An Analysis of Portion Cap Rules with a Multi-Product Seller

Jose G. Nuno-Ledesma [1]

January 2021

Link to most recent version

Abstract

I study the impacts of limiting the quantity of one product in a two-goods market with privately-informed buyers. The goal is to explore effects on consumer surplus and consumption of both products. The main finding is that following moderate enough caps, a standard nonlinear pricing model predicts an increase in surplus for the buyer with low preference for the regulated good and high valuation for the unregulated product. Specific changes in allocation depend on the model's parameters. Consumption of the limited good is reduced, while consumption of the unregulated item often falls. Changes in screening are also interesting. Severe enough caps cause the seller to move from full separation to bunching some types. These results are applicable to food retail and other industries, where portion cap rules are a regulatory alternative and the impact on consumer surplus is an important consideration.

PRELIMINARY VERSION. PLEASE, DO NOT CITE.

Keywords: Adverse selection; nonlinear pricing; portion cap; quantity restriction; multiple products.

JEL classification: D82, I18, L51

[1] Assistant professor, University of Guelph. `jnuno@uoguelph.ca`. I benefited from comments by Steven Wu; Joseph Balagtas; Timothy Cason; Bhagyashree Katare; Guanming Shi; Gerald Shively; Alfons Weersink, and by participants in seminars held at the Center for Research and Teaching in Economics (CIDE), Purdue University, and Indiana University. All errors are my own. Financial support from USDA-NIFA grant award no. 2014-67023-21852 is gratefully acknowledged.

1 Introduction

Portion cap rules (caps) are policies restricting the default maximum quantities at which products can be offered. In light of studies linking larger portions to increased consumption, foods and ingredients judged to have deleterious impacts on human health are natural targets of potential caps (Flood et al. 2006; Rolls et al. 2006; Ledikwe et al. 2005; Young and Nestle 2002). One example is the so-called New York City (NYC) "soda ban". The advanced plan intended to prohibit vendors regulated by NYC from selling sugar-sweetened beverages in containers exceeding 16 ounces (Kansagra, 2012).[1] Ultimately, the measure was struck down in court. Nonetheless, the proposal sparked vibrant public discussions around its potential ills and virtues. Most debates departed from the premise that caps unavoidably hurt consumers.[2] I argue that even if a regulation reduces consumption of the target product, it does not necessarily follow that buyers are worse-off. This is because sellers implement sophisticated pricing strategies and are likely to modify them to accommodate the intervention, with outcomes difficult to anticipate. In this paper, I use a standard nonlinear pricing model to show that, when the seller offers two products, capping one of the goods *benefits* the consumer who is more interested in the unregulated product but not as much in the regulated good.

In the case of food portion caps similar to the discussed "soda ban", the analysis in this paper would be more directly applicable in settings where products are sold in combos or "value meals", such as in quick-service restaurants. Caps as presented in this document are also analogous to other policies in the field, including credence labeling, and limits on

[1] As a reference, the "small", "medium", and "large" cup sizes typically found in popular American fast-food restaurants contain around 16, 21, and 32 ounces each.

[2] Notes in the media where this premise is present to some degree abound. These are two examples: Grynbaum (2012); and Grynbaum and Connelly (2012).

ingredients per serving. I discuss two more examples for which the present analysis can be informative to varying degrees.

First, consider credence labeling. Region of origin, organic status, and fairness of trade, among other attributes, are greatly appreciated by some consumers, but not by others. This leads to consumer heterogeneity as to the differentiating trait. Typically, a producer must satisfy rules regarding the amount or proportion of ingredients according to a government-defined standard to feature the label on their product.[3] This is akin to quantity cap on ingredients that do not follow the requirements. The work in this paper can illuminate the effects of such mandates if sellers desire to exhibit the label in their products and were to strictly follow the requisites.

Second, some stand-alone food products can be thought of as bundles.[4] For example, consider a sugary beverage manufacturer deciding sugar-"rest of ingredients" (call them goods A and B) combinations. The umbrella component called "rest of ingredients" includes ingredients such as fruit extract, artificial flavoring substances, carbon dioxide gas, and ingredients other than sugar. The "package" or bundle is a bottle of soda with a particular sugar-"rest of ingredients" ratio. To the degree that buyers have preferences over sweetness due to sugar (good A) on the one hand, and flavor, gas content or other characteristics (good B) on the other, sellers tailor alternatives to screen consumers.[5] The same logic can

[3]Take the Canadian example. Only food products with 95% or more content satisfying their organic standards may carry the Canadian Organic Label. Similarly, a food producer may show the label "Product of Canada" on their goods when, at most, 2% of the ingredients are non-Canadian.

[4]This interpretation coincides with the rationality behind random utility models (widely used in the literature leveraging discrete choice experiments), where total utility derived from consuming a product equals the sum of independent utilities provided by each of the attributes it embodies (Louviere and Woodworth 1983; Lancaster 1966). The idea of treating single products as bundles is present in some of the first papers in the screening literature (Adams and Yellen, 1976)

[5]Think for example of the presence of both "classic" and "diet" versions of the same soda brand, and how "classic" versions have different amount of sugar in different markets sometimes (it has been reported that before January 2015, a 355 millilitres can of regular Coke in Canada had 42 grams of sugar, but 39 grams

be applied to other products (e.g. flavored milk contains sweeteners, flavoring agents, and milk; breakfast cereals are composed of cereal grains, sugar, and sometimes raisins, nuts and other add-ons). Regulations designed to reduce the amount of a specific ingredient present in food servings can be informed by the model in this paper. One example is the United Kingdom's sugar reduction program. In 2016, the government of the U.K. set out to reduce overall sugar by 20% (taking 2015 levels as the benchmark) in a range of products widely consumed by children (Public Health England, 2017). The means to achieve the reduction in sugar content include both smaller portion sizes, and product reformulation to reduce sugar levels in a given portion. Although the program started as largely voluntary, my analysis can be informative to the degree that price-discriminating food producers decide to strictly comply with the program by reformulating their products.

I present an analysis of the effects of caps in a two-product market with four types of privately-informed buyers. Supposing that one of the items is subject to a portion cap, I aim to learn i) whether the intervention affects the consumption of both the target and the unregulated goods, ii) the effect on consumer surplus defined as gross utility from consumption net of the price paid, and iii) the effect on the seller's segmentation strategy. The model incorporates three stylized observations. First, preferences are private information to the buyers and are taken as exogenous by the seller. It is fair to assume that in most markets taste can be considered exogenous and that sellers design incentive-compatible menus before any transaction occurs. Second, the seller offers more than one product. This reflects what is observed in the field, where most retailers are either multi-product vendors or offer one product which itself is made of more than one component. Lastly, the seller decides the

in the U.S. (Schwartz, 2015)).

quantities and prices that characterize each package on the menu. In other words, she does not adopt a passive pricing scheme. Following a restriction in quantities, there is no reason to assume that the seller will not try to endogenously modify the menu to accommodate the intervention in ways that would impact how gains from trade are divided.

The multidimensional character of the model is an important feature. It highlights how changes in consumer surplus and allocation are contingent on the way a compliant seller adapts her pricing scheme while maintaining incentive compatibility. Moreover, by considering a two-product market, I can show that one customer type benefits from the limit due to a peculiar interaction between the regulation and the seller's bundling strategy:[6] without intervention, the customer is offered a portion of the target good larger than his first best; the cap moves the portion closer to his ideal. This paper complements previous work focusing on the one-dimensional case, where caps cannot increase buyers' surplus because the bundling device is absent.

2 Relevant literature

The design of incentive-compatible menus by sellers aiming to segment demand is a well-documented phenomenon. The single-product case has received more attention than the multidimensional scenario.[7] Perhaps because multidimensional nonlinear pricing is notorious

[6]Commodity bundling is the screening mechanism wherein the price of a package containing various items in combination is lower than the sum of the prices for the stand-alone products. Alternatively, and more relevant to the model in this paper, if two goods are always sold together in combos containing both components (the scheme known as pure bundling), they are said to be bundled is the variance in price across different packages is not entirely explained by differences in marginal cost of production. I provide a formal definition that applies to my analysis in the section where I introduce the model.

[7]See Myerson (1979), Maskin and Riley (1984), Mussa and Rosen (1978), Armstrong and Vickers (1993), and Sonderegger (2011)

for being a source of research problems easy to state but difficult to solve analytically or otherwise, the literature on the topic has remained relatively small.[8] With this paper, I expand the set of works studying the behavior of multi-product sellers.

Additionally, I contribute to the corpus of studies researching the effects of regulating price-discriminating firms. The regulatory measures analyzed using frameworks related in spirit to mine include minimum quality standards (Saitone and Sexton 2010; Besanko et al. 1988); food product quality and certification (Crespi and Marette 2003; Marette and Crespi 2003), and maximum prices (Corts 1995; Besanko et al. 1988). This paper looks at maximum quantity caps, a relatively less studied intervention.

My work directly contributes to the strand of literature examining the effects of quantity caps leveraging nonlinear pricing theory. In particular, my work complements findings by Bourquard and Wu (2019). They take the NYC "soda ban" as a motivating example. In their paper, one retailer offers a single product to two discrete buyer types with private taste. The H-type is willing to pay more for the product, while the L-type does not value it as much. Without regulation, the seller offers two options, one large and another small, to serve each buyer separately. The H-type retains a positive surplus. The L-type earns the reservation value. If the quantity limit distorts only the large portion, the H-type consumes less, but consumer surplus does not change.

The insight of Bourquard and Wu (2019) is that, with one product and discrete types, the lowest type is not affected by the cap, while higher types are not impacted if the limit does not distort the portion immediately below that of their choice. With their framework, a cap

[8]For discussions about of the literature on design of multi-product pricing and bundling and its evolution, see Armstrong (1996), and Armstrong (2016)

cannot make any buyer better-off. My paper explores the degree to which these conclusions extrapolate to the bi-dimensional case. The only screening device available to the one-product seller is a price schedule, concave with respect to quantity, to provide self-selection incentives to higher types. With more than one product, bundling becomes accessible to the seller as a tool to segment demand. Bourquard and Wu do not explore the outcomes resulting from the interaction between a possible bundling scheme and a reduce quantity space due to regulation. I examine this interaction.

My study also speaks to the empirical literature documenting how buyers respond to portion caps (Ahn and Lusk 2020; John et al. 2017; Wilson et al. 2013). For example, Wilson et al. (2013) conduct a behavioral study to determine how a limit on sugary drink portions might affect consumption patterns. They ask human subjects to indicate how much product they would hypothetically consume when presented two types of menus. The regulated menu replaces large (32oz) drinks with two smaller cups (e.g. 16oz each), leaving other options the same. They find that customers end up buying more ounces of soda when presented with the regulated menu. In essence, they document a potential framing effect resulting from the policy. My paper complements this literature by examining the seller's theoretical response to the policy (as opposed to that of the consumer). A complete explanation of the consequences of an intervention ought to include analyses of reactions from both buyers and sellers.

3 Model

The model is largely based on Armstrong and Rochet (1999). The seller (she) is a monopolist offering goods A and B in contracts $\{q^A,\ q^B,\ p\}$, where p is the price charged for a package containing q^A and q^B units of the respective products. The monopoly assumption does not entirely drive the results I delineate below. In nonlinear pricing models, the need to screen buyers emerges because the seller does not have complete information. Introducing competition would reduce consumption distortions for lower types, but would not eliminate this adverse selection problem (Stole, 2007).[9]

The ij-type buyer (he) has private preference i for good A, and j for B. For each item, he can either have a high (H) or a low (L) preference. There are four buyer types: HH, HL, LH, and LL. The ij-type buyer is characterized by the vector of taste parameters $(\theta_i^A,\ \theta_j^B)$ for $i,\ j =$ H, L. I assume $\theta_H > \theta_L$. I do not require $\theta_H^A \equiv \theta_H^B$ or $\theta_L^A \equiv \theta_L^B$. If the ij-type buyer pays price p_{ij} for a package containing quantities q_{ij}^A and q_{ij}^B, he earns consumer surplus:

$$R_{ij} = \theta_i^A u(q_{ij}^A) + \theta_j^B u(q_{ij}^B) - p_{ij} \tag{1}$$

The subscripts i and j under R, q^A, q^B, and p indicate the buyer type. I assume away interactions between the components. This is to highlight the interactions between the multidimensional incentive constraints and the seller's pricing behavior across policy scenarios. This assumption provides a neutral background where changes across regulatory environments can confidently be attributed to the impact of quantity restrictions on pricing behavior without the confounding effects of complementarity. Thus, I show that a cap

[9]Further, most second-degree price discrimination examples occur in oligopolisitc markets (Tirole, 1988). Thus the use of this pricing strategy does not imply either complete absence or prevalence of competition.

changes allocation and increases consumer surplus for a type *even in the absence of complementarity.*[10]

I do not include health benefits in the consumer welfare analysis that would be relevant in the case of regulations in the food retailing industry. This modeling decision implies the assumption that health benefits from caps are null. I keep this assumption for three main reasons. First, it would be difficult to argue a precise measure of consumer welfare improvement due to the cap. Attempting to do so, would require me to adopt arbitrary assumptions regarding how precisely reduced consumption translates into consumer welfare gains. Second, the set of assumptions could be strategically chosen to generate any desired outcome. Lastly, by ignoring welfare effects from reduced consumption, I am likely underestimating the gains for buyers. The most salient result of the paper is that a subset of consumers is benefited by the cap *even if there are no additional health benefits from lowered consumption of the presumed unhealthy food.*

I assume the utility function $u(\cdot)$ to be continuous, also $u(0) = 0$, $u'(q) > 0$ and $u''(q) < 0$. The seller has reservation value of zero. I assume both goods to have the same differentiable, increasing and convex cost function $c(\cdot)$. Also, $\theta_H u'(0) > c'(0)$ and $\lim_{q\to\infty}\theta_H u'(q) < c'(q)$ for both products, so that trade is possible at least with the HH-type, and total quantity supplied is finite. Hereafter, to simplify notation throughout the paper, I make the following substitution: $q' \equiv u(q)$.

The probability that a given buyer is an ij-type is denoted by β_{ij}, so $\sum_{ij}\beta_{ij} = 1$. The seller ignores the particular type of any particular consumer, but knows the distribution of

[10]I explore the case with related components (i.e. when goods are complements or substitutes in consumption) in a note available upon reasonable request. Introducing complementarity does not alter the essence of the findings regarding surplus distribution.

preferences. Expected profit is $\mathbb{E}[\pi] = \sum_{ij} \beta_{ij} \Big[p_{ij} - c(q_{ij}^A) - c(q_{ij}^B) \Big]$. It is useful to represent expected profit in terms of total and consumer surpluses:

$$\mathbb{E}[\pi] = \underbrace{\sum_{ij} \beta_{ij}[\theta_i^A q_{ij}'^A + \theta_j^B q_{ij}'^B - c(q_{ij}^A) - c(q_{ij}^B)]}_{\text{Expected total surplus}} - \underbrace{\sum_{ij} \beta_{ij}[\theta_i^A q_{ij}'^A + \theta_j^B q_{ij}'^B - p_{ij}]}_{\text{Expected consumer surplus}} \tag{2}$$

The seller considers one set of participation constraints (PC), and another of incentive-compatibility restrictions (IC). The satisfaction of PC implies that all types are at least indifferent between buying and opting-out from trade while earning a reservation value $\bar{u}$. Their general form is $R_{ij} \geq \bar{u} \; \forall \; ij$.

The IC restrictions are self-selection conditions providing incentives for the ij-type buyer not to purchase a package originally intended to serve the kl-type (either $i \neq k$ or $j \neq l$, or both inequalities hold). In other words, at the optimum, quantities and prices are such that the ij-type buyer is weakly better-off by choosing contract $\{q_{ij}^A, q_{ij}^B, p_{ij}\}$ over an alternative $\{q_{kl}^A, q_{kl}^B, p_{kl}\}$. The seller designs a menu of options such that the ij-type receives a positive surplus in the form of a temptation payoff. This is exactly equal to the extraordinary benefit the ij-type would have gained had he chosen the alternative intended for the kl-type. The IC constraints take the following general form:

$$R_{ij} \geq R_{kl} + q_{kl}'^A(\theta_i^A - \theta_k^A) + q_{kl}'^B(\theta_j^B - \theta_l^B) + \bar{u} \;\; \forall \; ij \text{ and } kl; i \neq k \text{ and } j \neq l \tag{3}$$

The seller designs options $\{q_{ij}^A, q_{ij}^B, p_{ij}\}$ to maximize expected profit (2) subject to PC and IC constraints. The resulting pricing mechanism is incentive-compatible if it satisfies monotonicity conditions: $q_{HH}^A \geq q_{LH}^A$, $q_{HL}^A \geq q_{LL}^A$, $q_{HH}^B \geq q_{HL}^B$, and $q_{LH}^B \geq q_{LL}^B$. These mean that the quantity of either good weakly increases with the corresponding valuation.

The complete optimization program includes 4 PC and 12 IC restrictions. In a "relaxed" version of the profit maximization problem, only a subset of the constraints is included. In this version of the program, only the "downward" incentive restrictions are relevant.[11] Similarly, as long as the PC restriction for the LL-type is satisfied, the rest of PCs are also met. In the appendix, I show that the solution to the simplified problem is the solution to the fully constrained program. The relevant IC constraints are graphically depicted in panel (a) of figure 1, where solid lines denote equations binding with equality, and dashed lines indicate restrictions that may or may not bind with strict equality depending on the problem's parameters.

[Figure 1 about here]

Before discussing the baseline results stemming from the maximization problem described above, I dedicate a few paragraphs to the phenomenon of commodity bundling.

3.1 Commodity bundling

If in the resulting menu of alternatives, the quantity of good i increases with the preference for component j, the seller is said to have bundled the products.

***Definition** 1. In this model, the seller is said to implement **bundling** when, for a given menu of contracts, the quantity of product i increases with preference for product j, i.e. when*

[11] In other words, the seller does not consider the possibility of the HL-type choosing the packages intended for either LH or LL buyers; or the LL-type choosing any option other than his; or the LH buyer choosing the alternative designed to serve the HL and/or the HH-types.

$q^A_{LL} < q^A_{LH}$, *and/or* $q^A_{HL} < q^A_{HH}$, *and/or* $q^B_{LL} < q^B_{HL}$, *and/or* $q^B_{LH} < q^B_{HH}$.

The seller can adopt a version of pure bundling: the items are always offered together.[12] The reader can tell if bundling is present by comparing consumption by types. An indication would be if, for instance, the HL-type purchases more good B than the LL-type (notice the only difference between these types is their preference for A, not B). From the point of view of the HL-type, to get more of the good he highly desires (A), he has to purchase more of the product he does nor value as much (B). This holds for the other types.

The distribution of types is described by the following correlation: $\rho = \beta_{HH}\beta_{LL} - \beta_{HL}\beta_{LH}$. One of the main intuitions in the early screening literature is that it is in the seller's best interest to bundle whenever this correlation is weak enough (Armstrong and Rochet 1999; McAfee et al. 1989; Adams and Yellen 1976). In this model, bundling is profitable as long as $\rho < \frac{\beta_{HL}\beta_{LH}}{\beta_{LL}}$. I will assume that $\rho < 0 < \frac{\beta_{HL}\beta_{LH}}{\beta_{LL}}$, which is when the incentive is strongest. When these inequalities hold, "medium types" (HL and LH) are relatively more common. It pays off to separate them from the "extreme types" (HH and LL). The way to screen is bundling: to a given medium type, offer more of the good he mildly prefers, and the increased production cost will be counterbalanced by increased sales.[13]

I study a seller that engages in bundling because even casual observation suggests that this practice is rampant in the food retail industry. Moreover, the model replicates an empirical

[12] Another well known bundling scheme is mixed bundling, where products can either buy the products by themselves or in combos. Under this strategy, the price of the combo is less than than the sum of the prices of individual goods.

[13] On the other hand, if $\frac{\beta_{HL}\beta_{LH}}{\beta_{LL}} < \rho$, the extreme types are proportionally more common and it is not worth it to separate the scarce medium types. Under these circumstances, the allocation and surplus distribution would be just as if the seller had solved, for each good separately, a single-product nonlinear pricing problem with two buyer types. The results from Bourquard and Wu (2019) would apply.

regularity documented in the marketing literature: with bundling, buyers consume more of items they would normally either not purchase or get small quantities of them (Sharpe and Staelin, 2010). Thus, findings derived from this model are useful in a regulatory context.

3.2 Baseline outcomes and incentive compatibility structures

Critically, the surplus R_{ij} received by each buyer type depends on the combination of active incentive constraints. Each set of binding IC restrictions corresponds to a particular incentive compatibility structure (IC-structure or ICS). Without a cap, there are four possible IC-structures (call them ICS Γ, Υ, Ψ, and Ω). Under all of them, the surplus earned by the LL, LH, and HL types are the following:

$$\begin{aligned} R_{HL} &= \Delta^A q_{LL}'^A + \bar{u} \\ R_{LH} &= \Delta^B q_{LL}'^B + \bar{u} \\ R_{LL} &= \bar{u} \end{aligned} \tag{4}$$

Where $\Delta^g \equiv \theta_H^g - \theta_L^g$ for $g = A, B$. Regarding the specific form of the HH-type's IC constraint, it varies across ICS in the following manner:

$$\begin{aligned} &\text{ICS } \Gamma\text{: } R_{HH} = \Delta^A q_{LH}'^A + \Delta^B q_{HL}'^B + \bar{u} \\ &\text{ICS } \Upsilon\text{: } R_{HH} = \Delta^A q_{LH}'^A + \Delta^B q_{LL}'^B + \bar{u} \\ &\text{ICS } \Psi\text{: } R_{HH} = \Delta^A q_{LL}'^A + \Delta^B q_{HL}'^B + \bar{u} \\ &\text{ICS } \Omega\text{: } R_{HH} = \Delta^A q_{LL}'^A + \Delta^B q_{LL}'^B + \bar{u} \end{aligned} \tag{5}$$

Although each ICS implies solutions with different point predictions regarding allocation and consumer surplus, some patterns are consistent. Without regulation, consumption weakly increases with type; quantity of product i increases when the preference for good j rises; and consumer rents weakly increase with preferences, with the LL-type earning the reservation value, the HH-type gaining the largest surplus, and the medium HL and LH-types earning rents in between. Naturally, the effects of a quantity limit are contingent on the IC-structure characterizing the regulation-free market. I will pay attention to this detail when describing the effects of the regulation.

Both panels (a) of figures 2 and 3 are helpful illustrations of the resulting allocation and surplus distribution, correspondingly. Throughout the paper, I discuss outcomes from the symmetric case ($\theta_H^A \equiv \theta_H^B$, $\theta_L^A \equiv \theta_L^B$, and $\beta_{HL} = \beta_{LH}$). However, the essence and logic of the result is the same for non-symmetric cases I omit scale labels along the vertical axis of both figures because the specific values of these variables depend on the chosen parameters. For some parameter combinations, the LL-type is excluded from participation, and surplus earned by the LL, LH, and HL types equal the reservation value $\bar{u}$. The horizontal lines in panel (a) figure 3 mark the levels of surplus received by the HH and LH types, to make the comparison with the cap scenario easier.

[Figures 2 and 3 about here]

To offer an interpretation of the baseline outcomes, consider the "combo meal" reading of the model. Figure 2 guides the discussion. Products A and B would stand for, say, soda and fries. The goods are offered in "large", "medium", and "small" portions. The HH-type is offered a combo with large options of both goods; the HL-type buys a meal with a large soda and medium fries; the LH gets a medium soda and large fries, and the LL buyer purchases small portions of both. I now proceed to describe the outcomes that come with the imposition of a quantity cap.

4 Model with portion cap rule

Without loss of generality, suppose that the quantity limit applies to good A. With the enacted regulation, bundles cannot have quantities of A larger than $\bar{q}$. The seller's objective is to maximize expected profit subject to the same participation and incentive-compatibility constraints, plus the following quantity cap (QC) restriction:

$$\text{QC:} \;\; q_{ij}^{A} \leq \bar{q} \;\text{ for } i,\, j = L,\, H \tag{6}$$

I analyze three levels of severity at which the cap can be set. The limit can be 1) mild if the limit is set below the largest quantity available without regulation but above the quantity of the second-largest unregulated alternative; 2) moderate if the cap is set below the second-largest unregulated option of good A, but also over the quantity contained in the smallest regulation-free alternative, or 3) severe if the limit on quantity is set at a level lower than the small alternative without the cap. I present the first order conditions characterizing the solution for each level of severity in the online appendix.

It is important to pay attention to how the caps distort the set of binding IC constraints in the expected profit maximization program. These are shown in panel (b) of figure 1.

***Proposition** 1.* *A mild cap does not modify the set of binding constraints regardless of the original regulation-free IC-structure. With moderate and severe caps, the set of IC constraints* *to consider in the seller's problem includes:* $\{LH \to LL;\ LH \to HL;\ HL \to LL\}$*, each of them binding with equality.*

Proofs for this and later propositions are in the appendix. Panel (b) in figure 1 shows the set of active incentive constraints under moderate and severe restrictions. It is important to notice that a mild cap does not modify the set of binding constraints, and the seller offers as much of A as she is allowed for the HH and HL buyers. The moderate restriction lies below the size offered to the LH type but above the option serving the lowest type in the regulation-free baseline. The moderate and severe limits modify the set of binding IC constraints. Under moderate and severe regulations, the size of product A shrinks so much that the seller no longer separates the HH and LH customers. Therefore, the seller defaults to offer a single option to serve both of these buyers. This "bunching" of types occurs regardless of the original regulation-free ICS. As the quantity of product A becomes smaller due to more restrictive caps, the seller has to be mindful of the possibility of the LH buyer misrepresenting himself as an HL-type and increases the temptation payoff accordingly. With a moderate or severe cap, the incentive constraint preventing unfaithful representation of the LH as an HL-type is binding. This can be confirmed using equation 3. This modification renders the downward incentive constraints involving the HH-type redundant, thus it can be ignored.

The effects of the caps on the quantities of both products are summarized in Table 1, and listed in the proposition below. The proofs result from direct comparisons between the corresponding first order conditions, thus I omit them. The first order conditions are shown in the online appendix.

[Table 1 about here]

***Proposition** 2. With independent goods and regardless of the original ICS, a mild cap affects only the target product and the portion served to the HH-type diminishes. The effects on quantities consumed associated with moderate and severe caps are the following:*

- *Moderate and severe caps reduce the quantity of A consumed by all buyer types.*
- *When the regulation-base IC structures are either ICS Γ or ICS Ψ, types HL and LL consume less of product B. For the remaining ICS, only the HL-type buyer reduces his consumption of B.*

There are two salient effects on the quantities of the regulated good. First, a mild cap affects exclusively the largest options. Second, a moderate restriction not only directly reduces the large and medium options, as intended, but also indirectly impacts the portion of A in the small alternative. Because interactions between the products are absent, it is surprising to find that the model suggests changes in the quantity of the *unregulated* good B following moderate and severe caps. When passing from regulation-free ICS Γ or ICS Ψ to a moderate or severe cap, not only the corresponding quantities of good A are reduced, but q^B_{HL} and q^B_{LL} are as well. When passing from regulation-free ICS Υ or Ω to a moderate or severe cap, q^B_{HL} drops. To facilitate the interpretation of the results, panel (b) in figure 2 shows the allocation resulting from a moderate cap in the symmetric case. The horizontal line denotes the hypothetical quantity cap.

Again, the “combo meal” interpretation can help in the understanding of the outcomes. Take panels (a) and (b) of figure 2. Now the seller offers soda in two sizes: large (with the same pre-intervention quantity) and medium (smaller than the pre-cap scenario). French fries continue being offered in three portions, large (with unchanged quantity), medium (lower than without regulation), and small (even smaller than the baseline). The HH and the LH types buy the same combo with medium soda and large fries. The HL consumer purchases the combo with medium soda and fries. The LL-type consumes the meal with small portions of both products.

I continue with a discussion about the effect on consumer surplus. Only moderate and more restrictive caps modify consumer rents because a mild cap does not distort any downward incentive compatibility constraint.

***Proposition** 3. When products are unrelated, only moderate and severe caps modify consumer surplus.*

***Proposition** 4. When products are unrelated, moderate and severe caps have the following effect on the rents earned by the LL, LH, and HL types:*

- *R_{LL} is unaffected.*
- *R_{HL} is negatively affected by both moderate and severe cap.*
- *R_{LH} increases following either a moderate or a severe cap.*

Additionally, the LH and HH types are bunched together and served the same alternative.

The LH-type buyer is benefited by moderate and severe caps on A. As shown by the equations in 4, the LH-type earns $\Delta^B q'^B_{LL}$ without regulation (and with a mild cap) regardless of the original ICS. Let wide tildes $(\widetilde{\cdot})$ denote solutions under a moderate restriction, and wide hats $(\widehat{\cdot})$ denote solutions under severe caps. With moderate and severe caps, the specific functional form of the LH-type rents change in the following manner:

$$\begin{aligned} &\textbf{Moderate cap: } \widetilde{R_{LH}} = \Delta^A[\widetilde{q'^A_{LL}} - \bar{q}'] + \Delta^B \widetilde{q'^B_{LL}} + \Delta^B \widetilde{q'^B_{HL}} + \bar{u} \\ &\textbf{Severe cap: } \widehat{R_{LH}} = \Delta^B \widehat{q'^B_{LL}} + \Delta^B \widehat{q'^B_{HL}} + \bar{u} \end{aligned} \tag{7}$$

Notice that $\widehat{R_{LH}} > \widetilde{R_{LH}}$ because $\Delta^A[\widetilde{q'^A_{LL}} - \bar{q}'] < 0$. In short, compared to the regulation-free baseline, R_{LH} increases with moderate and severe caps. Panel (b) in figure 3 is helpful illustration. The horizontal lines mark the levels where the HH and LH-type rents were before the enactment of a moderate cap. Compared to panel (a) of figure 3, surplus earned by the LH-type is larger. This is the most salient result in the paper. Unlike in the one-dimensional case where buyers are unaffected at best, with two products a customer type benefits from the imposition of a cap.

The intuition is the following. Compared to the first-best quantities for the LH-type, the distortions introduced by the regulation-free separating strategy results in a larger portion of product A (soda), and the ideal quantity of B (fries). In the unregulated baseline, the LH-type is purchasing a "medium" portion A for which he has a *low* preference. This buyer would rather buy a "small-large" A-B package, but the closest option in the unregulated baseline is a price-discounted "medium-large" combo (the "small-small" alternative has too little of B, whereas the "large-large" package is just too expensive for this buyer). A quantity limit on good A shapes the set of alternatives and moves them closer to this buyer's ideal scenario.

On the other hand, the HH-type is negatively impacted by the intervention. Recall that without regulation, the HH-type's rents can take different values depending on the original

IC structure. Under moderate and severe caps, this is no longer the case. With these quantity limits, the specific form of R_{HH} is the same regardless of the original regulation-free IC-structure. The surplus earned by the HH-type are negatively affected as long as the regulation-free rents from the corresponding ICS in 5 are strictly less than the following rents under regulation:

$$\begin{aligned} &\text{Moderate cap} : \widetilde{R_{HH}} = \Delta^A \widetilde{q_{LL}^{\prime A}} + \Delta^B \widetilde{q_{LL}^{\prime B}} + \Delta^B \widetilde{q_{HL}^{\prime B}} + \bar{u} \\ &\text{Severe cap} : \widehat{R_{HH}} = \Delta^A \bar{q}' + \Delta^B \widehat{q_{LL}^{\prime B}} + \Delta^B \widehat{q_{HL}^{\prime B}} + \bar{u} \end{aligned} \tag{8}$$

A brief discussion on the mechanism driving the model's outcomes may be beneficial. Once the cap is implemented, the seller's desire to price-discriminate continues. The restriction merely reduces her choice space. To accommodate the policy while continuing to segment the demand, the seller has to modify all of the endogenous variables at her disposal, including quantities of product B and the segmentation pattern (i.e. whether to bunch or separate buyer types).

According to the model, the HL-type buyer is offered less of B when the cap is moderate or severe. The LL buyer might receive less of B depending on the original ICS. I first discuss the adjustments made to the small package. In essence, these are driven by the LL-type's participation constraint and the need to provide incentives to LH to purchase his own package. Without regulation, surplus for the LH-type is driven by a larger quantity of A, compared to the level received by the LL-type buyer. With a moderate cap, the LH-type consumes less of the regulated good (the same occurs with the HL and HH buyers). However, the profit-maximizing seller still needs to provide positive earnings to the LH-type to make sure that he will not purchase the small combo designed to serve LL. Because there is a limit on A, the only way the seller can increase the difference in quantity of A offered to the LL and LH types is by decreasing the quantity of A served to the LL buyer. Thus, the LL receives less product A. To maintain LL-type's participation constraint, the seller modifies the quantity of B served to this type.

I now turn to discuss the changes in the package sold to the HL-buyer. These are explained by modifications in the smallest package (served to the LL-type), and the fact that the need to separate the HL from the LL-type remains, albeit the incentives need not be as strong under regulation. Due to the cap, the seller is unable to offer the first best quantity of A to the HL buyer. Indeed, the HL-type purchases considerably less compared to the baseline. The seller still needs to provide incentives to the HL-type in the form of a larger portion of B compared to the LL package. Because the quantity of A contained in the smallest package (that serving the LL-type) is low and indeed smaller compared to the baseline unregulated case, the extra amount of B granted to the HL-type consumer to generate surplus need not be as large.

5 Conclusion

I present a theoretical analysis of the effects of portion caps on consumer surplus, consumption, and segmentation scheme when the seller offers two products to serve privately-informed buyers. Moderate and severe caps do reduce consumption of the target good by all buyer types. Also, depending on the original IC-structure, the consumption of the non-target good B can decrease as well. Additionally, moderate and severe limits cause the seller to bunch together some buyer types that previously were offered tailored alternatives.

Quantity limits increase consumer surplus for a subset of buyers. The benefited buyer has low valuation for the regulated product but high preference for the unregulated good. Absent a quantity limit, the seller has an incentive to engage in commodity bundling and offers a relatively larger quantity of the product this buyer values lowly. The cap reduces the extent to which bundling can be leveraged as a sorting device. The offered alternative under a cap is closer to the customer's first-best option.

The analysis I present in this paper is based on a partial equilibrium model with no market failures. This limits the degree to which the conclusions can be applied to normative questions about whether government intervention is granted based on social welfare. As one of the first studies on the effects of quantity caps in multi-product markets, this paper serves as a starting point for assessing conclusions that, without careful analysis, seem plausible and therefore influence policy design. One example is the intuitive idea that caps necessarily hurt consumers. This type of analysis is important because the success of a policy hinges on the ability of proponents to stress benefits over costs, when supporting normative arguments.

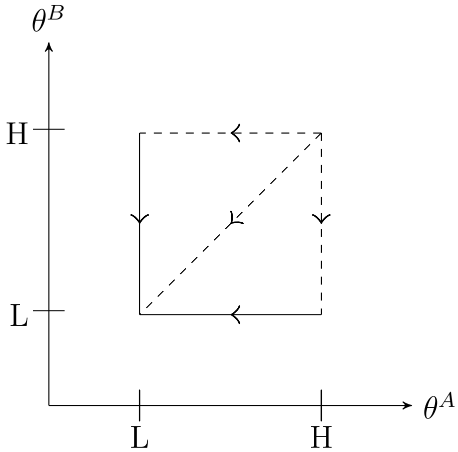
Other researchers can expand the analysis in this paper in multiple forms. For example, by including behavioral effects, studying caps with a general equilibrium framework, or by conducting empirical research (either observational or experimental) to evaluate the theoretical predictions. Health economists and researchers in adjacent areas can conduct studies that incorporate health considerations to the analysis. Future work can produce formal comparisons between the impacts of portion cap rules and other popular regulatory measures such as per-unit taxes. Additionally, both theoretical and empirical work on the effects of caps when the seller can practice mixed bundling are natural extensions.

Tables and figures

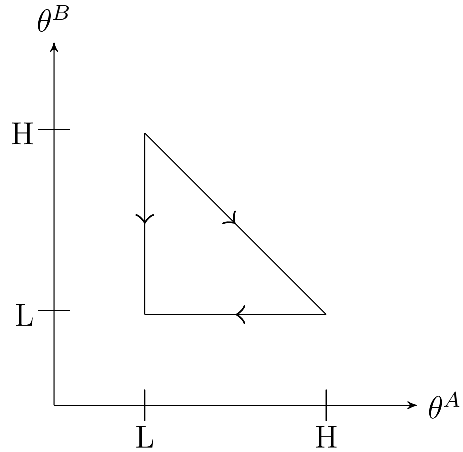
Table 1: Change in quantities

	q_{HH}^A	q_{HH}^B	q_{HL}^A	q_{HL}^B	q_{LH}^A	q_{LH}^B	q_{LL}^A	q_{LL}^B
IC-Structure Γ								
Mild	$\downarrow$	$=$	$\downarrow$	$=$	$=$	$=$	$=$	$=$
Moderate	$\downarrow$	$=$	$\downarrow$	$\downarrow$	$\downarrow$	$=$	$\downarrow$	$\downarrow$
Severe	$\downarrow$	$=$	$\downarrow$	$\downarrow$	$\downarrow$	$=$	$\downarrow$	$\downarrow$
IC-Structure Υ								
Mild	$\downarrow$	$=$	$\downarrow$	$=$	$=$	$=$	$=$	$=$
Moderate	$\downarrow$	$=$	$\downarrow$	$\downarrow$	$\downarrow$	$=$	$\downarrow$	$=$
Severe	$\downarrow$	$=$	$\downarrow$	$\downarrow$	$\downarrow$	$=$	$\downarrow$	$=$
IC-Structure Ψ								
Mild	$\downarrow$	$=$	$\downarrow$	$=$	$=$	$=$	$=$	$=$
Moderate	$\downarrow$	$=$	$\downarrow$	$\downarrow$	$\downarrow$	$=$	$\downarrow$	$\downarrow$
Severe	$\downarrow$	$=$	$\downarrow$	$\downarrow$	$\downarrow$	$=$	$\downarrow$	$\downarrow$
IC-Structure Ω								
Mild	$\downarrow$	$=$	$\downarrow$	$=$	$=$	$=$	$=$	$=$
Moderate	$\downarrow$	$=$	$\downarrow$	$\downarrow$	$\downarrow$	$=$	$\downarrow$	$=$
Severe	$\downarrow$	$=$	$\downarrow$	$\downarrow$	$\downarrow$	$=$	$\downarrow$	$=$

In each case, the comparison is against the baseline scenario.

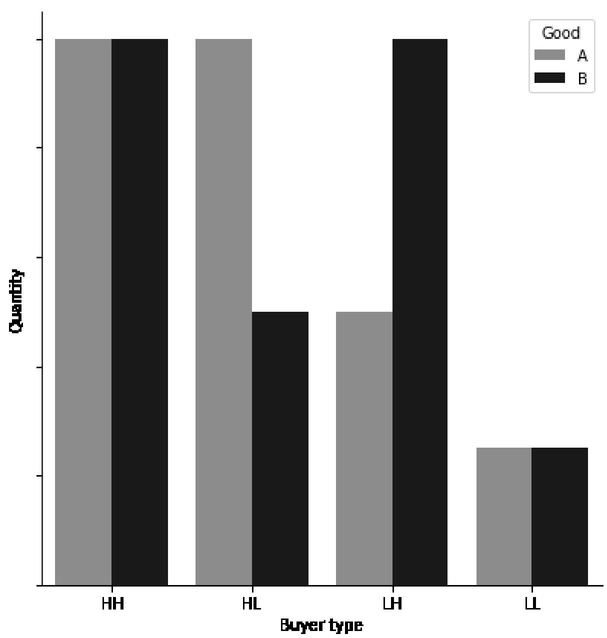


(a) Baseline

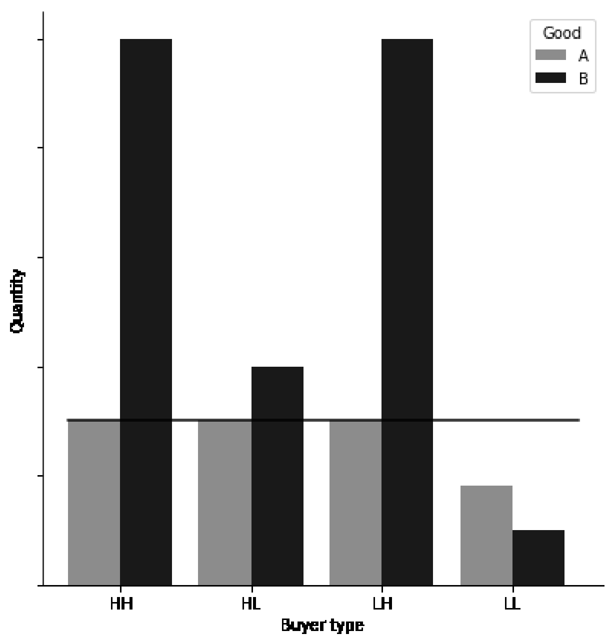


(b) Cap (moderate and severe)

Figure 1: Set of IC constraints

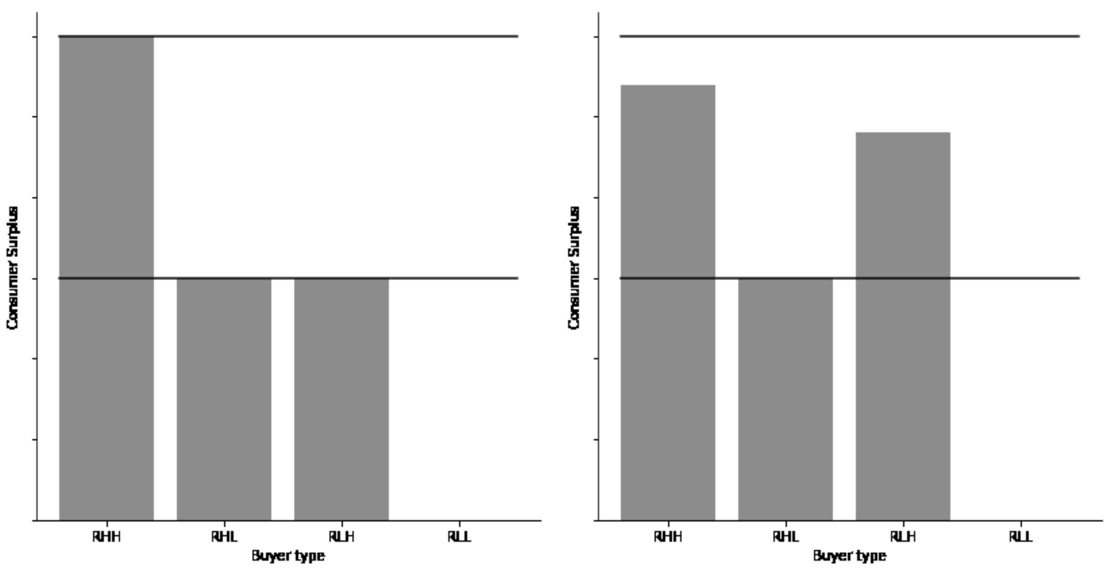


(a) Baseline



(b) Cap

Figure 2: Allocation by Buyer Type



(a) Baseline

(b) Cap

Figure 3: Consumer Surplus by Buyer Type

References

Adams, W. J. and Yellen, J. L. (1976). Commodity Bundling and the Burden of Monopoly. *The Quarterly Journal of Economics*, 90(3):475.

Ahn, S. and Lusk, J. L. (2020). Non-Pecuniary Effects of Sugar-Sweetened Beverage Policies. *American Journal of Agricultural Economics*, 103(1):53–69.

Armstrong, M. (1996). Multiproduct Nonlinear Pricing. *Econometrica*, 64(1):51.

Armstrong, M. (2016). Nonlinear Pricing. *Annual Review of Economics*, 8(1):583–614.

Armstrong, M. and Rochet, J.-C. (1999). Multi-dimensional Screening: A User's Guide. *European Economic Review*, 43(4-6):959–979.

Armstrong, M. and Vickers, J. (1993). Price Discrimination, Competition and Regulation. *The Journal of Industrial Economics*, 41(4):335.

Besanko, D., Donnenfeld, S., and White, L. J. (1988). The Multiproduct Firm, Quality Choice, and Regulation. *The Journal of Industrial Economics*, 36(4):411.

Bourquard, B. A. and Wu, S. Y. (2019). An Analysis of Beverage Size Restrictions. *American Journal of Agricultural Economics*, 102(1):169–185.

Corts, K. S. (1995). Regulation of a Multi-Product Monopolist: Effects on Pricing and Bundling. *The Journal of Industrial Economics*, 43(4):377.

Crespi, J. M. and Marette, S. (2003). "Does contain" vs."does not contain": Does it matter which GMO label is used? *European Journal of Law and Economics*, 16(3):327–344.

Flood, J. E., Roe, L. S., and Rolls, B. J. (2006). The effect of increased beverage portion size on energy intake at a meal. *Journal of the American Dietetic Association*, 106(12):1984–1990.

Grynbaum, M. M. (2012). Soda Makers Begin Their Push Against New York Ban - The New York Times. Retrieved from http://www.nytimes.com.

Grynbaum, M. M. and Connelly, M. (2012). 60 in City Oppose Bloomberg's Soda Ban, Poll Finds. Retrieved from http://www.nytimes.com.

John, L. K., Donnelly, G. E., and Roberto, C. A. (2017). Psychologically informed implementations of sugary-drink portion limits. *Psychological science*, 28(5):620–629.

Kansagra, S. (2012). Maximum size for sugary drinks: Proposed amendment of article 81. *New York: Bureau of Chronic Disease Prevention and Tobacco Control: New York Department of Health and Mental Hygiene*.

Lancaster, K. J. (1966). A New Approach to Consumer Theory. *Journal of Political Economy*, 74(2):132–157.

Ledikwe, J. H., Ello-Martin, J. A., and Rolls, B. J. (2005). Portion Sizes and the Obesity Epidemic. *The Journal of Nutrition*, 135(4):905–909.

Louviere, J. J. and Woodworth, G. (1983). Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data. *Journal of Marketing Research*, 20(4):350–367.

Marette, S. and Crespi, J. M. (2003). Can Quality Certification Lead to Stable Cartels? *Review of Industrial Organization*, 23(1):43–64.

Maskin, E. and Riley, J. (1984). Monopoly with Incomplete Information. *The RAND Journal of Economics*, 15(2):171.

McAfee, R. P., McMillan, J., and Whinston, M. D. (1989). Multiproduct Monopoly, Commodity Bundling, and Correlation of Values. *The Quarterly Journal of Economics*, 104(2):371.

Mussa, M. and Rosen, S. (1978). Monopoly and Product Quality. *Journal of Economic Theory*, 18(2):301–317.

Myerson, R. B. (1979). Incentive Compatibility and the Bargaining Problem. *Econometrica*, 47(1):61.

Public Health England (2017). Sugar Reduction: Achieving the 20%. A technical report outlining progress to date, guidelines for industry, 2015 baseline levels in key foods and next steps. Technical report.

Rolls, B. J., Roe, L. S., and Meengs, J. S. (2006). Larger Portion Sizes Lead to a Sustained Increase in Energy Intake Over 2 Days. *Journal of the American Dietetic Association*, 106(4):543–549.

Saitone, T. L. and Sexton, R. J. (2010). Impacts of minimum quality standards imposed through marketing orders or related producer organizations. *American Journal of Agricultural Economics*, 92(1):164–180.

Schwartz, D. (2015). Why Coke is lowering its sugar levels in Canada. Retrieved from http://www.cbc.ca.

Sharpe, K. M. and Staelin, R. (2010). Consumption effects of bundling: Consumer perceptions, firm actions, and public policy implications. *Journal of Public Policy & Marketing*, 29(2):170–188.

Sonderegger, S. (2011). Market Segmentation with Nonlinear Pricing. *The Journal of Industrial Economics*, 59(1):38–62.

Stole, L. A. (2007). Chapter 34 “Price Discrimination and Competition”. In *Handbook of Industrial Organization*, pages 2221–2299. Elsevier.

Tirole, J. (1988). *The Theory of Industrial Organization*. MIT press.

Wilson, B. M., Stolarz-Fantino, S., and Fantino, E. (2013). Regulating the way to obesity: Unintended consequences of limiting sugary drink sizes. *PLoS ONE*, 8(4):e61081.

Young, L. R. and Nestle, M. (2002). The Contribution of Expanding Portion Sizes to the US Obesity Epidemic. *American Journal of Public Health*, 92(2):246–249.

Appendix A: Proofs for online publication.

The solution to the relaxed problem is the solution to the fully constrained program.

I show this is the case for IC-Structure Γ. The proofs for the other IC Structures are very similar. For the purposes of this proof, I assume symmetry, that is: $\theta_i^A = \theta_i^L = \theta_i$ for $i =$ H, L. Proofs for asymmetric cases are similar.

The seller maximizes expected profit subject to the following restrictions:

$$R_{LL} = \bar{u}$$

$$R_{LH} = q_{LL}^{\prime B}\Delta + \bar{u}$$

$$R_{HL} = q_{LL}^{\prime A}\Delta + \bar{u}$$

$$R_{HH} = \Delta[q_{LL}^{\prime A} + q_{LL}^{\prime B}] + \Delta[q_{LH}^{\prime A} - q_{LL}^{\prime A}] + \Delta[q_{HL}^{\prime B} - q_{LL}^{\prime B}] + \bar{u}$$

$$q_{HH}^{A} \geq q_{LH}^{A},\ q_{HL}^{A} \geq q_{LL}^{A},\ q_{HH}^{B} \geq q_{HL}^{B}, q_{LH}^{B} \geq q_{LL}^{B}$$

Where: $\Delta \equiv \theta_H - \theta_L$

- Goods A and B are served in a "small" portions. $q_{LL}^{*} \equiv q_{LL}^{A*} = q_{LL}^{B*}$.

- There is a "medium" portion. $q_{LH}^{*} \equiv q_{LH}^{A*} = q_{HL}^{B*}$.

- There are "large" portions. $q_{HL}^{*} \equiv q_{HL}^{A*} = q_{LH}^{B*}$ and $q_{HH*} \equiv q_{HH}^{A*} = q_{HH}^{A*}$.

- The quantities consumed by the LL, LH, HL and HH, respectively are: (q_{LL}^{*}, q_{LL}^{*}), (q_{LH}^{*}, q_{HL}^{*}), (q_{HL}^{*}, q_{LH}^{*}), and (q_{HH}^{*}, q_{HH}^{*}).

Consumer Surplus:

$$R_{LL}^{*} = \bar{u}$$

$$R_{LH}^{*} = \Delta q_{LL}^{'*} + \bar{u}$$

$$R_{HL}^{*} = \Delta q_{LL}^{'*} + \bar{u}$$

$$R_{HH}^{*} = 2\Delta q_{LH}^{'*} + \bar{u}$$

Proposition *5. I closely follow the proofs in Armstrong and Rochet (1999). Maximizing 2 subject to 5 gives the solution to the seller's fully constrained problem.*

Proof of proposition 5:

Together, $R_{LL} = 0$, the monotonicity constraints, and the four downward binding constraints imply the satisfaction of the omitted IC constraints:

$$R_{LL} > R_{LH} + q'_{LH}(\theta_L - \theta_H)$$

$$R_{LL} > R_{HL} + q'_{HL}(\theta_L - \theta_H)$$

$$R_{LL} > R_{HH} + 2[q'_{HH}(\theta_L - \theta_H)]$$

From the corresponding first order conditions, it is straightforward to conclude that $q_{HL} > q_{LH}$, thus:

$$R_{LH} > R_{HL} + q'_{HL}(\theta_L - \theta_H) + q'_{LH}(\theta_H - \theta_L)$$

$$R_{HL} > R_{LH} + q'_{LH}(\theta_H - \theta_L) + q'_{HL}(\theta_L - \theta_H)$$

Lastly, the single crossing condition implies:

$$R_{LH} > R_{HH} + q'_{HH}(\theta_H - \theta_L)$$

$$R_{HL} > R_{HH} + q'_{HH}(\theta_L - \theta_H)$$

Proof of proposition 1

There are three parts to show:

First, a mild cap does not change the set of relevant IC constraints. A mild cap is defined as that where the limit is set strictly below the maximum unregulated portion of good A, and at or above the unregulated second largest portion of A. Effectively, the only quantity affected is q^A_{HH}. A new IC constraint would be added to the set to consider if the cap causes it to start binding with equality. The only downward IC restrictions affected are the following, all of which contain q^A_{HH} in the left hand side of the inequality:

$$R_{HH} \geq R_{LL} + (\theta^A_H - \theta^A_L)q'^A_{LL} + (\theta^B_H - \theta^B_L)q'^B_{LL} + \bar{u}$$

$$R_{HH} \geq R_{LH} + (\theta^A_H - \theta^A_L)q'^A_{LH} + (\theta^B_H - \theta^B_H)q'^B_{LH} + \bar{u}$$

$$R_{HH} \geq R_{HL} + (\theta^A_H - \theta^A_H)q'^A_{HL} + (\theta^B_H - \theta^B_L)q'^B_{HL} + \bar{u}$$

These are the same IC already included in the original regulation-free IC-structures. The reduction in q^A_{HH} is small enough that none of them change from potentially binding to always binding.

Second, $LH \to LL$ and $HL \to LL$ remain unchanged and binding in all caps because these IC restrictions do not involve q^A_{HH}.

Lastly, with moderate and severe caps, $LH \to HL$ binds with equality and substitutes $HH \to LL$, $HH \to LH$, and $HH \to HL$. The analysis below concerns to the moderate cap.

To proof this part, first recall that the general form of the IC constraints is the following:

$$R_{ij} \geq R_{kl} + \bar{u} + q_{kl}^{\prime A}(\theta_i^A - \theta_k^A) + q_{kl}^{\prime B}(\theta_j^B - \theta_l^B) \; \forall \; ij \text{ and } kl; i \neq k \text{ and } j \neq l$$

With a moderate cap the specific forms of the all IC program in the complete program are the following:

$$\begin{aligned}
LL \rightarrow LH: &\quad R_{LL} \geq R_{LH} - \Delta^B q_{LH}^{\prime B} \\
LL \rightarrow HL: &\quad R_{LL} \geq R_{HL} - \Delta^A \bar{q}^{\prime} \\
LL \rightarrow HH: &\quad R_{LL} \geq R_{HH} - \Delta^A \bar{q}^{\prime} - \Delta^B q_{HH}^{\prime B} \\
LH \rightarrow LL: &\quad R_{LH} \geq R_{LL} + \Delta^B q_{LL}^{\prime B} \\
LH \rightarrow HL: &\quad R_{LH} \geq R_{HL} - \Delta^A \bar{q}^{\prime} + \Delta^B q_{HL}^{\prime B} \\
LH \rightarrow HH: &\quad R_{LH} \geq R_{HH} - \Delta^A \bar{q}^{\prime} \\
HL \rightarrow LL: &\quad R_{HL} \geq R_{LL} + \Delta^A q_{LL}^{\prime A} \\
HL \rightarrow LH: &\quad R_{HL} \geq R_{LH} + \Delta^A \bar{q}^{\prime} - \Delta^B q_{LH}^{\prime B} \\
HL \rightarrow HH: &\quad R_{HL} \geq R_{HH} - \Delta^B q_{HH}^{\prime B} \\
HH \rightarrow LL: &\quad R_{HH} \geq R_{LL} + \Delta^A q_{LL}^{\prime A} + \Delta^B q_{LL}^{\prime B} \\
HH \rightarrow LH: &\quad R_{HH} \geq R_{LH} + \Delta^A \bar{q}^{\prime} \\
HH \rightarrow HL: &\quad R_{HH} \geq R_{HL} + \Delta^B q_{HL}^{\prime B}
\end{aligned}$$

Because profit maximization necessitates the satisfaction of the participation constraint for the lowest type, the rents earned by the LL-type continue to be $R_{LL} = 0$.

The cap could have changed the set of binding constraints. Regarding the medium types HL and LH, there are three possible candidates for the form of their surplus functions following the cap:

$$\text{Candidate 1:} \begin{cases} R_{LH} = \Delta^B q'^B_{LL} + \bar{u} \\ R_{HL} = \Delta^A q'^A_{LL} + \bar{u} \end{cases}$$

This first candidate set implies that $LH \to LL$ and $HL \to LL$ are the only binding constraints regarding the medium types (HL and LH types) and the LL-type buyer.

$$\text{Candidate 2:} \begin{cases} R_{LH} = \Delta^B q'^B_{LL} + \bar{u} \\ R_{HL} = \Delta^A q'^A_{LL} + \Delta^B q'^B_{LL} + \Delta^A \bar{q'} - \Delta^B q'^B_{LH} + \bar{u} \end{cases}$$

This second candidate set implies that $LH \to LL$, $HL \to LL$, and $HL \to LH$ are binding constraints regarding the medium types (HL and LH types) and the LL-type buyer.

$$\text{Candidate 3:} \begin{cases} R_{LH} = \Delta^B q'^B_{LL} + \Delta^A q'^A_{LL} + \Delta^B q'^B_{HL} - \Delta^A \bar{q'} + \bar{u} \\ R_{HL} = \Delta^A q'^A_{LL} + \bar{u} \end{cases}$$

This last candidate set implies that $LH \to LL$, $HL \to LL$, and $LH \to HL$ are binding constraints regarding the medium types (HL and LH types) and the LL-type buyer.

Candidates 1 and 2 would result in the violation of IC $HH \to LL$. Candidate 2 does not violate the set of IC constraints. Thus, as a result of the cap, the IC constraints $LH \to LL$, $HL \to LL$, and $LH \to HL$ bind with equality.

Notice that $HH \to LH$ and $LH \to HH$ are equivalent under the cap. This implies the bunching together of buyers HH and LH. The IC restrictions $HH \to LH$ and $HH \to HL$

are satisfied as long as $LH \rightarrow LL$, $HL \rightarrow LL$, and $LH \rightarrow HL$ bind with equality. In other words, if the downward IC restrictions involving the LH-type are satisfied, the HH-type will not purchase an option intended to serve either the HL or the LL types.

In other words, as the quantity of product A becomes smaller due to more and more restrictive cap rules, the incentive constraint $LH \rightarrow HL$ becomes relevant.

The analysis for the severe cap is analogous.

Proof of proposition 3

To proof this proposition, it is sufficient to notice that a mild cap does not distort any downward incentive compatibility constraint. A mild cap only reduces q^A_{HH}, and this variable is absent in all IC restrictions shown in 4 and 5.

Proof of proposition 4

This proof has two parts. First, I will show that a moderate cap (where the regulation is intended to reduce the portion below the second largest alternative) impacts q^A_{HH}, q^A_{HL}, and q^A_{LH}. This facilitates showing the welfare effects. The second part addresses each of the three claims within proposition 4 separately.

First part: With unrelated goods and without intervention, the second largest portion is q^A_{LH}. This can be verified as follows. By examining the corresponding first order conditions, it can be deduced that $q^A_{HH} = q^A_{HL}$ in all IC structures. Thus, the candidates for "second largest" portion are either q^A_{LH} and q^A_{LL}. We can find ot which one is largest by comparing the corresponding FOCs.

In IC structures Ψ and Ω, it is straightforward to find out that $q^A_{LH} > q^A_{LL}$.

Moving to IC structures Γ and Υ, bundling requires $\beta_{HH}\beta_{LL} - \beta_{LH}\beta_{HL} < 0$, this implies $\frac{\beta_{HH}}{\beta_{LH}} < \frac{\beta_{HL}}{\beta_{LL}}$. Bearing this in mind, comparing the corresponding FOCs implies $q^A_{LH} > q^A_{LL}$.

Second part: There are three claims within proposition 4.

First, R_{LL} is unaffected. This is straightforward to corroborate, because the participation constraint of this buyer type always binds with equality. In other words, this buyer was already receiving no rents before the cap. Because the LL-type's outside option is zero, no restriction can push R_{LL} below this value.

Second, R_{HL} is affected by the both moderate and severe caps. This is straightforward to corroborate as $R_{HL} = \Delta^B q'^A_{LL}$ and q^A_{LL} is reduced either indirectly or directly by the moderate and severe caps.

Third, R_{LH} increases following either a moderate or severe cap. Before the cap, $R_{LH} = \Delta^B q'^B_{LL}$. As shown in the proof of proposition 1, the rents earned by the LH-type with caps are the following:

$$\text{With moderate cap: } R_{LH} = \Delta^B u(\widetilde{q^B_{LL}}) + \Delta^A \widetilde{q'^A_{LL}} + \Delta^B \widetilde{q'^B_{HL}} - \Delta^A \bar{q'} + \bar{u}$$

$$\text{With severe cap: } R_{LH} = \Delta^B \widehat{q'^B_{LL}} + \Delta^B \widehat{q'^B_{HL}} + \bar{u}$$

Both of them are strictly larger than the base unregulated R_{LH}.

Impact of moderate and severe caps on R_{HH}

***Proposition** 6.* *Moderate and severe caps reduce R_{HH}, the rents earned by the HH-type are negatively affected as long as the following inequalities hold:*

$$\textbf{\textit{ICS }}\Gamma\textbf{\textit{:}} \begin{cases} \textit{Moderate: } \Delta^A q'^A_{LH} + \Delta^B q'^B_{HL} > \Delta^A \widetilde{q'^A_{LL}} + \Delta^B \widetilde{q'^B_{LL}} + \Delta^B \widetilde{q'^B_{HL}} + \bar{u} \\ \textit{Severe: } \Delta^A q'^A_{LH} + \Delta^B q'^B_{HL} > \Delta^A \bar{q'} + \Delta^B \widehat{q'^B_{LL}} + \Delta^B \widehat{q'^B_{HL}} + \bar{u} \end{cases}$$

$$\textbf{\textit{ICS }}\Upsilon\textbf{\textit{:}} \begin{cases} \textit{Moderate: } \Delta^A q'^A_{LH} + \Delta^B q'^B_{LL} > \Delta^A \widetilde{q'^A_{LL}} + \Delta^B \widetilde{q'^B_{LL}} + \Delta^B \widetilde{q'^B_{HL}} + \bar{u} \\ \textit{Severe: } \Delta^A q'^A_{LH} + \Delta^B q'^B_{LL} > \Delta^A \bar{q'} + \Delta^B \widehat{q'^B_{LL}} + \Delta^B \widehat{q'^B_{HL}} + \bar{u} \end{cases}$$

$$\textbf{\textit{ICS }}\Psi\textbf{\textit{:}} \begin{cases} \textit{Moderate: } \Delta^A q'^A_{LL} + \Delta^B q'^B_{HL} > \Delta^A \widetilde{q'^A_{LL}} + \Delta^B \widetilde{q'^B_{LL}} + \Delta^B \widetilde{q'^B_{HL}} + \bar{u} \\ \textit{Severe: } \Delta^A q'^A_{LL} + \Delta^B q'^B_{HL} > \Delta^A \bar{q'} + \Delta^B \widehat{q'^B_{LL}} + \Delta^B \widehat{q'^B_{HL}} + \bar{u} \end{cases}$$

$$\textbf{\textit{ICS }}\Omega\textbf{\textit{:}} \begin{cases} \textit{Moderate: } \Delta^A q'^A_{LL} + \Delta^B q'^B_{LL} > \Delta^A \widetilde{q'^A_{LL}} + \Delta^B \widetilde{q'^B_{LL}} + \Delta^B \widetilde{q'^B_{HL}} + \bar{u} \\ \textit{Severe: } \Delta^A q'^A_{LL} + \Delta^B q'^B_{LL} > \Delta^A \bar{q'} + \Delta^B \widehat{q'^B_{LL}} + \Delta^B \widehat{q'^B_{HL}} + \bar{u} \end{cases}$$

The proof is a simple comparison between the rents earned by the HH-type before and after the cap, so I omit it.

Appendix B: First Order Conditions

IC-Structure Γ baseline:

$$FOC[q^A_{HH}] : \theta^A_H \frac{\partial q'^A_{HH}}{\partial q^A_{HH}} = \frac{\partial c(q^A_{HH})}{\partial q^A_{HH}} \qquad FOC[q^B_{HH}] : \theta^B_H \frac{\partial q'^B_{HH}}{\partial q^B_{HH}} = \frac{\partial c(q^B_{HH})}{\partial q^B_{HH}}$$

$$FOC[q^A_{HL}] : \theta^A_H \frac{\partial q'^A_{HL}}{\partial q^A_{HL}} = \frac{\partial c(q^A_{HL})}{\partial q^A_{HL}} \qquad FOC[q^B_{HL}] : \theta^B_L \frac{\partial q'^B_{HL}}{\partial q^B_{HL}} = \frac{\frac{\partial c(q^B_{HL})}{\partial q^B_{HL}}}{\left(1 - \frac{\beta_{HH}}{\beta_{HL}} \frac{\Delta^B}{\theta^B_L}\right)}$$

$$FOC[q^A_{LH}] : \theta^A_L \frac{\partial q'^A_{LH}}{\partial q^A_{LH}} = \frac{\frac{\partial c(q^A_{LH})}{\partial q^A_{LH}}}{\left(1 - \frac{\beta_{HH}}{\beta_{LH}} \frac{\Delta^A}{\theta^A_L}\right)} \qquad FOC[q^B_{LH}] : \theta^B_H \frac{\partial q'^B_{LH}}{\partial q^B_{LH}} = \frac{\partial c(q^B_{LH})}{\partial q^B_{LH}}$$

$$FOC[q^A_{LL}] : \theta^A_L \frac{\partial q'^A_{LL}}{\partial q^A_{LL}} = \frac{\frac{\partial c(q^A_{LL})}{\partial q^A_{LL}}}{\left(1 - \frac{\beta_{HL}}{\beta_{LL}} \frac{\Delta^A}{\theta^A_L}\right)} \qquad FOC[q^B_{LL}] : \theta^B_L \frac{\partial q'^B_{LL}}{\partial q^B_{LH}} = \frac{\frac{\partial c(q^B_{LL})}{\partial q^B_{LL}}}{\left(1 - \frac{\beta_{LH}}{\beta_{LL}} \frac{\Delta^B}{\theta^B_L}\right)}$$

IC-Structure ϒ baseline:

$$FOC[q^A_{HH}] : \theta^A_H \frac{\partial q'^A_{HH}}{\partial q^A_{HH}} = \frac{\partial c(q^A_{HH})}{\partial q^A_{HH}} \qquad FOC[q^B_{HH}] : \theta^B_H \frac{\partial q'^B_{HH}}{\partial q^B_{HH}} = \frac{\partial c(q^B_{HH})}{\partial q^B_{HH}}$$

$$FOC[q^A_{HL}] : \theta^A_H \frac{\partial q'^A_{HL}}{\partial q^A_{HL}} = \frac{\partial c(q^A_{HL})}{\partial q^A_{HL}} \qquad FOC[q^B_{HL}] : \theta^B_L \frac{\partial q'^B_{HL}}{\partial q^B_{HL}} = \frac{\partial c(q^B_{HL})}{\partial q^B_{HL}}$$

$$FOC[q^A_{LH}] : \theta^A_L \frac{\partial q'^A_{LH}}{\partial q^A_{LH}} = \frac{\frac{\partial c(q^A_{LH})}{\partial q^A_{LH}}}{\left(1 - \frac{\beta_{HH}}{\beta_{LH}} \frac{\Delta^A}{\theta^A_L}\right)} \qquad FOC[q^B_{LH}] : \theta^B_H \frac{\partial q'^B_{LH}}{\partial q^B_{LH}} = \frac{\partial c(q^B_{LH})}{\partial q^B_{LH}}$$

$$FOC[q^A_{LL}] : \theta^A_L \frac{\partial q'^A_{LL}}{\partial q^A_{LL}} = \frac{\frac{\partial c(q^A_{LL})}{\partial q^A_{LL}}}{\left(1 - \frac{\beta_{HL}}{\beta_{LL}} \frac{\Delta^A}{\theta^A_L}\right)} \qquad FOC[q^B_{LL}] : \theta^B_L \frac{\partial q'^B_{LL}}{\partial q^B_{LL}} = \frac{\frac{\partial c(q^B_{LL})}{\partial q^B_{LL}}}{\left(1 - \frac{\beta_{LH}+\beta_{HH}}{\beta_{LL}} \frac{\Delta^B}{\theta^B_L}\right)}$$

IC-Structure Ψ baseline:

$$FOC[q^A_{HH}] : \theta^A_H \frac{\partial q'^A_{HH}}{\partial q^A_{HH}} = \frac{\partial c(q^A_{HH})}{\partial q^A_{HH}} \qquad FOC[q^B_{HH}] : \theta^B_H \frac{\partial q'^B_{HH}}{\partial q^B_{HH}} = \frac{\partial c(q^B_{HH})}{\partial q^B_{HH}}$$

$$FOC[q^A_{HL}] : \theta^A_H \frac{\partial q'^A_{HL}}{\partial q^A_{HL}} = \frac{\partial c(q^A_{HL})}{\partial q^A_{HL}} \qquad FOC[q^B_{HL}] : \theta^B_L \frac{\partial q'^B_{HL}}{\partial q^B_{HL}} = \frac{\frac{\partial c(q^B_{HL})}{\partial q^B_{HL}}}{\left(1 - \frac{\beta_{HH}}{\beta_{HL}} \frac{\Delta^B}{\theta^B_L}\right)}$$

$$FOC[q^A_{LH}] : \theta^A_L \frac{\partial q'^A_{LH}}{\partial q^A_{LH}} = \frac{\partial c(q^A_{LH})}{\partial q^A_{LH}} \qquad FOC[q^B_{LH}] : \theta^B_H \frac{\partial q'^B_{LH}}{\partial q^B_{LH}} = \frac{\partial c(q^B_{LH})}{\partial q^B_{LH}}$$

$$FOC[q^A_{LL}] : \theta^A_L \frac{\partial q'^A_{LL}}{\partial q^A_{LL}} = \frac{\frac{\partial c(q^A_{LL})}{\partial q^A_{LL}}}{\left(1 - \frac{\beta_{HL}+\beta_{HH}}{\beta_{LL}} \frac{\Delta^A}{\theta^A_L}\right)} \qquad FOC[q^B_{LL}] : \theta^B_L \frac{\partial q'^B_{LL}}{\partial q^B_{LL}} = \frac{\frac{\partial c(q^B_{LL})}{\partial q^B_{LL}}}{\left(1 - \frac{\beta_{LH}}{\beta_{LL}} \frac{\Delta^B}{\theta^B_L}\right)}$$

IC-Structure Ω baseline:

$$FOC[q^A_{HH}] : \theta^A_H \frac{\partial q'^A_{HH}}{\partial q^A_{HH}} = \frac{\partial c(q^A_{HH})}{\partial q^A_{HH}} \qquad FOC[q^B_{HH}] : \theta^B_H \frac{\partial q'^B_{HH}}{\partial q^B_{HH}} = \frac{\partial c(q^B_{HH})}{\partial q^B_{HH}}$$

$$FOC[q^A_{HL}] : \theta^A_H \frac{\partial q'^A_{HL}}{\partial q^A_{HL}} = \frac{\partial c(q^A_{HL})}{\partial q^B_{HL}} \qquad FOC[q^B_{HL}] : \theta^B_L \frac{\partial q'^B_{HL}}{\partial q^B_{HL}} = \frac{\partial c(q^B_{HL})}{\partial q^B_{HL}}$$

$$FOC[q^A_{LH}] : \theta^A_L \frac{\partial q'^A_{LH}}{\partial q^A_{LH}} = \frac{\partial c(q^A_{LH})}{\partial q^A_{LH}} \qquad FOC[q^B_{LH}] : \theta^B_H \frac{\partial q'^B_{LH}}{\partial q^B_{LH}} = \frac{\partial c(q^B_{LH})}{\partial q^B_{LH}}$$

$$FOC[q^A_{LL}] : \theta^A_L \frac{\partial q'^A_{LL}}{\partial q^A_{LL}} = \frac{\frac{\partial c(q^A_{LL})}{\partial q^A_{LL}}}{\left(1 - \frac{\beta_{HL}+\beta_{HH}}{\beta_{LL}} \frac{\Delta^A}{\theta^A_L}\right)} \qquad FOC[q^B_{LL}] : \theta^B_L \frac{\partial q'^B_{LL}}{\partial q^B_{LL}} = \frac{\frac{\partial c(q^B_{LL})}{\partial q^B_{LL}}}{\left(1 - \frac{\beta_{LH}+\beta_{HH}}{\beta_{LL}} \frac{\Delta^B}{\theta^B_L}\right)}$$

Mild IC-Structure Γ:

$$FOC[\bar{q}] : \theta_H^A \frac{\partial \bar{q}'}{\partial \bar{q}} = \frac{\partial c(\bar{q})}{\partial \bar{q}} \qquad FOC[q_{HH}^B] : \theta_H^B \frac{\partial q_{HH}'^B}{\partial q_{HH}^B} = \frac{\partial c(q_{HH}^B)}{\partial q_{HH}^B}$$

$$FOC[q_{HL}^A] : \theta_H^A \frac{\partial q_{HL}'^A}{\partial q_{HL}^A} = \frac{\partial c(q_{HL}^A)}{\partial q_{HL}^A} \qquad FOC[q_{HL}^B] : \theta_L^B \frac{\partial q_{HL}'^B}{\partial q_{HL}^B} = \frac{\frac{\partial c(q_{HL}^B)}{\partial q_{HL}^B}}{\left(1 - \frac{\beta_{HH}}{\beta_{HL}} \frac{\Delta^B}{\theta_L^B}\right)}$$

$$FOC[q_{LH}^A] : \theta_L^A \frac{\partial q_{LH}'^A}{\partial q_{LH}^A} = \frac{\frac{\partial c(q_{LH}^A)}{\partial q_{LH}^A}}{\left(1 - \frac{\beta_{HH}}{\beta_{LH}} \frac{\Delta^A}{\theta_L^A}\right)} \qquad FOC[q_{LH}^B] : \theta_H^B \frac{\partial q_{LH}'^B}{\partial q_{LH}^B} = \frac{\partial c(q_{LH}^B)}{\partial q_{LH}^B}$$

$$FOC[q_{LL}^A] : \theta_L^A \frac{\partial q_{LL}'^A}{\partial q_{LL}^A} = \frac{\frac{\partial c(q_{LL}^A)}{\partial q_{LL}^A}}{\left(1 - \frac{\beta_{HL}}{\beta_{LL}} \frac{\Delta^A}{\theta_L^A}\right)} \qquad FOC[q_{LL}^B] : \theta_L^B \frac{\partial q_{LL}'^B}{\partial q_{LL}^B} = \frac{\frac{\partial c(q_{LL}^B)}{\partial q_{LL}^B}}{\left(1 - \frac{\beta_{LH}}{\beta_{LL}} \frac{\Delta^B}{\theta_L^B}\right)}$$

Mild cap IC-Structure Υ:

$$FOC[\bar{q}] : \theta_H^A \frac{\partial \bar{q}'}{\partial \bar{q}} = \frac{\partial c(\bar{q})}{\partial \bar{q}} \qquad FOC[q_{HH}^B] : \theta_H^B \frac{\partial q_{HH}'^B}{\partial q_{HH}^B} = \frac{\partial c(q_{HH}^B)}{\partial q_{HH}^B}$$

$$FOC[q_{HL}^A] : \theta_H^A \frac{\partial q_{HL}'^A}{\partial q_{HL}^A} = \frac{\partial c(q_{HL}^A)}{\partial q_{HL}^A} \qquad FOC[q_{HL}^B] : \theta_L^B \frac{\partial q_{HL}'^B}{\partial q_{HL}^B} = \frac{\partial c(q_{HL}^B)}{\partial q_{HL}^B}$$

$$FOC[q_{LH}^A] : \theta_L^A \frac{\partial q_{LH}'^A}{\partial q_{LH}^A} = \frac{\frac{\partial c(q_{LH}^A)}{\partial q_{LH}^A}}{\left(1 - \frac{\beta_{HH}}{\beta_{LH}} \frac{\Delta^A}{\theta_L^A}\right)} \qquad FOC[q_{LH}^B] : \theta_H^B \frac{\partial q_{LH}'^B}{\partial q_{LH}^B} = \frac{\partial c(q_{LH}^B)}{\partial q_{LH}^B}$$

$$FOC[q_{LL}^A] : \theta_L^A \frac{\partial q_{LL}'^A}{\partial q_{LL}^A} = \frac{\frac{\partial c(q_{LL}^A)}{\partial q_{LL}^A}}{\left(1 - \frac{\beta_{HL}}{\beta_{LL}} \frac{\Delta^A}{\theta_L^A}\right)} \qquad FOC[q_{LL}^B] : \theta_L^B \frac{\partial q_{LL}'^B}{\partial q_{LL}^B} = \frac{\frac{\partial c(q_{LL}^B)}{\partial q_{LL}^B}}{\left(1 - \frac{\beta_{LH}+\beta_{HH}}{\beta_{LL}} \frac{\Delta^B}{\theta_L^B}\right)}$$

Mild cap IC-Structure Ψ:

$$FOC[\bar{q}] : \theta_H^A \frac{\partial \bar{q}'}{\partial \bar{q}} = \frac{\partial c(\bar{q})}{\partial \bar{q}} \qquad FOC[q_{HH}^B] : \theta_H^B \frac{\partial q_{HH}'^B}{\partial q_{HH}^B} = \frac{\partial c(q_{HH}^B)}{\partial q_{HH}^B}$$

$$FOC[q_{HL}^A] : \theta_H^A \frac{\partial q_{HL}'^A}{\partial q_{HL}^A} = \frac{\partial c(q_{HL}^A)}{\partial q_{HL}^A} \qquad FOC[q_{HL}^B] : \theta_L^B \frac{\partial q_{HL}'^B}{\partial q_{HL}^B} = \frac{\frac{\partial c(q_{HL}^B)}{\partial q_{HL}^B}}{\left(1 - \frac{\beta_{HH}}{\beta_{HL}} \frac{\Delta^B}{\theta_L^B}\right)}$$

$$FOC[q_{LH}^A] : \theta_L^A \frac{\partial q_{LH}'^A}{\partial q_{LH}^A} = \frac{\partial c(q_{LH}^A)}{\partial q_{LH}^A} \qquad FOC[q_{LH}^B] : \theta_H^B \frac{\partial q_{LH}'^B}{\partial q_{LH}^B} = \frac{\partial c(q_{LH}^B)}{\partial q_{LH}^B}$$

$$FOC[q_{LL}^A] : \theta_L^A \frac{\partial q_{LL}'^A}{\partial q_{LL}^A} = \frac{\frac{\partial c(q_{LL}^A)}{\partial q_{LL}^A}}{\left(1 - \frac{\beta_{HL}+\beta_{HH}}{\beta_{LL}} \frac{\Delta^A}{\theta_L^A}\right)} \qquad FOC[q_{LL}^B] : \theta_L^B \frac{\partial q_{LL}'^B}{\partial q_{LL}^B} = \frac{\frac{\partial c(q_{LL}^B)}{\partial q_{LL}^B}}{\left(1 - \frac{\beta_{LH}}{\beta_{LL}} \frac{\Delta^B}{\theta_L^B}\right)}$$

Mild IC-Structure Ω:

$$FOC[\bar{q}] : \theta_H^A \frac{\partial \bar{q'}}{\partial \bar{q}} = \frac{\partial c(\bar{q})}{\partial \bar{q}}$$

$$FOC[q_{HH}^B] : \theta_H^B \frac{\partial {q'}_{HH}^B}{\partial q_{HH}^B} = \frac{\partial c(q_{HH}^B)}{\partial q_{HH}^B}$$

$$FOC[q_{HL}^A] : \theta_H^A \frac{\partial {q'}_{HL}^A}{\partial q_{HL}^A} = \frac{\partial c(q_{HL}^A)}{\partial q_{HL}^A}$$

$$FOC[q_{HL}^B] : \theta_L^B \frac{\partial {q'}_{HL}^B}{\partial q_{HL}^B} = \frac{\partial c(q_{HL}^B)}{\partial q_{HL}^B}$$

$$FOC[q_{LH}^A] : \theta_L^A \frac{\partial {q'}_{LH}^A}{\partial q_{LH}^A} = \frac{\partial c(q_{LH}^A)}{\partial q_{LH}^A}$$

$$FOC[q_{LH}^B] : \theta_H^B \frac{\partial {q'}_{LH}^B}{\partial q_{LH}^B} = \frac{\partial c(q_{LH}^B)}{\partial q_{LH}^B}$$

$$FOC[q_{LL}^A] : \theta_L^A \frac{\partial {q'}_{LL}^A}{\partial q_{LL}^A} = \frac{\frac{\partial c(q_{LL}^A)}{\partial q_{LL}^A}}{\left(1 - \frac{\beta_{HL}+\beta_{HH}}{\beta_{LL}} \frac{\Delta^A}{\theta_L^A}\right)}$$

$$FOC[q_{LL}^B] : \theta_L^B \frac{\partial {q'}_{LL}^B}{\partial q_{LL}^B} = \frac{\frac{\partial c(q_{LL}^B)}{\partial q_{LL}^B}}{\left(1 - \frac{\beta_{LH}+\beta_{HH}}{\beta_{LL}} \frac{\Delta^B}{\theta_L^B}\right)}$$

Moderate cap:

$$FOC[\bar{q}] : \theta_H^A \frac{\partial \bar{q'}}{\partial \bar{q}} = \frac{\beta_{HH} + \beta_{HL} + \beta_{LH}}{\beta_{HH} + \beta_{HL} + \beta_{LH}\left(\frac{\theta_L^A}{\theta_H^A} + \frac{\Delta^A}{\theta_H^A}\right)} \cdot \frac{\partial c(\bar{q})}{\partial \bar{q}}$$

$$FOC[q_{HH}^B] : \theta_H^B \frac{\partial {q'}_{HH}^B}{\partial q_{HH}^B} = \frac{\partial c(q_{HH}^B)}{\partial q_{HH}^B}$$

$$FOC[q_{HL}^B] : \theta_L^B \frac{\partial {q'}_{HL}^B}{\partial q_{HL}^B} = \frac{\frac{\partial c(q_{HL}^B)}{\partial q_{HL}^B}}{\left(1 - \frac{\beta_{LH}+\beta_{HH}}{\beta_{HL}} \frac{\Delta^B}{\theta_L^B}\right)}$$

$$FOC[q_{LH}^B] : \theta_H^B \frac{\partial {q'}_{LH}^B}{\partial q_{LH}^B} = \frac{\partial c(q_{LH}^B)}{\partial q_{LH}^B}$$

$$FOC[q_{LL}^A] : \theta_L^A \frac{\partial {q'}_{LL}^A}{\partial q_{LL}^A} = \frac{\frac{\partial c(q_{LL}^A)}{\partial q_{LL}^A}}{\left(1 - \frac{\beta_{LH}+\beta_{HL}+\beta_{HH}}{\beta_{LL}} \frac{\Delta^A}{\theta_L^A}\right)}$$

$$FOC[q_{LL}^B] : \theta_L^B \frac{\partial {q'}_{LL}^B}{\partial q_{LL}^B} = \frac{\frac{\partial c(q_{LL}^B)}{\partial q_{LL}^B}}{\left(1 - \frac{\beta_{LH}+\beta_{HH}}{\beta_{LL}} \frac{\Delta^B}{\theta_L^B}\right)}$$

Severe cap

$$FOC[\bar{q}] : \theta_H^A \frac{\partial \bar{q}'}{\partial \bar{q}} = \frac{\theta_H^A}{\theta_L^A} \frac{\partial c(\bar{q})}{\partial \bar{q}}$$

$$FOC[q_{HH}^B] : \theta_H^B \frac{\partial q_{HH}'^B}{\partial q_{HH}^B} = \frac{\partial c(q_{HH}^B)}{\partial q_{HH}^B}$$

$$FOC[q_{HL}^B] : \theta_L^B \frac{\partial q_{HL}'^B}{\partial q_{HL}^B} = \frac{\frac{\partial c(q_{HL}^B)}{\partial q_{HL}^B}}{\left(1 - \frac{\beta_{HH}+\beta_{LH}}{\beta_{HL}} \frac{\Delta^B}{\theta_L^B}\right)}$$

$$FOC[q_{LH}^B] : \theta_H^B \frac{\partial q_{LH}'^B}{\partial q_{LH}^B} = \frac{\partial c(q_{LH}^B)}{\partial q_{LH}^B}$$

$$FOC[q_{LL}^B] : \theta_L^B \frac{\partial q_{LL}'^B}{\partial q_{LL}^B} = \frac{\frac{\partial c(q_{LL}^B)}{\partial q_{LL}^B}}{\left(1 - \frac{\beta_{HH}+\beta_{LH}}{\beta_{LL}} \frac{\Delta^B}{\theta_L^B}\right)}$$