types $i$ to $j - 1$ get alternative $\bar{a}$ with probability $\epsilon$ (and the outside option with probability $1 - \epsilon$) and pay $\epsilon v^i_{\bar{a}}$. Types $j$ to $n$ have the same allocation as in $M$, but their payment is decreased by $\epsilon (v^j_{\bar{a}} - v^i_{\bar{a}})$ relative to their payment in $M$. See Figure 7.

Mechanism $M'$ has a higher revenue than $M$ for $\epsilon > 0$. Compared to $M$, $M'$ generates an additional revenue of $\epsilon v^i_{\bar{a}}$ from all types $i' \geq i$, and loses a revenue of $\epsilon v^j_{\bar{a}}$ from all types $i' \geq j$. The difference in revenue is $\epsilon v^i_{\bar{a}} \Pr[i' \geq i] - \epsilon v^j_{\bar{a}} \Pr[i' \geq j]$, which is $\epsilon$ times the difference in the revenue of mechanism $N^i$ compared to the revenue of mechanism $N^j$. This difference is strictly positive by the assumption that $N^i$ is optimal but $N^j$ is not. It remains to show that $M'$ is IR and IC for small enough $\epsilon > 0$.

IR holds for types 1 to $i - 1$, because they are excluded in $M'$. A type $i' = i, \ldots, j - 1$ has utility $\epsilon v^i_{\bar{a}} - \epsilon v^i_{\bar{a}} \geq 0$, and a type $i' \geq j$ has a higher utility in $M'$ than in $M$. Thus, IR holds for any $\epsilon > 0$.

For IC, observe that $M'$ coincides with $M$ in the limit as $\epsilon$ goes to zero. Thus, if an IC constraint is slack (holds strictly) in $M$, then it is satisfied in $M'$ for small enough $\epsilon$. Now, in mechanism $M$ a type $i'$ strictly prefers not to mimic another type $i''$ in two cases: (1) if $i' > j$ and $i'' < j$; (2) if $i' < j$ and $i'' \geq j$. In case (1), type $i'$ has a strictly positive utility in $M$. This is because type $j$ must get utility of zero in $M$ (otherwise $M$ is not optimal as prices can increase), and type $i'$ can mimic type $j$ and get $j$'s allocation and price but has a strictly higher valuation for any alternative $a \neq 0$ by assumption. Thus $i'$ strictly prefers not mimic type $i''$ (and get utility zero) in $M$. In case (2), type $i'$ gets a strictly negative utility from mimicking $i''$ since $v^i \cdot x(i'') - p(i'') < v^j \cdot x(i'') - p(i'') \leq 0$, where the last inequality followed since the utility of type $j$ is zero.

We next verify the remaining IC constraints in mechanism $M'$. Consider a type $i' < j$. As discussed in case (2) above, such a type $i'$ does not benefit from mimicking types $i'' \geq j$. 

![Figure 7: Construction of mechanism M' from mechanism M in the proof of Lemma 5.](image-url)
Type $i'$ prefers the allocation of types $1, \ldots, i - 1$ (the outside option) to the allocation of types $i, \ldots, j - 1$ if and only if $\epsilon(v_{i'}^i - v_{i}^i) \leq 0$, that is, $i' \leq i$. Thus truthtelling maximizes the utility of a type $i' < j$. For a type $i' \geq j$, note that mimicking a type $j, \ldots, n$ is not beneficial since $M$ is IC and all such types get the same additional payment in $M'$. From case (1) above, a type $i' > j$ does not benefit from mimicking types $1, \ldots, j - 1$. Finally, the utility of type $j$ in $M'$ is at least $\epsilon(v_{j}^j - v_{i}^i) > 0$, which is the utility it would get by mimicking types $i, \ldots, j - 1$, and is no lower than the utility of zero it would get by mimicking types $1, \ldots, i - 1$. ■

To complete the proof of Proposition 3, we observe that because mechanism $N^j$ is optimal for every segment $f'$, $N^j$ is also optimal for the original market $f$. Otherwise, in market $f$ mechanism $M$, which is optimal for $f$, would give a higher revenue than mechanism $N^j$, so the same must be true for at least one segment $f'$ (for any fixed mechanism, the revenue in market $f$ is the weighted average of the revenues in the segments). But then $N^j$ would not be optimal for $f'$, a contradiction. The proof of the following result formalizes this argument.

**Lemma 6** For any mechanism $M$, the set $\mathcal{F}(M)$ is convex.

We now prove Proposition 3.

**Proof of Proposition 3.** Consider a market $f$ and a segmentation $\mu$ of $f$ that achieves first best consumer surplus. By Lemma 3, there exists a mechanism $M$ that is optimal for market $f$ and also all segments $f'$ of $\mu$. Let $j$ the the lowest type that gets a non-empty allocation in $M$, and notice that $\hat{j}(f') \leq j$ for any segment $f'$, where $\hat{j}(f')$ is the lowest type in the support of market $f'$. Indeed, if the inequality were reversed, all types in the support of $f'$ would obtain a strictly positive utility, so $M$ would not be optimal for $f'$.

By Lemma 3, for any segment $f'$ both $N^{\hat{j}(f')}$ and $M$ are optimal. Lemma 5 implies that $N^j$ is also optimal for $f'$. Since $N^j$ is optimal for every segment $f'$, Lemma 6 shows that it is also optimal for $f$. Thus, we have shown that if first best consumer surplus is achievable for a market $f$, then there exists a non-discriminating mechanism, $N^j$, that is optimal for $f$. ■

5.1 Achievability Of First Best Consumer Surplus And The Surplus Triangle

Proposition 3 implies the following result, which states that first best consumer surplus, and in fact the surplus triangle, are achievable for all markets with a given set of types if and only if non-
screening is optimal for all markets with that set of types. The second statement is equivalent to saying that the sets \( \mathcal{F}(N^i) \) collectively cover the set of all markets, that is, \( \cup_i \mathcal{F}(N^i) = \Delta(T) \).

**Proposition 4** For any set of types \( T \), the following are equivalent:

1. First best consumer surplus is achievable for every market.
2. The surplus triangle is achievable for every market.
3. For every market, \( N^i \) is optimal for some \( i \).

**Proof.** (2) \( \rightarrow \) (1): By definition.

(3) \( \rightarrow \) (2): When offering only \( \bar{a} \) is optimal for all markets, we can think of the problem as one with a single alternative \( \bar{a} \). We can therefore apply a segmentation into equal-revenue markets (Definition 5) that uses Corollary 1 and is identical to that of Bergemann et al. (2015). We defer the details to Appendix C.3.

(1) \( \rightarrow \) (3): Immediate from Proposition 3.

**Proposition 4** characterizes the achievability of first best consumer surplus and the surplus triangle for all markets in terms of whether non-screening is optimal for all markets. Whether non-screening is optimal for a given market is in general difficult to ascertain. But Haghpanah and Hartline (2020) provide a simple characterization of the sets of types for which non-screening is optimal for all markets. For the characterization, let \( r^i_a = v^i_a/v^i_{\bar{a}} \) be the ratio between type \( i \)'s valuations of alternatives \( a \) and \( \bar{a} \).

**Proposition 5** (Haghpanah and Hartline, 2020) For any set of types \( T \), the following are equivalent:

1. For every market, \( N^i \) is optimal for some \( i \).
2. The ratio \( r^i_a \) is non-decreasing in \( i \) for all \( a \).

From Proposition 4 and Proposition 5 we have the following result, which is illustrated in Figure 8.

**Theorem 1** For any set of types \( T \), the following are equivalent:

1. First best consumer surplus is achievable for every market.
Figure 8: (a) For every market, a non-screening mechanism is optimal (Statement (3) of Theorem 1). (b) The ratio of valuations increases in the valuation for the most preferred alternative (Statement (4) of Theorem 1).

2. The surplus triangle is achievable for every market.

3. For every market, $N^i$ is optimal for some $i$.

4. The ratio $r_a^i$ is non-decreasing in $i$ for all $a$.

Theorem 1 relates achievability of first best consumer surplus to the optimality of non-screening and to the ratio of valuations for different types. Let us relate the theorem to our two type analysis in Section 4. Recall that with two types, the set of markets for which mechanism $N^1$ is optimal is $[0, q_1]$, and the set of markets for which mechanism $N^2$ is optimal is $[q_2, 1]$, where $q_1 \leq q_2$. For either $N^1$ or $N^2$ to be optimal for every market, we must have $q_1 = q_2$. Thus, with two types, Theorem 1 states that first best consumer surplus and the surplus triangle are achievable for every market if and only if $q_1 = q_2$. This generalizes parts of Proposition 1 and Proposition 2. However, Proposition 1 and Proposition 2 show that if $q_1 < q_2$, then first best consumer surplus is unachievable for all inefficient markets and the surplus triangle is unachievable for all non-singleton markets. Theorem 1 does not make such a statement. The next subsection provides necessary and sufficient conditions for first best consumer surplus to be unachievable for all inefficient markets and for the surplus triangle to be unachievable for all non-singleton markets.

5.2 Unachievability Of First Best Consumer Surplus And The Surplus Triangle

We now characterize when first best consumer surplus is unachievable for all inefficient markets and the surplus triangle is unachievable for all non-singleton markets. By Theorem 1, for first
Figure 9: (a) The sets of markets for which two different mechanisms $N^i$ and $N^j$ are optimal do not intersect (Statement (3) of Theorem 2). (b) The ratio of valuations decreases in the valuation for the most preferred alternative (Statement (4) of Theorem 2).

best consumer surplus and the surplus triangle to be achievable for all markets, non-screening must be optimal for all markets. Recall that with two types (and any number of alternatives), if screening is optimal for some market, then first best consumer surplus is unachievable for all inefficient markets (Proposition 1) and the surplus triangle is unachievable for all non-singleton markets (Proposition 2). But this is not true with more than two types: it may be that screening is optimal for some markets and yet first best consumer surplus and the surplus triangle are achievable for other markets with the same set of types. To identify the condition for unachievability, let us interpret the third statement of Theorem 1 as stating that the sets of markets for which non-screening is optimal collectively include all markets. As we show below, first best consumer surplus and the surplus triangle are unachievable if the sets of markets for which non-screening is optimal do not intersect, that is, the set of markets for which screening is optimal separates the sets of markets for which non-screening is optimal. This is illustrated in Figure 9.

**Theorem 2** For any set of types $T$, the following are equivalent:

1. First best consumer surplus is unachievable for every inefficient market.
2. The surplus triangle is unachievable for every non-singleton market.
3. For every market, $N^i$ is optimal for at most one $i$.
4. For every pair of types $i < j$, there exists an alternative $a$ such that $r^i_a > r^j_a$.

The proof uses the following key lemma, which states that if there is a pair of types $i < j$ such that $r^i_a > r^j_a$ for some $a$, then there exists no market for which both mechanisms $N^i$ and
$N^i$ and $N^j$ are optimal. Recall from our two type analysis in Section 4 that if $r_a^i > r_a^j$, then $N^i$ and $N^j$ are not both optimal for any market with support $\{i, j\}$. The lemma generalizes this statement to all markets, regardless of their supports. In contrast to our two type analysis, the proof of the lemma does not rely on identifying optimal mechanisms. Instead, assuming both $N^i$ and $N^j$ are optimal, the lemma constructs a mechanism that outperforms $N^j$.

**Lemma 7** Consider a pair of types $i < j$ such that $r_a^i > r_a^j$ for some $a$. Then, for any market $f$, mechanisms $N^i$ and $N^j$ cannot both be optimal.

We defer the complete proof of Theorem 2 to Appendix C.4. To develop a partial intuition, suppose that the third statement of Theorem 2 holds, namely, for every market, $N^i$ is optimal for at most one $i$. If this is the case, then there exists no equal-revenue market (Definition 5) other than markets that consist of only a single type of consumers (singleton markets). As a result, Corollary 1 cannot be applied since it concerns only segmentations in which all segments are equal-revenue markets. This, of course, is insufficient to prove that no segmentation achieves first best consumer surplus. For this, Lemma 7 is instrumental.

Theorem 1 and Theorem 2 generalize our two type analysis (Proposition 1). Theorem 1 for the two type case states that first best consumer surplus and the surplus triangle are achievable for all inefficient markets if and only if $q_1 = q_2$ (Figure 6). Theorem 2 for the two type case states that first best consumer surplus is unachievable for any inefficient market and the surplus triangle is unachievable for any non-singleton market if and only if $q_1 < q_2$.

Thus, with two types, either non-screening is optimal for all markets ($q_1 = q_2$), or the sets of markets for which non-screening is optimal do not intersect ($q_1 < q_2$). This is shown in Figure 6. With more than two types, this is no longer the case. It is possible that screening is optimal for some markets, and yet two sets of markets over which non-screening is optimal intersect, as illustrated in Figure 10. The next subsection addresses this remaining case.

### 5.3 The Remaining Case

The following result is an immediate corollary of Theorem 1 and Theorem 2. The first statement in the result is true if and only if the first statements of both Theorem 1 and Theorem 2 are false, and similarly for the other statements.

**Theorem 3** For any set of types $T$, the following are equivalent:
Figure 10: (a) Screening is optimal for some markets, but the screening region does not separate non-screening regions (Statement (3) of Theorem 3). (b) The ratio of valuations decreases in the valuation for the most preferred alternative for some but not all pairs of types (Statement (4) of Theorem 3).

1. First best consumer surplus is achievable for some but not all inefficient markets.

2. The surplus triangle is achievable for some but not all non-singleton markets.

3. There exists a market for which screening is optimal, that is, $N^i$ is not optimal for any $i$, and there exists a market for which $N^i$ and $N^j$ are both optimal for some $i \neq j$.

4. There exists a pair of types $i < j$ such that $r_{a}^{i} \leq r_{a}^{j}$ for all $a$, and there exists a pair of types $i' < j'$ such that $r_{a}^{i'} > r_{a}^{j'}$ for some $a$.

In Appendix C.5, we discuss an example of markets for which first best consumer surplus is achievable when conditions (2) and (3) from Theorem 3 hold.

6 Economic Interpretation

By Theorem 1, when the ratio between the value of each alternative and value of the best alternative is higher for higher types, consumers in any market can achieve first best consumer surplus by providing certain information to the seller, who then uses this information to segment the market and offer a profit-maximizing menu in each segment. Under the same conditions, a social planner can achieve any point in the surplus triangle. The reason is that under these conditions on the ratios, screening is not profitable in any market, so the seller offers only the best alternative in every market segment and we can use the results of Bergemann et al. (2015) for markets with a single product. Intuitively, screening is not optimal for any market when the ratios are ranked because selling a less desirable alternative to some type reduces the revenue of
the seller from this type by more than the resulting decrease in information rents from higher
types. In contrast, Theorem 2 shows that if the ratio between the value of each alternative and
value of the best alternative is lower for higher types, then the seller optimally screens types in a
set of markets that separates the sets of markets in which screening is not optimal. In this case,
consumers cannot provide information to achieve first best consumer surplus in any inefficient
market, and a planner cannot achieve the surplus triangle in any inefficient market. The seller’s
ability to screen consumers when screening is profitable interferes with the achievability of first
best consumer surplus and the surplus triangle.

When are the ratios likely to be increasing? The higher the ratio is for a given alternative, the
lower the relative value is of replacing that alternative with the best alternative. Thus, ratios are
increasing in type when individuals who have higher absolute values have lower relative values to
replacing alternatives with the best alternative. As a simple example consider a movie streaming
service and a newly-released movie for which the service can offer a rental option and a purchase
option. Suppose there are two types of consumers. The first type does not like movies that
much, but prefers to purchase the movies he watches and rewatch parts of them from time to
time. The second type likes to watch movies, but does not much like to watch the same movie
twice. This type’s willingness to pay for renting or buying the movie is higher than the first
type’s, but he is willing to pay relatively less to buy the movie (compared to renting it) than the
first type. Then, regardless of the proportions of the two types in the market, it is optimal for
the streaming service to offer only the purchase option, and first best consumer surplus and the
surplus triangle are achievable. If, on the other hand, the type that likes movies more also likes
to rewatch movies and the other type does not, then screening is optimal for some markets, so
first best consumer surplus is not achievable for any inefficient market and the surplus triangle
is not achievable for any non-singleton market.

In this example, it is the pleasure they derive from watching movies that makes one type
willing to pay more than the other type. A similar ranking in types’ willingness to pay can also
arise because of wealth differences: a wealthier person, whose marginal utility for a dollar is lower
than that of a less wealthy person, may be willing to pay more for any alternative. The ranking of
ratios can also be interpreted easily in a bundling setting when the best alternative corresponds
to the grand bundle. The ratio $r^i_a$ can be thought of as the inverse relative value of “topping up”
bundle $a$ by adding the remaining products: the higher $r^i_a$ is, the lower the relative willingness to
pay for these additional products. Haghpanah and Hartline (2020) provide additional examples
and discussion of the ratio conditions.

7 Conclusion

We studied first best consumer surplus and the surplus triangle in a multi-product environment. A key feature of our model is that the seller may find it profitable to screen consumers in a market by offering multiple bundle-price pairs, thus combining second- and third-degree price discrimination. With two consumer types, we provided a complete characterization of when first best consumer surplus and the surplus triangle are achievable. With more than two types, we provided a characterization of when first best consumer surplus is achievable for all markets or no inefficient market and when the surplus triangle is achievable for all markets or no non-singleton market. These characterizations relate achievability to the optimality of screening. We then provided an economic interpretation for these characterizations using the ratios of the values of alternatives to the value of the best alternative.

References


A Appendix for Section 3

A.1 Proof of Lemma 1

Proof of Lemma 1. We argued in the text that $CS(f) \leq MCS(f)$. To see that $MCS(f) \leq E_{i \sim f}[v_i^a] - ER(f)$, consider any market $f'$ and its optimal mechanism $M' = (x', p')$ such that $CS(f') = EU(f', M')$. We have

$$CS(f') + ER(f') = E_{i \sim f'}[v^i \cdot x'(i) - p'(i)] + E_{i \sim f'}[p'(i)] = E_{i \sim f'}[v^i \cdot x'(i)] \leq E_{i \sim f'}[v_i^a].$$  \hspace{1cm} (3)

Now consider a market $f$ and a segmentation $\mu$ of $f$. From (3) we have

$$CS(\mu) + E_{f' \sim \mu}[ER(f')] = E_{f' \sim \mu}[CS(f') + ER(f')] \leq E_{f' \sim \mu}[E_{i \sim f'}[v_i^a]] = E_{i \sim f'}[v_i^a],$$  \hspace{1cm} (4)

where the equality followed since $E_{f' \sim \mu}[f'] = f$. Now consider any optimal mechanism $(x, p)$ of $f$. We have

$$E_{f' \sim \mu}[ER(f')] \geq E_{f' \sim \mu}[\sum_i f'_i p(i)] = \sum f_i p(i) = ER(f).$$  \hspace{1cm} (5)
Combining (4) and (5), we conclude that
\[ CS(\mu) \leq E_{i \sim f}[v^i_a] - ER(f) \]
for all \( \mu \), and therefore \( MCS(f) \leq E_{i \sim f}[v^i_a] - ER(f) \). ■

A.2 Proof of Lemma 3

Proof of Lemma 3. Consider any market \( f' \) and its optimal mechanism \( M' = (x', p') \) such that \( CS(f') = EU(f', M') \). We have
\[ CS(f') + ER(f') = E_{i \sim f'}[v^i \cdot x'(i) - p'(i)] + E_{i \sim f'}[p'(i)] = E_{i \sim f'}[v^i \cdot x'(i)] \leq E_{i \sim f'}[v^i_a], \] (6)
Now consider a market \( f \) and a segmentation \( \mu \) of \( f \). From (6) we have
\[ CS(\mu) + E_{f' \sim \mu}[ER(f')] = E_{f' \sim \mu}[CS(f') + ER(f')] \leq E_{f' \sim \mu}[E_{i \sim f'}[v^i_a]] = E_{i \sim f'}[v^i_a], \] (7)
where the equality followed since \( E_{f' \sim \mu}[f'] = f \). Now consider any optimal mechanism \( (x, p) \) of \( f \). We have
\[ E_{f' \sim \mu}[ER(f')] \geq E_{f' \sim \mu}\sum_i f'_i p(i) = \sum_i f_i p(i) = ER(f) \]. (8)
Combining (7) and (8), we conclude that
\[ CS(\mu) \leq E_{i \sim f}[v^i_a] - ER(f), \]
with equality if and only if both (7) and (8) hold with equality. For (7) to hold with equality, it is necessary and sufficient that \( x_a(i) = 1 \) for all \( i \) in the support of every segment of \( \mu \). That is, every segment of \( \mu \) is efficient (Definition 2). For (8) to hold with equality, it is sufficient that some optimal mechanism \( (x, p) \) of \( f \) is optimal for every segment of \( \mu \), and is necessary that all optimal mechanisms \( (x, p) \) of \( f \) is optimal for every segment of \( \mu \). ■

B Appendix for Section 4

B.1 Optimal Mechanisms For Two Types

The following lemma provides a characterization of optimal mechanisms for two types. A mechanism is optimal if and only if the following conditions hold. First, the allocation of type 1 is a
distribution over alternatives that maximize $v^1_a - qv^2_a$, and the allocation of type 2 is alternative $\bar{a}$. Second, the payments are calculated via the binding IR constraint of type 1 and the binding IC constraint of type 2.

**Lemma 8** With two types, a mechanism $(x, p)$ is optimal if and only if

1. $x(1) \in \Delta(\arg\max_a v^1_a - qv^2_a)$, $x_{\bar{a}}(2) = 1$.
2. $p(1) = v^1 \cdot x(1)$, $p(2) = p(1) + v^2 \cdot (x(2) - x(1))$.

**Proof.** Consider any mechanism $(x', p')$. The IR constraint for type 1 is

$$v^1 \cdot x'(1) - p'(1) \geq 0, \quad (9)$$

Similarly, the IC constraint for type 2 is

$$v^2 \cdot x'(2) - p'(2) \geq v^2 \cdot x'(1) - p'(1), \quad (10)$$

Substituting (9) and (10), the expected revenue is

$$(1 - q)p'(1) + qp'(2) = p'(1) + q(p'(2) - p'(1))$$

$$\leq v^1 \cdot x^2 \cdot (x'(2) - x'(1))]$$

$$= x'(1) \cdot (v^1 - qv^2) + qx'(2) \cdot v^2 \quad (11)$$

Thus, maximum revenue across all mechanisms is at most the maximum of (11) across all $x'(1)$ and $x'(2)$, which is obtained by setting $x(1) \in \Delta(\arg\max_a v^1_a - qv^2_a)$ and $x_{\bar{a}}(2) = 1$. Furthermore, note that the expected revenue of mechanism $(x, p)$ equals $x(1) \cdot (v^1 - qv^2) + qx(2) \cdot v^2$ if the two constraints (9) and (10) hold with equality, which is obtained by setting $p(1) = v^1 \cdot x(1)$ and $p(2) = p(1) + v^2 \cdot (x(2) - x(1))$. Thus, if mechanism $(x, p)$ is IC and IR, it is optimal. Furthermore, if such a mechanism is IC and IR, then all optimal mechanisms must satisfy the conditions specified by the lemma.

To verify that the mechanism $(x, p)$ is IC and IR, note that (9) and (10) imply that the IR constraint for type 2 is satisfied. The only remaining constraint is the incentive constraint for
type 1. The utility of type 1 for deviating to type 2 is
\[ v^1 \cdot x(2) - p(2) = v^1 \cdot x(2) - p(1) - v^2 \cdot (x(2) - x(1)) \]
\[ = v^1 \cdot x(2) - v^1 \cdot x(1) - v^2 \cdot (x(2) - x(1)) \]
\[ = (v^2 - v^1) \cdot x(1) - (v^2 - v^1) \cdot x(2) \]
\[ = \sum_a x_a(1)((v^2_a - v^1_a) - (v^2 - v^1) \cdot x(2)) \tag{12} \]
where the last equality followed since \( \sum_a x_a(1) = 1 \). Now consider any alternative \( a \) such that \( x_a(1) > 0 \). By definition, it must be that \( a \in \arg \max_a v^1_a - qv^2_a \). Thus, in particular, \( v^1_a - qv^2_a \geq v^1_a - qv^2_a \). Since \( q \leq 1 \), we have \( v^2_a - v^2_a \geq v^1_a - v^1_a \). Therefore, for any \( a \) where \( x_a(1) > 0 \),
\[ (v^2_a - v^1_a) - (v^2 - v^1) \cdot x(2) \leq 0. \]
Now note that for any \( a \), either \( x_a(1) = 0 \), or \( x_a(1) > 0 \) in which case the above inequality holds. In either case, we have
\[ x_a(1)((v^2_a - v^1_a) - (v^2 - v^1) \cdot x(2)) \leq 0. \]
Summing over all \( a \), and given (12), we have
\[ v^1 \cdot x(2) - p(2) = \sum_a x_a(1)((v^2_a - v^1_a) - (v^2 - v^1) \cdot x(2)) \leq 0. \]
Thus, the incentive constraint for type 1 is satisfied.

The following lemma reveals the structural properties of \( \arg \max_a v^1_a - qv^2_a \), which identifies the allocation of type 1 by Lemma 8. The problem is one of maximizing over functions \( v^1_a - qv^2_a \) that are linear in \( q \). Thus, the set of markets \([0, 1]\) can be divided into intervals, each assigned a unique alternative, such that the alternative is the unique maximizer of \( v^1_a - qv^2_a \) in the interior of the interval, and is an optimizer of \( v^1_a - qv^2_a \) at the endpoints of the interval. Additionally, \( \max_a v^1_a - qv^2_a \) is convex in \( q \). See Figure 11.

**Lemma 9** For some \( m \leq k \), there exist thresholds \( \tau_0 < \ldots < \tau_{m+1} \), \( \tau_0 = 0 \), \( \tau_{m+1} = 1 \) and an injective function \( g : \{0, \ldots, m\} \to A \), \( g(0) = \bar{a} \), \( g(m) = 0 \), such that for all \( j \leq m \), \( g(j) \in \arg \max_a v^1_a - qv^2_a \) if and only if \( q = [\tau_j, \tau_{j+1}] \). Additionally, \( v^2_{g(j)} - v^1_{g(j)} \) is strictly decreasing in \( j \).
Figure 11: An example with $k = 4$, $m = 3$, $\tau_1 = a$, and $\tau_2 = a'$.

**Proof.** Consider the problem of maximizing $v_a^1 - qv_a^2$. For each $q$, the problem is to maximize over $k$ linear functions. The graph is drawn in Figure 11. Note the properties of the solution. First, the set of markets $[0, 1]$ can be divided into intervals with thresholds $\tau_0 < \ldots < \tau_{m+1}$, with each $\tau_j$ assigned a unique alternative $g(j)$, such that for all $j \leq m$, $g(j) \in \arg \max_a v_a^1 - qv_a^2$ if and only if $q = [\tau_j, \tau_{j+1}]$. Since $v_a^1 > v_{\bar{a}}^1$ for all $a \neq \bar{a}$, $g(0) = \bar{a}$. Similarly, since $v_a^1 - v_{\bar{a}}^1 < 0$ for all $a \neq 0, g(m) = 0$. Additionally, consider $j$ such that $g_{j-1} = a$ and $g_j = a'$. We must have $v_a^1 - qv_a^2 > v_{\bar{a}}^1 - qv_{\bar{a}}^2$ for all $q < \tau_j$ and $v_a^1 - qv_a^2 < v_{\bar{a}}^1 - qv_{\bar{a}}^2$ for all $q > \tau_j$. Thus, $v_a^2 - v_{\bar{a}}^2 > v_{\bar{a}}^2 - v_a^1$.

**B.2 Proof of Lemma 4**

**Proof of Lemma 4.** By Lemma 8, the mechanism $N^1$ is optimal if and only if $\bar{a} \in \arg \max_a v_a^1 - qv_a^2$. By Lemma 9, the set of such markets is $[0, q_1]$ for $q_1 = \tau_1 > 0$. Furthermore, $N^1$ is the unique optimal mechanism for $[0, q_1)$.

Similarly, by Lemma 8 the mechanism $N^2$ is optimal if and only if $0 \in \arg \max_a v_a^1 - qv_a^2$. By Lemma 9, the set of such markets is $[q_2, 1]$ for $q_2 = \tau_m < 1$. Furthermore, $N^2$ is the unique optimal mechanism for $(q_2, 1]$.

Finally, consider a mechanism $M$ that is optimal for $(q_1, q_2)$. Since neither $N^1$ nor $N^2$ is optimal for any market $(q_1, q_2)$, $M \neq N^1, N^2$, and thus $M$ cannot be optimal for $[0, q_1)$ or $(q_2, 1]$. 

35
B.3 Proof of Proposition 2

Proof of Proposition 2. Suppose that \( q_1 = q_2 \). As noted by Bergemann et al. (2015), the same segmentation that achieves first best consumer surplus also achieves the surplus triangle.

Now suppose that \( q_1 < q_2 \). Consider any inefficient market \( q > q_1 \). By Proposition 1, first best consumer surplus, and therefore the surplus triangle, is unachievable for such a market. Now consider an efficient market \( q \leq q_1 \). In this case the lowest possible total surplus is unachievable. That is, there exists no segmentation such that the consumer surplus is zero and the revenue is equal to that of \( q \).

C Appendix for Section 5

C.1 Proof of Lemma 6

Proof of Lemma 6. Consider two markets \( f', f'' \) for which a mechanism \((x', p')\) is optimal. Let \( f = \alpha f' + (1 - \alpha) f'' \). For any mechanism \((x, p)\), we have

\[
E_{i \sim f}[p(i)] = \alpha E_{i \sim f'}[p(i)] + (1 - \alpha) E_{i \sim f''}[p(i)] \leq \alpha E_{i \sim f'}[p'(i)] + (1 - \alpha) E_{i \sim f''}[p'(i)] = E_{i \sim f}[p'(i)].
\]

That is, for market \( f \), the revenue of any mechanism \((x, p)\) is at most the revenue of \((x', p')\). Thus, \((x', p')\) is optimal for \( f \).

C.2 Proof of Lemma 7

Proof of Lemma 7. Assume for contradiction that \( r^i_a > r^j_a \) for some \( i < j \) and \( a \), and \( N^i \) and \( N^j \) are both optimal for a market \( f \). Denote by \( q_i \) the fraction of types \( i \) and higher in market \( f \), and by \( q_j \) the fraction of types \( j \) and higher in market \( f \). For \( N^i \) and \( N^j \) to be both optimal, we must have \( v^i_a q_i = v^j_a q_j \), that is \( q_i = \frac{v^j_a q_j}{v^i_a} \). Thus we can write

\[
v^i_a q_i = v^i_a (\frac{v^j_a q_j}{v^i_a}) = (\frac{v^j_a v^i_a}{v^i_a}) q_j > v^j_a q_j. \tag{13}
\]

where the inequality followed from the assumption that \( r^i_a > r^j_a \) (that is, \( v^i_a / v^i_a > v^j_a / v^j_a \)).

Construct a mechanism \( M \) that improves upon \( N^j \) as follows. Types \( i, \ldots, j - 1 \) get alternative \( a \) with probability \( \epsilon \) and pay \( \epsilon v^i_a \). Types \( j, \ldots, n \) get alternative \( \bar{a} \) and pay \( v^j_{\bar{a}} - \epsilon (v^j_{\bar{a}} - v^i_a) \). This is illustrated in Figure 12.
Let us compare the revenue of $M$ with the revenue of $N^j$. Types $i, \ldots, j - 1$ pay $\epsilon v^i_a$ more in $M$ than in $N^j$. Types $j$ and higher pay $\epsilon(v^j_a - v^i_a)$ less in $M$ than in $N^j$. The difference in expected revenue is

$$
\epsilon v^i_a (q_i - q_j) - \epsilon(v^j_a - v^i_a)q_j = \epsilon(v^i_a q_i - v^j_a q_j) > 0,
$$

where the inequality followed from inequality 13. So to complete the proof, we show that $M$ is IC and IR, which contradicts the assumption that $N^j$ is optimal.

Mechanism $M$ is IR. Types lower than $i$ are excluded. A type $i'$ from $i$ to $j - 1$ has utility $\epsilon(v^i_a - v^i_a) \geq 0$. Types $j$ and higher have a higher utility in $M$ than in $N^j$.

For IC, observe similarly to the proof of Lemma 5 that if an incentive constraint is slack in $N^j$, then it is satisfied in $M$ for small enough $\epsilon > 0$. In particular, (1) a type $i' > j$ does not benefit from mimicking a type $i'' < j$, (2) a type $i' < j$ does not benefit from mimicking a type $i'' \geq j$.

We now verify the remaining incentive constraints. A type $i' < j$ prefers the allocation of types $i, \ldots, j - 1$ to the outside option if and only if $\epsilon(v^i_a - v^i_a) \geq 0$, that is, $i' \geq i$. Thus the incentive constraints are satisfied for types $i' < j$. For types $i' \geq j$, note that mimicking any type $j, \ldots, n$ is not beneficial since all such types have the same allocation and payment. Finally, the utility of type $j$ in $M$ is $\epsilon(v^j_a - v^i_a)$, which is the utility it would receive by mimicking types $i, \ldots, j - 1$, and is strictly more than the utility it would receive by mimicking types $1, \ldots, i - 1$.

C.3 Proof of Proposition 4

We start with the following lemma from Bergemann et al. (2015).

**Lemma 10** For any market $f$ and $i \in \arg \max_j v^j_a (\sum_{i' \geq j} f_{i'})$, there exists a segmentation $\mu$ of $f$ such that every segment $f'$ of $\mu$ includes $i$ in its support and for all $i'$ in the support of $f'$,
\[ i' \in \arg \max_j v^j_a(\sum_{i' \geq j} f_{i'}). \]

We now complete the proof of Proposition 4.

**Proof of Proposition 4** To complete the proof, we show that (2) implies (1). Consider a market \( f \). By assumption, the mechanism \( N^i \) is optimal for some \( i \). In particular, \( N^i \) is optimal for any \( i \in \arg \max_j v^j_a(\sum_{i' \geq j} f_{i'}). \) By Lemma 10, there exists a segmentation \( \mu \) such that every segment \( f' \) of \( \mu \) includes \( i \) in its support and for all \( i' \) in the support of \( f' \), \( i' \in \arg \max_j v^j_a(\sum_{i' \geq j} f_{i'}). \) Since \( N^j \) is optimal for \( f' \) for some \( j \), \( N^i \) is optimal for all \( i' \) in the support of \( f' \). That is, \( f' \) is an equal-revenue market. Thus, by Corollary 1, \( \mu \) achieves first best consumer surplus.

**C.4 Proof of Theorem 2**

**Proof of Theorem 2.**

(3) \( \rightarrow \) (1) and (2): To see that (3) implies (1), suppose for contradiction that (1) is violated, that is, some segmentation \( \mu \) of an inefficient market \( f \) achieves first best consumer surplus. From Proposition 3, \( N^i \) must be optimal for market \( f \) for some \( i \). By Lemma 3, \( N^i \) is optimal for every segment \( f' \) of \( \mu \). The lowest type \( \bar{i}(f') \) in the support of \( f' \) must satisfy \( \bar{i}(f') \leq i \), otherwise all types in the support of \( f' \) get strictly positive utility in the optimal mechanism \( N^i \). Moreover, at least one segment \( f' \) must satisfy \( \bar{i}(f') < i \). Otherwise, if \( \bar{i}(f') = i \) for all segments \( f' \), then \( \bar{i}(f) = i \) and \( f \) is efficient. Now consider a segment \( f' \) such that \( \bar{i}(f') < i \). By Lemma 3 for market \( f' \), \( N^{\bar{i}(f')} \) is optimal. That is, for \( \bar{i}(f') < i \), \( N^i \) and \( N^{\bar{i}(f')} \) are both optimal for some market \( f' \). Therefore, (3) is violated. We have thus shown that (3) implies (1). That (3) implies (1) also shows that if (3) holds, then the surplus triangle is not achievable for any inefficient market.

To see that (3) implies (2), it remains to show that the surplus triangle is not achievable for any non-singleton efficient market. Consider a non-singleton efficient market \( f \) and suppose that a segmentation \( \mu \) achieves the lowest possible total surplus (Definition 4). Consider a segment \( f' \) whose support includes type \( \bar{i}(f) \), the highest type in the support of \( f \). Because \( \mu \) achieves the lowest possible total surplus, consumer surplus in \( f' \) is 0, so \( N^{\bar{i}(f)} \) is optimal for \( f' \). And since \( f \) is efficient, \( N^{\bar{i}(f)} \) is optimal for \( f \), where \( \bar{i}(f) \) is the lowest type in the support of \( f \). So by Lemma 3, \( N^{\bar{i}(f)} \) is also optimal for \( f' \). Thus, both \( N^{\bar{i}(f)} \) and \( N^{\bar{i}(f)} \) are optimal for \( f' \), and these are distinct mechanisms because \( f \) is a non-singleton market. This contradicts (3).
Figure 13: The set of inefficient markets for which first best consumer surplus is achievable is shaded.

(4) → (3): Directly from Lemma 7.

(1) → (4) and (2) → (4): Suppose for contradiction that for some $i < j$, $r^i_a \leq r^j_a$ for all $a$. By Proposition 5, either $N^i$ or $N^j$ is optimal for any market with support in $\{i, j\}$. By Proposition 1 and Proposition 2, first best consumer surplus and the surplus triangle are achievable for every market with support in $\{i, j\}$.

C.5 Theorem 3: An Example

Theorem 3 identifies conditions under which first best consumer surplus is achievable for some, but not all, inefficient markets. In this section we identify with an example the set of markets for which first best consumer surplus is achievable.

Consider an example with 3 types and its set of optimal mechanisms as shown in Figure 13. The set of inefficient markets for which first best consumer surplus is achievable is shaded in Figure 13. Markets in $\hat{F}(N^1)$ are efficient. First best consumer surplus is unachievable for any market for which screening is optimal by Proposition 3. Thus, if first best consumer surplus is achievable for an inefficient market, then the market must be in $\hat{F}(N^2) \cup \hat{F}(N^3)$. Suppose that a segmentation $\mu$ of a market $f$ in $\hat{F}(N^2)$ achieves first best consumer surplus. By Lemma 3, two conditions must hold for every segment $f'$ of $\mu$. First, $f'$ must be in $\hat{F}(N^2)$. Second, $N^i(f')$ must be optimal for $f'$. The set of such markets $f'$ is shown in green in Figure 13. Market $f$ must be in the convex hull of such market, which is shaded in Figure 13. Similarly, if first best consumer surplus is achievable for a market $f \in \hat{F}(N^3)$, then $f$ must be in the convex hull of the set of markets that are shown in red in Figure 13.