Competitive Provision of Digital Goods[*]

Elia Sartori[†]

November 14, 2018

Job Market Paper

Abstract

Digital goods are produced along a quality ranking and can be both *duplicated* and *damaged* at zero marginal cost. Consumers' valuation of quality consists of a common decreasing returns component and an heterogeneous component that gives sellers a motive for screening. The monopolist problem is naturally divided into an acquisition and a distribution stage; two interdependent sources of inefficiency, underprovision and quality damaging, emerge. Competition is modeled as a two stage game of perfect information, in which active firms acquire market power through an irreversible investment in quality. The monopolistic allocation emerges as one equilibrium but there are also equilibria with active competition. The welfare comparison between monopoly and duopoly is ambiguous: additional underacquisition and double spending favor the former, undoing damaging inefficiencies by distributing a positive quality for free favors the latter.

[*]I am indebted to my advisor, Stephen Morris, for encouragement and support in developing this project. I am also grateful to Faruk Gul and Wolfgang Pesendorfer for insightful conversations. For helpful comments I would like to thank Pierpaolo Battigalli, Massimo Marinacci, Michele Fornino, Pietro Ortoleva, Franz Ostrizek, Alessandro Pavan, Giorgio Saponaro, and seminar audiences at Princeton.

[†]Princeton University, esartori@princeton.edu

1 Introduction

We study the distribution of goods that are produced along a quality ranking and that can be both duplicated and damaged at zero marginal cost. A firm that creates a version of the good of quality q can, at no additional cost, sell arbitrary amounts of any quality below q. Consumers demand at most one version of the good, they agree on the quality ranking but have different tastes for quality. Such heterogeneity makes producers with market power willing to engage in inefficient quality damaging for screening purposes as in the literature of multiproduct monopolist as in Mussa and Rosen (1978), hereafter referred to as MR.

Two markets whose functioning is well approximated by this model are the market for digital content, (computer software, mobile apps, digital audio and video content), and some portions of the market for information (weather forecasts, non-strategic financial information). The former is a large and growing sector of advanced economies, while the latter is of interest because it is natural to assume that sellers have access to a free garbling technology that allows for damaging of information structures. Until Section 4, where we formalize the application to information markets, we will not discuss how particular results explain phenomena observed in those markets, and generically refer to *digital goods* as products that have the following characteristics:

1. They are *non-rival* but *excludable* through a pricing system;[1]
2. Produced and damaged along *single-dimensional quality ranking* that is given exogenously and on which all consumers agree. This excludes the possibility of horizontal differentiation across consumers. However,
3. Consumers have *heterogenous tastes* for quality.
4. When multiple firms are active, their products are *homogenous*: Individuals never want to combine qualities sold by different producers.
5. On the production side, *replication and damaging* of a version occurs at negligible cost, and damaging must lie on the pre-specified quality ranking.

We will assume consumers' preferences for quality (point 3 above) are separable in a common decreasing returns component and an heterogeneous constant returns component. We adopt this specification for two reasons. First, the interpretation we give in Section 2.1.1 in which agents use the digital good to perform two tasks (basic and professional activities) may be a reasonable description of the demand for digital goods (we indeed motivate it with an example of software consumption). Second, it is a parsimonious specification that allows for rich empirical implications of the model. The standard linear preferences used in MR, which will be presented as a subcase, are unable to generate an optimal contract that displays non-trivial damaging: we show that if the decreasing returns component were absent, then only one positive quality would be sold in the market. This is counterfactual, at least in some markets: to give just one example, Figure 1.1 below shows a set of packages for statistical software that differ in their computational power. Likewise, digital content

[1] An essential non-rivalry arises because of the free replicability. In general, no consumption externalities are allowed, which is a particularly restrictive assumption in the market for information where a recent literature focused on the fact that the value of information is an equilibrium object. The excludability issue is also critical; a whole literature (Muto (1986), Varian (2000), Polanski (2007) among others) focuses on the distribution of non-excludable "information goods" that in their definition include also software and books.

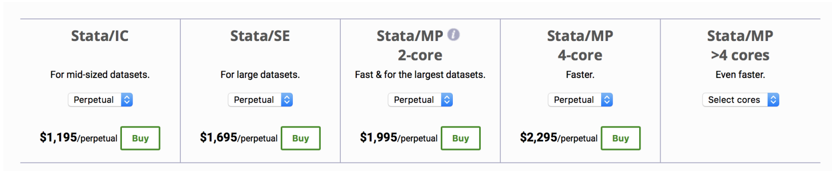

Figure 1.1: Non-trivial damaging in the distribution of Stata

is often offered in SD or HD packages and Section 4 gives examples of non-trivial screening in information markets.

An additional observation motivating our analysis is that goods of positive quality are distributed for free in many of the markets with the characteristics described above. An enormous amount of information is available at no (monetary) cost. The same is true for online services, ranging from e-mails to document storage and digital contents. As for computer software, it is interesting to notice that Open Office was released in 2002, 12 years after Microsoft sold the first Office package, possibly as a consequence of increased competition. Many mobile apps are also sold for free, even though premium options are often present and, unlike (some) computer software, are immune to failures of non-excludability.[2] The literature offers some explanations for the free-quality phenomenon, ranging from creating costumer fidelization to be exploited in parallel markets (bundling and cross-fidelization), to earning profits from individual attention (through advertising). In this paper we build a simple model that can both create, under certain assumptions about primitives and the competition structure, the positive implications discussed above (non-trivial damaging and distribution of free quality), and that is flexible enough to answer some questions like: Is such free distribution socially desirable? What is the welfare impact of some policies in this framework (damaging prohibition, linear taxation, patent protection)?

1.1 Outline of the paper and preview of results

Section 2 formalizes the primitives and characterizes the provision of digital goods by a monopolist under both perfect and asymmetric information. As in standard first-degree price discrimination problems, the former setting induces the first-best allocation which features no quality damaging. The latter problem is more complicated. We show that it can be conveniently rewritten as the maximization over a sequence of MR problems parametrized by the quality cap constraining the monopolist's allocation function. Each problem in this sequence, not the original one, can be solved applying standard monopolist screening techniques with a particular (zero) cost function. We show that the optimal contract conditional on a quality cap allocates each type the minimum between an increasing type-dependent function and the quality cap itself. A bunching threshold moves as the quality cap increases, which raises rents of high valuation types while leaving rents of lower types unchanged. This property makes it simple to characterize the marginal revenue function and hence solve the quality-acquisition problem. The two-stage nature of the monopolist problem generates two sources of inefficiency: an acquisition inefficiency similar to standard underprovision with

[2] Many online apps profit from matching demand and supply for either transportation (Uber, Lyft) or for food delivery (Deliveroo, Foodora). They also actively engage in some sort of screening: ride hailing apps charge a higher price for larger or more comfortable cars, while food delivery apps may propose early delivery for a surcharge. Those are not examples that fit our description, since both "premium" services (bigger car and fast delivery) require a higher *marginal* cost to the producer: the delivery guy has to run a motorbike rather than a bike and, similarly, the premium car has a higher depreciation/fuel cost.

market power, and a damaging inefficiency from asymmetric information. The two are interdependent: distribution obviously depends on the quality constraint, and incentives to acquire depend on the revenues that damaging can achieve. In particular the efficiency at the top typical of standard screening problems is limited to a distributional efficiency: a positive measure of types never receives a damaged quality but even the highest type receives a quality below what he gets in the first best. Curvature of the common component implies that at low quality levels optimal distribution features no damaging, and that in all contracts all types receive positive quality (and surplus). With linear preferences (no concave component) a "no haggling" result holds: a cap-invariant set of types is always served the undamaged quality while others are fully excluded (given $q = 0$), so marginal revenues are constant. We conclude Section 2 by analyzing the impact of a No Screening (NS) policy, namely to prohibit the seller from engaging in quality damaging for screening purposes. When binding, the policy is proved to always worsen the underacquisition inefficiency. As for its impact on damaging, two forces operate: the NS policy mechanically prevents inefficient damaging, though it may induce the complete exclusion of some low types that received positive surplus in the unconstrained monopoly contract. We find conditions under which the NS policy is welfare improving.

Section 3 studies competition in digital goods markets as the equilibrium of a two stage game of perfect information. The first stage, investment in quality, determines firms' market power at the pricing stage. The second stage (pricing) equilibrium is easily characterized using the tools developed to solve the monopolist problem: the owner of the larger quality behaves indeed as a (interim-)monopolist on the quality spectrum he owns exclusively, while Bertrand forces drive the price of the second highest quality to zero. In the first stage there are multiple equilibria indexed by n, the number of firms that are active (i.e. choosing to acquire a positive quality with positive probability). With $n = 1$ the active firm is a monopolist for sure (the only pure strategy equilibrium) and agents receive the monopolist allocation of Section 2. With any $n \geq 2$ there is a symmetric equilibrium in which active firms randomize investment with full support ranging from zero to the monopolist quality. In the class of equilibria with active competition, we prove that every type in the economy is better-off with a lower intensity of competition (smaller n). The relevant welfare comparison is therefore between monopoly and duopoly. By only looking at the support of the mixed equilibrium we observe that the highest quality distributed under competition will be below the monopolist quality (increasing underprovision inefficiency). Also, active competition implies inefficient double spending: due to homogeneity, the development cost of an inferior quality is always socially wasteful. However, by distributing a positive quality for free, competition shrinks screening inefficiencies associated to each realized best quality. We prove that we cannot go beyond this qualitative comparison between monopoly and duopoly equilibria: different shapes of the cost function can shut down almost completely either the positive or the negative impact of competition. In particular, if the monopolist was not damaging, then competition unambiguously reduces total welfare. By contrast, if costs are extremely convex (approaching a fixed cost structure), then a competitive market induces stochastic allocations that converge to the flat allocation where everybody receives the quality produced (but not distributed) by a monopolist and hence dominates the monopolistic equilibrium.

Section 4 is independent of the other two and assesses the fitness of the framework presented before to study information markets, relative to existing approaches that model information acquisition. We discuss how modeling choices used in the paper translate into implicit assumptions on the type of information markets that can be analyzed. The key element is that production of information is decentralized and that the technology to convert the factor of production (attention) into state-signal structures must be taken as a primitive. We present a simple but exact microfoundation of the reduced form model studied in the paper and discuss how correlation in primary information structures can be used to create product heterogeneity which is necessary to avoid some implications of the model that are counterfactual in those market (only one firm makes profits). The dimensionality of feasible signal structures and of payoff relevant types, even with a "small" state space, poses significant tractability challenges that are beyond the scope of this paper but suggest directions for future research.

1.2 Literature review

In this section we review technical contributions related to the building blocks of the model: the demand side, the composition of replicability and damaging on the production technology, and a description of model of competition with screening. Models of information markets are reviewed separately in the last section.

Quality screening The demand side of the economy and the solution techniques for the quality-conditional problem are based on the literature on screening with a multiproduct monopolist pioneered by Mussa and Rosen (1978), advanced in Maskin and Riley (1984) and later in Wilson (1993). Assumptions in this paper make sure to avoid ironing and other technical complications *within* each MR problem (which are the focus of the original paper and Rochet and Choné (1998)), and to have a simple revenue comparison *across* different problems.

Free Replicability Motivated by the example of software and digital contents, a recent literature in computer science Goldberg and Hartline (2003); Goldberg et al. (2006); Hartline and Roughgarden (2008) studied how to design the revenue maximizing mechanism to allocate a good that is replicable for free to agents with heterogeneous valuations. In particular Goldberg et al. (2001) show that posted price mechanism performs surprisingly close to the optimal (possibly dynamic) incentive compatible auction. This result partially justifies my focus of screening through (a menu of) prices.[3]

Damaging The idea of damaging a good for screening purposes was originally introduced in Deneckere and Preston McAfee (1996).[4] The approach in this paper is different both in modeling choice and in the type of questions addressed. From a modeling perspective, beyond preserving a positive marginal cost from distribution,[5] Deneckere and McAfee (1996) take a binary set of qualities as exogenously fixed, thereby excluding an acquisition margin, and assume that the only way to produce the good of low quality is by damaging the high quality good. Marginal distribution costs are therefore *larger* for the low quality good.[6] They focus on the monopolist problem and address the following question: When is it the case that the possibility to damage benefits all agents in the economy?[7]

Product versioning through quality damaging has been explored also in the context of a durable good monopolist. In related papers, Inderst (2008) and Hahn (2006) consider an environment with two consumer

[3]

It is not a complete justification to my approach as they don't allow damaging of the replicable good. An extension of their results to cases in which the seller can damage the good would be interesting per se.

[4]Srinagesh and Bradburd (1989) offer a very general analysis for the case where there are two types of customer. McAfee (2007) provides an exact characterization in terms of marginal revenues of when damaging is profitable.

[5]Their motivating examples include processors, printers and other technological products.

[6]Clearly, under those assumptions it is a more startling fact that a monopolist is sometimes willing to engage in screening.

[7]Section 2.5 addresses a natural extension of this question within our framework.

types and a monopolist that sells different versions of a product over time and faces a Coasian commitment problem in price and quality.[8]

Availability of a damaged quality may also result from illegal activities such as piracy (assuming the copied version is somehow inferior to the original).[9] Peitz and Waelbroeck (2006) provides a critical overview of the theoretical literature that addresses the economic consequences of end-user copying, though focusing mostly on the non-excludability of low qualities that is induced by the illegal activity.[10]

Competition with screening A vast literature on competition with screening spurred from the seminal contribution of Rothschild and Stiglitz (1976) (RS) on insurance markets. RS equilibrium invokes a natural notion of stability induced by a free entry condition which can be interpreted as the Nash equilibrium of a contract[11] posting game among many (ex-ante) symmetric firms. Existence is not guaranteed. The advancement of game theory allowed a more formal analysis of the strategic interaction among competitors: Jaynes (1978) shows the RS equilibria always exist if sharing of information about customers is treated endogenously as part of the game among firms. Hellwig (1988) shows that this is true if each firm's communication strategy is conditioned on the set of contracts that are offered by the other firms.[12]

Departures from modeling equilibrium with asymmetric information as the outcome of an extensive form game produced some elegant characterizations. Dubey and Geanakoplos (2002) study the RS model by fixing an exogenous set of pools characterized by their limits on contributions. Households signal their reliability by choosing which to join. They put discipline on beliefs over pools that are not visited in equilibrium and prove existence (and uniqueness) of the separating RS equilibrium. Bisin and Gottardi (1999) and Bisin et al. (2011) extend the the model of general competitive equilibrium to economies with asymmetric information without having to explicitly model private information.[13]

Possibly due to a lack of tractability of the latter models, the literature even in recent years has kept analyzing competition under asymmetric information as the equilibrium of an extensive form game. This is the approach taken also in this paper.[14] Netzer and Scheuer (2010) extend the RS model to two-dimensional heterogeneity in both risk and patience where the latter is the endogenous result of optimal savings or labor supply decisions. In their model RS equilibria exist and equilibrium contracts can earn strictly positive profits because any contract that attracts good consumers would also attract bad risk types and become unprofitable. A similar result of profitable contracts in RS equilibria is obtained in the two dimensional screening model of Smart (2000) where both dimensions are exogenous.

However, in many cases RS equilibria fail to exist due to the many deviations available to the pool of

[8]In their setting the seller may optimally choose to engage in quality deterioration in the first period (and trade only occurs in this period) by selling the low-quality version below marginal cost in the first period to avoid later price concessions to high-valuation consumers.

[9]Takeyama (1994) argues that the loss in profits due to copying may be greater if dynamic effects are taken into account; however it is also possible that the seller benefits from being copied since this reduces his commitment problem.

[10]Hence connecting mostly with the literature on non-excludable "information goods" referred to in footnote 1.

[11]In their case, a contract is a set of insurance rate and price.

[12] With the use a dynamic games, the analysis of contracting with adverse selection was extended to the other issues, among others that of *renegotiation* (e.g. Hart and Tirole (1988)) and *recontracting* (Beaudry and Poitevin (1995) model a financial market where the informed party has the bargaining power even though competing uninformed parties make the offer).

[13]Wilson (1978), Dutta and Vohra (2005) and Vohra (1999) propose an extension of the core as a positive foundation of equilibria under asymmetric information.

[14]Although a strategic foundation may sometimes be an appealing feature of the model, it adds one degree of arbitrariness: by looking at the simplest model of competition and how Cournot and Bertrand equilibria differ in their implication we understand how crucial even the specification of the action space may be. Selection of the game structure is mainly driven by the tractability of equilibria it delivers; alternative specifications are discussed in Section 3.4.

potential entrants. One solution in this case is to give firms some market power at the pricing stage. Garrett et al. (2014) have a model in which market power is given exogenously by assuming that also consumers are imperfectly informed about the offers in the market (two-sided asymmetric information). They show that the intensity of competition decreases this source of market power, so in the limit the Bertrand equilibrium emerges. Having firms commit through an ex-ante irreversible investment is a second way of creating (this time endogenously) market power, used since Hotelling (1929). A standard reference for this approach is Kreps and Scheinkman (1983), who have firms commit to a *quantity* level before Bertrand competing on the realized investments. The setup closer to that of this paper is Champsaur and Rochet (1989) who analyze a MR duopoly where each competitor costlessly commits to a subset of qualities and then chooses a pricing function (paying the distribution costs at this stage). In section 3.4 we compare the properties of the equilibria of their models with results in this paper.

2 The Monopolist Problem

2.1 Primitives and Efficiency Benchmark

2.1.1 Demand

The economy is populated by a unit mass of consumers. Each consumer is characterized by a utility type $\theta \in \Theta$, where Θ is a compact subset of $\mathbb{R}_+$. $F : \Theta \to [0,1]$ strictly increasing and admitting a density is the population distribution function. Utility types describe an agent's cardinal rankings over different quality versions of the digital good. Consumers' valuation for quality are assumed to take the functional form

$$u(q,\theta) = g(q) + \theta q \tag{2.1}$$

where g is a concave function representing the relative curvature of the common component of quality ranking with respect to the type dependent one.[15] The degenerate case of $g \equiv 0$, i.e linear utility, will be an important subcase for two reasons: it is the utility specification adopted in MR among others and also it delivers an extreme version of our screening results. When non-degenerate, it is assumed that g satisfies the Inada conditions

$$\lim_{x\to 0} g'(x) = \infty, \qquad \lim_{x\to\infty} g'(x) = 0 \tag{2.2}$$

Agents also own a large amount of a numeraire good and have quasilinear preferences in this good. So the demand correspondence associated to a quality pricing function $\boldsymbol{p} : Q \to \mathbb{R}$ is given by

$$D_{\boldsymbol{p}}(\theta) = \arg\max_q u(q,\theta) - \boldsymbol{p}(q)$$

[15] As quality does not have a natural metric, we can consider a more general setting in which for two increasing function g_1, g_2

$$u(x,\theta) = g_1(x) + \theta g_2(x)$$

then define quality $q = g_2(x)$ with associated cardinal rankings

$$u(q,\theta) = g_1\left(g_2^{-1}(q)\right) + \theta q$$

The cost function over the new quality space can be redefined in a similar fashion. The qualitative results would not change, what is key is that the type independent component $g_1 \circ g_2^{-1}$ is concave or, that g_1 is "more concave" than g_2. The empirical content driving the results is that the common valuation of the quality is the additive separability in types and the fact that the function multiplying type is "less concave" than the common quality ranking.

Returns from quality: Interpretation and properties

The utility specification (2.1)-(2.2) is important to deliver some properties of the optimal contract.[16] We now justify it by offering an interpretation in the context of software consumption, and we identify the analytical properties that drive our results. For an interpretation, suppose consumers use the digital good to perform two tasks. All users perform the same basic task and measure returns to quality in the accomplishment of this task according to a common decreasing returns function; they also perform an advanced task, but they have heterogeneous constant marginal return types θ, which measure the intensity of individuals' tastes for quality in the accomplishment of such task. To substantiate the assumption, we use the software (OS) example and broadly define q as computational power. The set of basic tasks include simple calculations, text editing and other activities that are performed by everyone in essentially the same way and for which the returns to quality are very steep at the beginning, but then vanish. The advanced task is a professional activity in which each consumer specializes and that may be more or less computationally intensive. At a low level of q there is little (relative) variation in marginal utilities as everyone cares essentially about the steep improvement in the performance of the basic task, while at large q the demand for improvement is driven by the use one can make in the advanced task, which is heterogeneous.

From a technical standpoint, additive separability and concavity of g gives rise to a constant difference in the marginal utility $g'(q) + \theta$ between any two types. Yet, since g' is decreasing q, their ratio

$$\frac{g'(q) + \theta}{g'(q) + \theta'}$$

is also a decreasing function of q. The type dependent component θ within the marginal valuation $g'(q) + \theta$ becomes dominant as we climb up the quality ladder. As the marginal willingness to pay for a quality improvement for a high type relative to a low type increases in the level of quality, it becomes profitable to screen type θ from θ' only when the quality level is large enough.

2.1.2 Production and Sale

The digital good can be produced along a continuum of versions or qualities, where $Q = \mathbb{R}_+$ is the quality space. A producer creates a version of the good of quality q at cost $c(q)$, assumed to be increasing and convex and measured in the same units as revenues.[17] Then he can supply an arbitrary quantity of version q as well as all versions dominated by q. Formally, they operate the following production set[18]

$$Y = \{\mathbb{I}\{q' \leq q\}, -c(q)\}_{q \in Q} \tag{2.3}$$

[16] In section 2.4.1 we present less restrictive sufficient condition to preserve the structure of the result at the cost of lower tractability.

[17] Again, using the re-definition of the quality spectrum from footnote 15, the cost reads $c_1(q) = c\left(g_2^{-1}(q)\right)$, and the substantive assumption that $c \circ g_2^{-1}$ is convex is guaranteed since it is a composition of convex functions.

[18] Consumers are in unit measure and demand at most one version, so a supply of 1 is indeed "arbitrarily large".

After producing q the seller has to quote a feasible market, that is a pricing function on the restricted domain $\boldsymbol{p}:[0,q]\to\mathbb{R}$.[19] The profit maximization problem therefore reads

$$\max_{q,\boldsymbol{p}:[0,q]\to\mathbb{R}}\int_{\Theta}\boldsymbol{p}\left(D_{\boldsymbol{p}}\left(\theta\right)\right)f\left(\theta\right)\mathrm{d}\theta-c\left(q\right)$$

2.1.3 First Best and Perfect Information

We begin by stating and solving the first best problem of choosing a social allocation to maximize expected social utility net of acquisition costs.

Definition 2.1. The **efficient allocation** is the function$\rho^{eff}:\Theta\to Q$

that solves

$$\max_{\rho:\Theta\to Q}S\left(\rho\right)=\int_{\Theta}u\left(\rho\left(\theta\right),\theta\right)\mathrm{d}F\left(\theta\right)-c\left(\sup_{\theta}\rho\left(\theta\right)\right) \tag{2.4}$$

The following proposition characterizes the efficient allocation

Proposition 2.2. *The efficient allocation map is given by*

$$\begin{array}{cccc}\rho^{eff} & : & \Theta & \to & Q\\ & & \rho^{eff}\left(\theta\right) & \mapsto & q^{\star}\end{array}$$

where $q^{\star}$ is the unique solution to equation

$$g'\left(q\right)+\mathbb{E}_{F}\left[\theta\right]=c'\left(q\right) \tag{2.5}$$

It is natural that the efficient allocation has singleton image $q^{\star}$: since each individual's utility is increasing in q and the distribution of each quality below q costs the same, it will never be socially optimal to allocate a damaged good to any type. Equation (2.5) is the first order condition of problem (2.4) after noticing that the efficient allocation is flat; sufficiency is immediate. As is also standard, we notice that a monopolist that is not subject to information frictions, namely that observes each type and can charge different prices to different costumers will induce the efficient allocation rule ρ^{eff} and extract all the surplus.

Remark 2.3. (First Degree Price Discrimination) Suppose that the individual type θ were observable to the seller. Then the monopolist would replicate the efficient allocation characterized in Proposition 2.2, and make profits $S\left(\rho^{eff}\right)$.

2.2 Private information

In the remainder of the paper, we assume consumers have private information about their utility type. The monopolist must rely on incentive compatible market design to allocate different qualities to different types.[20]

[19] An equivalent restatement is to let the seller choose only a pricing function $\boldsymbol{p}:Q\to\mathbb{R}\cup\{\infty\}$ and define the cost function $\bar{c}$ on the space of pricing functions as:

$$\bar{c}\left(\boldsymbol{p}\right)=c\left(\sup\left\{q:\boldsymbol{p}\left(q\right)<\infty\right\}\right)$$

In this case, the monopolist would solve

$$\max_{\boldsymbol{p}:Q\to\mathbb{R}}\int_{\Theta}\boldsymbol{p}\left(D_{\boldsymbol{p}}\left(\theta\right)\right)f\left(\theta\right)\mathrm{d}\theta-\bar{c}\left(\boldsymbol{p}\right)$$

[20] It is assumed that the principal has no "screening device" as defined in Jaynes (2006), namely he cannot obtain additional information about valuation types so that each individual must be treated as a random draw from F (assumed known).

We set up the problem as a multi-agent mechanism design problem and appeal to the revelation principle to write the monopolist problem as choosing a pair of allocation and transfer rules[21]

$$(\rho, p) : \Theta \to Q \times \mathbb{R}$$

to maximize profits under incentive compatibility and rationality constraints. Compared to standard screening problems, the key novelty is that the cost of an allocation rule ρ no longer takes the additively separable form

$$\overline{c}(\rho) = \int_{\Theta} \widetilde{c}(\rho(\theta))\, \mathrm{d}\theta$$

for some primitive cost $\widetilde{c}$. By contrast, it solely depends on one statistic of the allocation rule, namely the maximum quality, so it can be written as:[22]

$$\overline{c}(\rho) = c\left(\sup_{\theta} \rho(\theta)\right) \tag{2.6}$$

With these observations at hand, we write the monopolist problem in the following way

$$\begin{array}{lcc} \max_{\rho,p:\Theta\to Q\times\mathbb{R}} & & \int_{\Theta} p(\theta)\, \mathrm{d}F(\theta) - c(\sup_{\theta} \rho(\theta)) \\ & \text{s.t.} & \\ & \text{IC} & u(\theta, \rho(\theta)) - p(\theta) \geq u(\theta, \rho(\theta')) - p(\theta') \;\; \forall \theta, \theta' \\ & \text{IR} & u(\theta, \rho(\theta)) - p(\theta) \geq 0 \; \forall \theta \end{array} \tag{2.7}$$

It is worth emphasizing that under the non additively separable cost function (2.6), we cannot solve (2.7) by piecewise maximization of an appropriately defined type dependent profit. However, the simple form of non-separability characterizing (2.6) suggests that problem (2.7) can be divided in two stages. First, revenues conditional on each quality cap are calculated, and then those revenues are compared with the cost of buying a quality cap. The following Lemma formalizes this intuition and introduces the revenue and cap-conditional allocation functions.

Lemma 2.4. *Define the quality constrained revenue function $V : Q \to \mathbb{R}$ given by*

$$\begin{array}{rcc} V(q) & \longmapsto & \max_{\rho,p:\Theta\to Q\times\mathbb{R}} \int_{\Theta} p(\theta)\, \mathrm{d}F(\theta) \\ & & IC\ ,\ IR \\ & & \rho(\theta) \leq q, \qquad \forall \theta \in \Theta \end{array} \tag{2.8}$$

and let $\rho_q : \Theta \to Q$ be the optimal quality allocation of problem (2.8). The solution to problem (2.7) is

Bergemann et al. (2015) analyze the limits of (third-degree) price discrimination induced by additional information on buyers' types and show that that information can be tailored to achieve any combination of surplus such that total surplus is below the efficient level, and producer and consumer surplus are above uniform monopoly pricing and zero, respectively.

[21] The steps for rewriting the monopolist problem presented in the previous section as the design of a direct mechanism are standard and therefore omitted. It should not create confusion that from now on the pricing function p has domain the type space Θ rather than the quality space Q.

[22] In a companion project (joint with Franz Ostrizek) we explore screening under the more general cost structure

$$\overline{c}(\rho) = \int_{\Theta} \widetilde{c}(\rho(\theta), \rho)\, \mathrm{d}\theta$$

that allows the cost of producing the quality sold to type θ to depend on the whole allocation rule. This is useful to describe less extreme versions of economies of scale, learning, or nontrivial cost of quality replication and versioning.

characterized by a quality cap q^M given by:

$$q^M = \arg\max_q V(q) - c(q) \tag{2.9}$$

and by an allocation

$$\rho^\star : \Theta \to Q = \rho_{q^M}(\theta)$$

Once we are given the V function from (2.8), it is clear we can solve (2.9) as a simple maximization in single variable (if we show V is concave, q^M will be characterized by a first order condition alone). The challenge is then to find function V and the cap-conditional allocation rule

$$\rho : Q \times \Theta \to Q$$

where $\rho(q, \theta)$ is the quality assigned to type θ when the quality cap is q.[23] This is the objective of the following section.

2.3 Solution of the Screening Problem

Characterizing the constraint-conditional optimal contract (and therefore the revenues) is in principle a complicated problem, since for each $q \in Q$ we need to solve for a function ρ_q. The main result of this section, Proposition 2.9, shows that whenever primitives satisfy regularity conditions, the set of constraint-conditional allocation rules take a simple form in which we just compare a function of the type with the constraint. Before moving to the general case with a continuous type space, we consider a simpler economy populated by only two types since it provides the intuition for the more general case and clarifies the nature of the regularity assumptions.

2.3.1 Two Types Example

Suppose the economy is populated by a fraction π of high marginal valuation types denoted by H and $1 - \pi$ low marginal valuation types L. The following proposition characterizes the optimal contract $V(q)$, $\rho_q(\cdot)$ for this economy.

Proposition 2.5. *Let $y^\star \in Q$ be the solution to*

$$\frac{g'(y) + \theta_L}{g'(y) + \theta_H} = \pi \tag{2.10}$$

The optimal allocation takes the simple form

$$\rho(L, q) = \begin{cases} q & if\ q < y^\star \\ y^\star & else \end{cases}, \qquad \rho(H, q) = q$$

firm's profits are given by

$$V(q) = \begin{cases} u(q, L) & if\ q < y^\star \\ u(y^\star, L) + \pi\left(u(H, q) - u(H, y^\star)\right) & else \end{cases}$$

[23] The policy function $\rho_q : \Theta \to Q$ associated to the quality constrained problem (2.8) is the section at q of the above defined ρ; in this sense my notation is consistent and I will use $\rho(q, \cdot)$ and $\rho_q(\cdot)$ interchangeably.

So that

$$V'(q)=\begin{cases}u'(q,L) & if\ q<y^{\star}\\ \pi u'(H,q) & else\end{cases}=c'(q)$$

characterizes per Lemma 2.4 the optimal quality cap.

Proposition 2.5 suggest how one may construct all cap-dependent contracts: using only the demand primitives of the model, i.e. the curvature of g and the distribution of types, we determine the point $y^{\star}$.[24] Then, one obtains the optimal allocation $\rho(\theta,q)$ by choosing the minimum between a *type dependent threshold* (in this case, $y^{\star}$ for low types, ∞ for high types) and the *quality cap* itself. By solving a single equation, (2.10), we can characterize the allocations for each quality cap and then infer transfer from the binding incentive constraints to construct the revenue and marginal revenue function. Finally, expression (2.10) suggests why a strictly concave g function is needed for a nontrivial allocation rule to realize. The following Remark formalizes this.

Remark 2.6. (Linear Utility). Under the linear utility specification equation (2.10) would read

$$\pi=\frac{\theta_L}{\theta_H}$$

which is not a function of q and cannot determine a threshold $y^{\star}$. In this case the monopolist would either always (i.e. at every quality cap) sell to both types, if $\theta_L>\theta_H\pi$, or always serve only high types. Notice expression (2.10) would also be trivial if we assumed any form multiplicatively separable utilities $u(q,\theta)=u_1(q)\cdot u_2(\theta)$. The fact that linear utility has no "service margin" as a function of quality will be preserved in the more general setting.

2.3.2 Continuous types

We now assume that the type space is continuous.

Definition 2.7. We say primitives are **regular** if the utility function takes the form (2.1) and the type distribution F has a monotonically increasing hazard rate

$$h(\theta)=\frac{1-F(\theta)}{f(\theta)}$$

Now consider the virtual valuation function

$$vv(\theta,q)=u(q,\theta)-h(\theta)\,u_{\theta}(q,\theta)$$

and define the correspondence of maximizers of the virtual valuation

$$\beta:\Theta\rightrightarrows Q\cup\{\infty\}$$

$$\beta(\theta)\longmapsto\arg\max vv(\theta,q)$$

where the abuse $\arg\max vv(\theta,q)=\infty$ is adopted whenever $vv(\theta,q)$ is strictly increasing in $q\in Q$.

[24] As $\lim_{q\to 0}h(q)=1$, $\lim_{q\to 0}h(q)=\frac{\theta_L}{\theta_H}$, then existence of threshold $y^{\star}$ requires that $\frac{\theta_L}{\theta_H}$ is not larger than π. Uniqueness follows from monotonicity of ratio of marginal utilities

The following Lemma, which follows from standard application of supermodular comparative statics, will be used in the proof of the main Proposition of this section.

Lemma 2.8. *If primitives are regular, then β is single valued, monotonically increasing and it is equal to ∞ on a set of positive measure $\left[\widetilde{\theta}, \overline{\theta}\right]$, where $\widetilde{\theta}$ is the unique solution to*

$$\theta - h(\theta) = 0$$

The general expression of β is given by

$$\beta(\theta) = \begin{cases} (g')^{-1}(h(\theta) - \theta) & \theta < \widetilde{\theta} \\ \infty & \theta \geq \widetilde{\theta} \end{cases}$$

Since g' is decreasing by assumption and so is $\theta - h(\theta)$, then β is an increasing continuous function that asymptotes to ∞ in the interior of Θ. Also, by the Inada condition and $h(\underline{\theta}) - \underline{\theta} > 0$ we obtain $\beta(0) > 0$, which will deliver important implications for the shape of the cap-contingent optimal contract. The graph below plots an example of β.

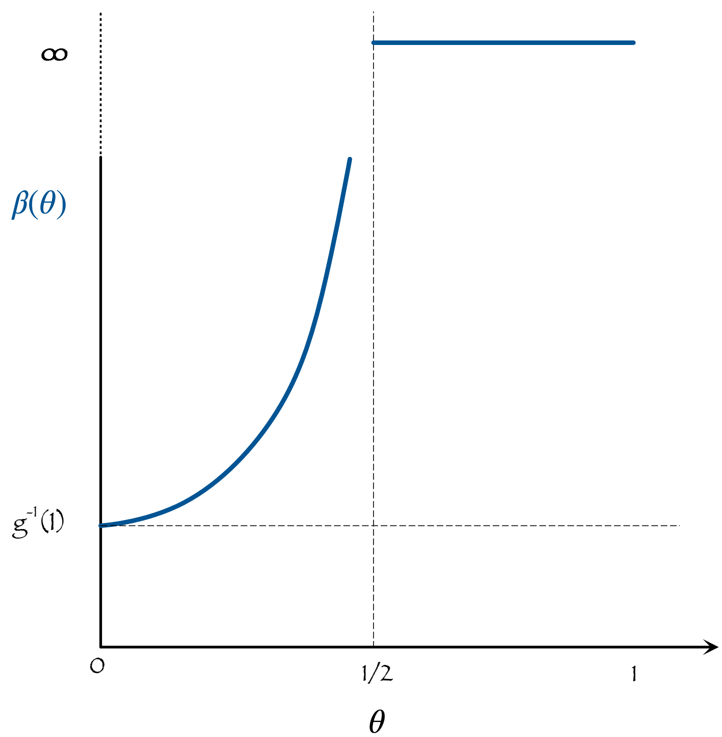

Figure 2.1: β function for concave preferences.

The threshold $\widetilde{\theta}$ takes value $\frac{1}{2}$ in the plot since in the remainder of this section we will assume $\boldsymbol{\theta} \sim \mathcal{U}[0,1]$. As Section 2.4.1 formalizes, qualitative results are unchanged though explicit formulas for revenues, surplus and inefficiencies would need to carry an additional transformation of the inverse hazard rate.[25] Under the

[25]We do not have a clear interpretation for the distribution of types, as their empirical content is not independent of the valuation function (2.1).

uniform distribution assumption, $h(\theta)-\theta = 1-2\theta$, $\beta(0) = (g')^{-1}(1) > 0$, and $\widetilde{\theta} = \frac{1}{2}$. We now proceed to state the proposition characterizing the monopolist allocation.

Theorem 2.9. *Suppose primitives are regular and types are uniformly distributed. Then,*

i) The ***quality-contingent*** *optimal contract takes the simple form*

$$\rho(q,\theta) = \min\{q, \beta(\theta)\} \tag{2.11}$$

ii) The revenue function V is concave with ***continuously differentiable*** *derivative given by*

$$V'(q) = \begin{cases} g'(q) & g'(q) > 1 \\ \left(\frac{1+g'(q)}{2}\right)^2 & g'(q) \leq 1 \end{cases} \tag{2.12}$$

iii) The monopolist quantity q^M is always strictly below $q^\star$.

The allocation rule (2.11) proves that the intuition in the two type example extends to regular continuous type spaces: the cap-contingent allocation of each type is simply determined by comparing a type dependent function (found at an ex-ante, constraint-free stage from Lemma 2.8), with the quality cap itself. It is only convenient to screen type θ if the quality cap exceeds $\beta(\theta)$, and when this happens the allocation $\rho(q,\theta)$ becomes unresponsive to further cap increments.

The shape of the optimal contract(s) helps interpreting also expression (2.12), i.e., the marginal revenue function. To understand this result, and for future reference, it is convenient to define $b : Q \to \left[0, \frac{1}{2}\right]$ as the inverse β function that returns for each quality level the lowest type that is bunched at the top. Under the uniform distribution we get

$$b(q) = \beta^{-1}(q) = \max\left\{0, \frac{1-g'(q)}{2}\right\} \tag{2.13}$$

At an intermediate step of the derivation of the marginal revenue we get

$$V'(q) = (1-b(q))\left[g'(q) + b(q)\right]$$

This expression has an intuitive interpretation: the term $g'(q)+b(q)$ is the marginal utility of the "marginally bunched" type $b(q)$, while $(1-b(q))$ is the mass of types above him.[26] By marginally increasing the quality cap, the monopolist does not alter the revenues made from optimal allocation of all lower qualities (given to the same types at the same price). He allocates the marginal quality to type $b(q)$ and to all those above, who increase their marginal transfer by

$$u_q(q, b(q)) = g'(q) + b(q)$$

Point *iii*) states that the top quality distributed by a monopolist is below the efficient level implied by (2.5). Although not obvious in this setting, the result suggests a natural parallelism with the classic underprovision of a good in the presence of market power. The "efficiency at the top" result, which is typical of quality screening problems, does not hold in this framework. Despite the highest type never receiving a damaged quality,[27] the quality cap produced under monopoly is below the level $q^\star$ that type, and everyone else, receives in the first best.

[26]Notice that when $b(q) = 0$, then $V'(q) = g'(q)$ delivering the first branch in (2.12). Substitution of the nontrivial expression for b delivers the second branch.

[27]So that *distributional* efficiency at the top realizes in our setting.

Comparing i) and iii) in Theorem 2.9 with Proposition 2.2 we notice that a monopolist induces two sources of inefficiency: one from damaging, because generically, $\rho(q,\theta) < q$ for a set o positive measure, and one from suboptimal acquisition. Although associated to different stages of the monopolist problem, these inefficiencies are to some degree interdependent as (1) the screening allocation is clearly constrained by the quality acquired and (2) the quality acquired depends on the screening possibilities. Expression (2.12) incorporates the optimal distribution of each maximal quality, which generally entails damaging. A graphic representation of the inefficiencies is given in the bottom panel of Figure 2.2, where green represents underacquisition and orange damaging. Their analytical expression will be derived in the next section.

2.4 Properties of the Monopolist Contract

We now list some properties of the optimal contract which follow immediately from Theorem (2.9). We begin with a description of the resulting quality allocations, separating the case of strictly concave and trivial g.

Corollary 2.10. *(Linear utility) Suppose $g \equiv 0$. Then, optimal allocations are given by:*

$$\rho(q,\theta) = \begin{cases} 0 & \theta \leq \frac{1}{2} \\ q & \theta > \frac{1}{2} \end{cases}$$

and the monopolist has constant marginal revenue $V'(q) = \frac{1}{4}$.

The result follows by applying allocation rule (2.11) to the maximizer of the virtual valuation, which in case of linear preferences is:[28]

$$\beta(\theta) = \begin{cases} 0 & \theta < \frac{1}{2} \\ \infty & \theta > \frac{1}{2} \end{cases}$$

When preferences are linear, screening is only performed by excluding (selling quality 0 to) an invariant set of types, so that only one positive quality will be offered. All the non-trivial screening observed in price discrimination with linear preferences is driven by the shape of the marginal cost curve (and from ironing a non-monotone hazard rate). Since the distribution problem lacks such cost curvature, a smooth screening contract must result from the specification of preferences. Allocations with concave g have the following properties

Corollary 2.11. *Suppose g is strictly concave. Then*

i) All types receive a positive quality in the optimal contract.

ii) At low quality caps the optimal contract features full bunching:[29]

$$\rho(q,\theta) = q \quad \forall\theta, q < (g')^{-1}(1)$$

[28] As in Remark 2.6, the shape of the optimal contract would be the same for any multiplicatively separable utility specification

$$u(q,\theta) = u_1(q) \cdot u_2(\theta)$$

which gives:

$$vv(q,\theta) = u_1(q) \cdot \left[u_2(\theta) - u_2'(\theta)\left(\frac{1-F(\theta)}{f(\theta)}\right)\right]$$

and $\beta(\theta)$ is again either 0 or ∞.

[29] With non-uniform distribution, the requirement is $q < (g')^{-1}(h(\underline{\theta}) - \underline{\theta})$

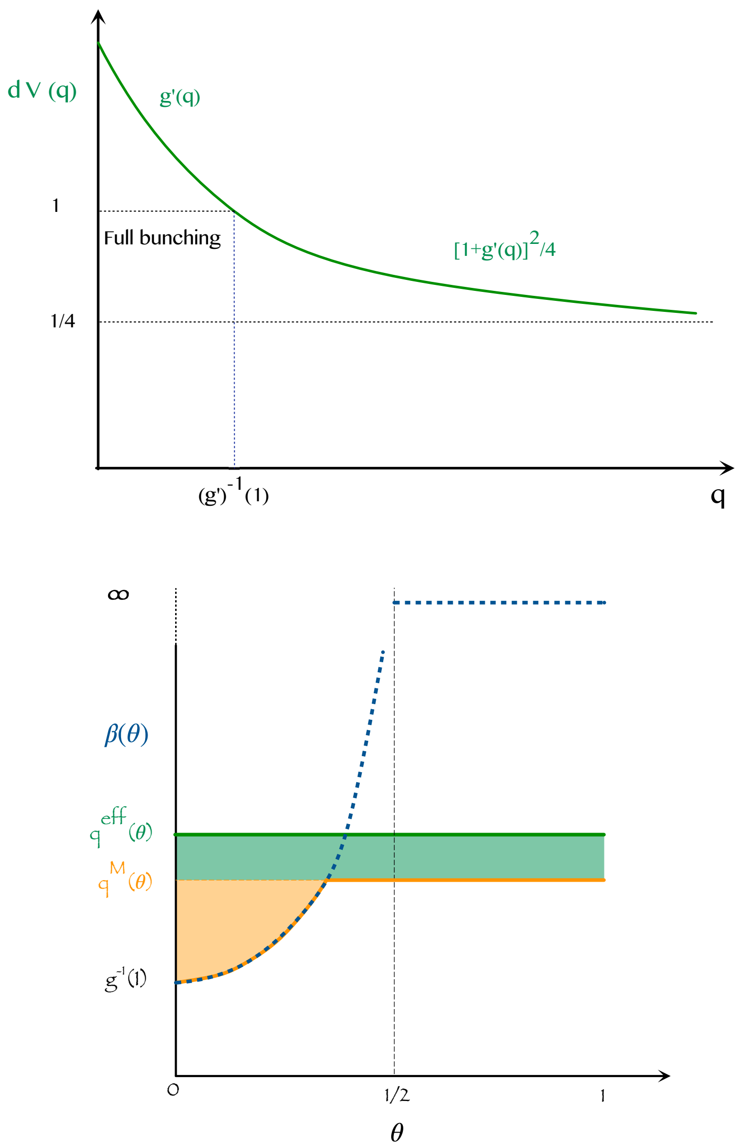

Figure 2.2: Marginal revenue (top) and the two inefficiencies (bottom) from monopolistic provision of digital goods.

iii) A positive measure of agents $\left[\frac{1}{2},1\right]$ *receives the highest quality good irrespectively of the quality cap.*[30]

$$\rho\left(q,\theta\right)=q\quad\forall\theta\in\left[\frac{1}{2},1\right]$$

Point $i)$ is implied by the Inada condition of g around 0: by giving a marginal quality to low valuation types who receive nothing the seller gets unbounded marginal revenues, which she can distribute as information rents to make sure IC constraints for higher types are satisfied. If the quality acquired is low enough, since there is little variation in *relative* marginal utilities it will not be optimal to screen any type, giving point $ii)$. An implication of $ii)$ is that sufficiently steep marginal cost shuts down one source of inefficiency, damaging, leaving only underacquisition active. This will allow us to isolate the impact of a policy on the underacquisition inefficiency alone by assuming monopolist was producing in this region. By point $iii)$, there is a set of types that are bunched at the top irrespectively of the quality cap. It should be noticed that those types are exactly the same that were sold a positive undamaged quality under linear preferences. Indeed, a concave g does not change the fact that above a certain threshold the virtual valuation is monotonically increasing, but it gives a nontrivial maximizer for other types. Also, we can now explain why (2.12) gives limit marginal revenues

$$\lim_{q\to\infty}V'\left(q\right)=\frac{1}{4}$$

As the quality grows higher, marginal increments are distributed as if preferences were linear since at high qualities agents use incremental units only in the performance of the advanced task.[31]

Figure 2.3 shows a graphical derivation of the optimal contract: the marginal revenue is crossed with the marginal cost function to determine q^M (left graphs with axes flipped for convenience), this level is reported on the vertical axis in the graph of β and the optimal allocation $\rho^\star\left(\cdot\right)$ then results from "slicing" β at q^M. The three panels describe full bunching, active screening and linear preferences.

We now compute total surplus under monopoly and give an analytic expression to the two sources of inefficiency described above.[32] If q^M is below $\left(g'\right)^{-1}\left(1\right)$, then computations are simple. Producer revenues are $V\left(q\right)=g\left(q\right)$, consumer surplus is $W\left(q\right)=\frac{1}{2}q$ and per Corollary 2.11 $i)$, there are no damaging inefficiencies (only underacquisition is active). The following proposition characterizes surpluses when instead $q^M>\left(g'\right)^{-1}\left(1\right)$.

Proposition 2.12. *(Decomposition of inefficiencies above* $\left(g'\right)^{-1}\left(1\right)$*)*

Total consumer surplus is given by

$$W\left(q^M\right)=\frac{1}{2}\left[\left(g'\right)^{-1}\left(1\right)+\int_{\left(g'\right)^{-1}\left(1\right)}^{q^M}\left(\frac{1+g'\left(q\right)}{2}\right)^2\mathrm{d}q\right]\tag{2.14}$$

We can therefore decompose monopolist inefficiencies as

$$\int_{\left(g'\right)^{-1}\left(1\right)}^{q^M}d\left(q\right)\mathrm{d}q+\int_{q^M}^{q^\star}\left[\frac{1}{2}+g'\left(q\right)-c'\left(q\right)\right]\mathrm{d}q\tag{2.15}$$

[30] Similarly, we would get the unconditional bunching region to be $\left[\widetilde{\theta},\overline{\theta}\right]$.

[31] This also means that to have a finite solution to the monopolist problem the marginal cost must have limit that exceeds $\frac{1}{4}$.

[32] We are currently missing an expression for consumers' rent.

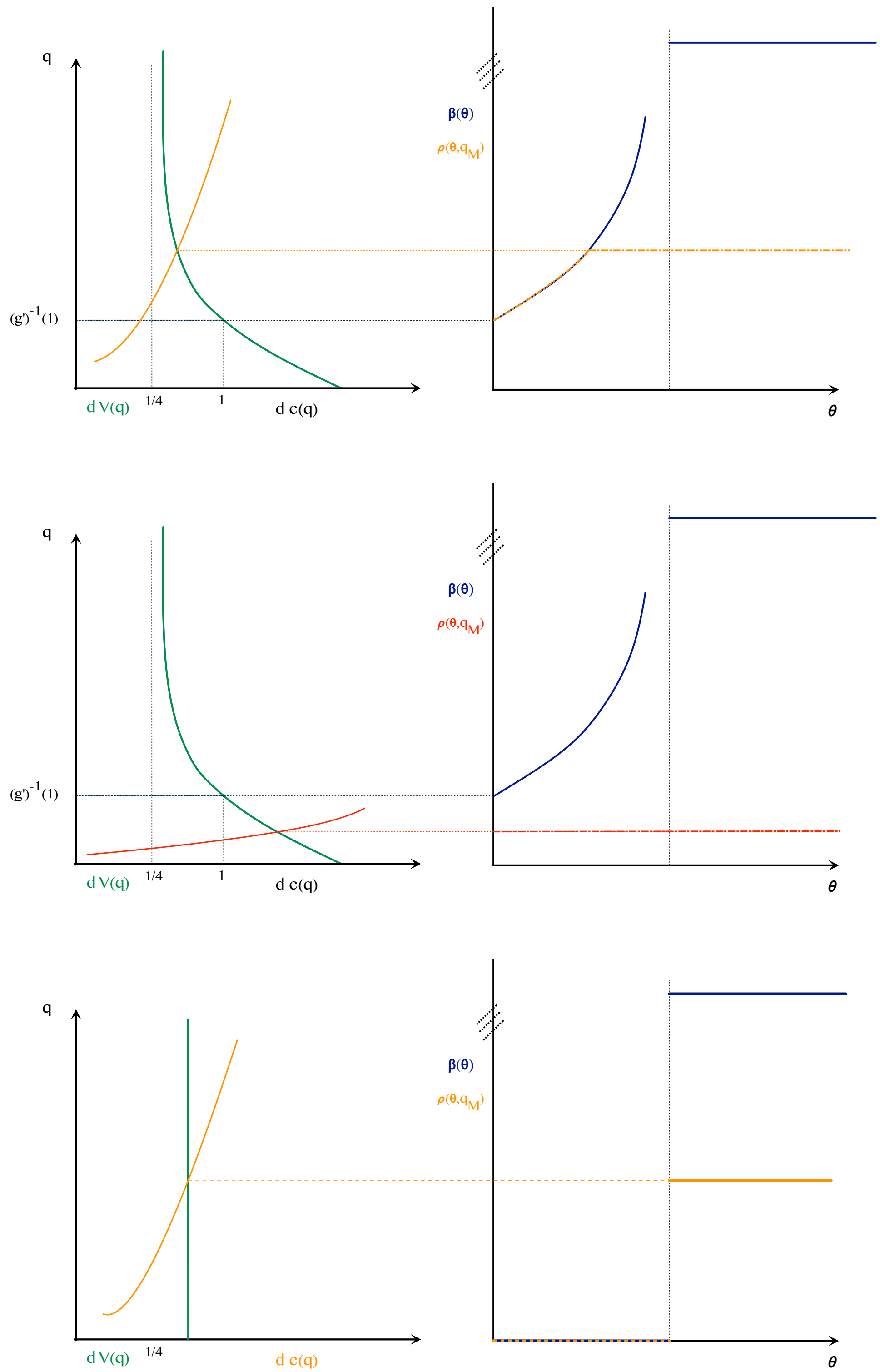

Figure 2.3: Graphical representation of optimal contract with high (top) and low (medium) marginal cost for concave g, and for linear preferences (bottom).

where
$$d(q) = \frac{1}{8}\left(1 + \left(2 - 3g'(q)\right)g'(q)\right) \tag{2.16}$$
are the marginal inefficiencies from damaging.

Analytic manipulation, shows that aggregate marginal information rent is given by

$$W'(q) = \frac{1}{2}\left(\frac{1+g'(q)}{2}\right)^2 = \frac{1}{2}V'(q)$$

Hence, total surplus under monopoly grows, above $(g')^{-1}(1)$ with slope given by

$$\underbrace{\left(\frac{1+g'(q)}{2}\right)^2 - c'(q)}_{\text{marginal profits}} + \underbrace{\frac{1}{2}\left(\frac{1+g'(q)}{2}\right)^2}_{\text{marginal rents}} = \frac{3}{2}\left(\frac{1+g'(q)}{2}\right)^2 - c'(q)$$

Which, combined with full bunching below $(g')^{-1}(1)$ gives the following expression for marginal monopolist surplus

$$m(q) = \begin{cases} g'(q) + \frac{1}{2} - c'(q) & g'(q) > 1 \\ \frac{3}{2}\left(\frac{1+g'(q)}{2}\right)^2 - c'(q) & g'(q) \leq 1 \end{cases} \tag{2.17}$$

If we didn't have the inefficient damaging, total surplus would grow with slope

$$\frac{1}{2} + g'(q) - c'(q) \tag{2.18}$$

Subtracting the two we get the expression for the marginal inefficiencies from damaging

$$d(q) = \frac{1}{8}\left(1 + \left(2 - 3g'(q)\right)g'(q)\right)$$

by integrating $d(q)$ from $(g')^{-1}(1)$ to q^M we get the first term in (2.15) (the area of the orange region in Figure 2.2). While we cannot give an intuitive interpretation as to why (2.16) represents marginal inefficiencies from damaging, because the analytical result depends on the uniformity assumption, we notice that $d(q) = 0$ when $g'(q) = 1$ (below that level there were no damaging inefficiencies), and that d is a positive hump shaped function in $g'(q)$ when smaller than 1. Underacquisition inefficiencies, the green area constitute instead the second summand in (2.15) and add to social surplus losses (2.18) in the underacquisition region $\left[q^M, q^\star\right]$.

2.4.1 Relaxing Demand Primitives

We have made restrictive assumptions about the specification of returns from quality and of the distribution of types. Returns belong to the family (2.1) and the type distribution is assumed uniform. The latter assumption is innocuous: the qualitative results of allocation rule (2.11), monopolist underprovision and decomposition of inefficiencies into damaging and underprovision always hold though the analytic expression of marginal revenues (2.12), consumer surplus and damaging inefficiencies (2.16) are modified to allow for a different inverse hazard rate. Under a general distribution characterized by hazard rate h, the full bunching threshold $(g')^{-1}(1)$ would be
$$\overline{q} = (g')^{-1}\left(\underline{\theta} - h\left(\underline{\theta}\right)\right)$$

where $\overline{q} > 0$ is guaranteed by the monotone hazard rate assumption and the Inada condition $\lim_{x\to 0} g'(x) = \infty$. Similarly, the inverse β function used to calculate the marginal type bunched at the top which gives marginal revenues and welfare is defined implicitly by the equation:

$$g'(q) = (h(b(q)) - b(q)) = \widetilde{h}(b(q))$$

In general, the b function would be given by

$$b(\cdot) = \max\left\{0, \widetilde{h}^{-1} \circ g'(\cdot)\right\}$$

which reduces to (2.13) in the uniform case as$\widetilde{h}^{-1}(x) = {}^{(1-x)}/_{2}$.The cutoff type assigned the undamaged quality when preferences are linear will be $\widetilde{\theta}$, the zero of $\widetilde{h}(\theta) = \theta - h(\theta)$,[33] and marginal revenues would be $\widetilde{\theta}\left[1 - F\left(\widetilde{\theta}\right)\right]$, which are $\frac{1}{2}$ and $\frac{1}{4}$ respectively in the uniform case.

The additive separability assumption is more substantial; its empirical content is discussed in the introduction, together with a plausible microfoundation. The Inada conditions are essential to get the full bunching region at low qualities, while identifying preferences with function g allows to characterize revenues and welfare only as a function of its curvature. For the bunching at the top property of the cap-contingent contract (Proposition 2.9, i)) it would be sufficient that β is an increasing function. This would be guaranteed by concavity in q and supermodularity of the virtual valuation which is in turns ensured by the standard conditions on mixed derivatives

$$u_{qq} < 0, \qquad u_{q\theta} > 0, \qquad u_{q\theta\theta} \leq 0$$

and a monotone hazard rate.

2.5 Impact of a No-Screening policy

Corollaries 2.11 ii) and 2.10 present two cases in which only one positive quality is distributed under monopoly: either preferences are linear, a constant mass of agents is served the undamaged quality and others are excluded, or g is concave but the optimal quality cap is low enough (steep marginal cost) to induce full bunching. The aim of this section is to evaluate the positive and normative implications of a regulation that prohibits the monopolist from selling damaged goods. This exercise is useful for two reasons. First, it is the natural extension to this framework of the Deneckere and McAfee (1996) normative question “when is the possibility of screening beneficial for all types in the economy?”. Our different specification of the cost function and the fact that available qualities are not pre-determined add different channels through which the NS policy can impact allocations and welfare. Second, this is a first pass at evaluating the impact of a policy on the two inefficiency sources isolated in the previous section, and will provide a useful benchmark of comparison for the welfare implications of competition.

2.5.1 The NS Problem

The superscript NS will denote objects associated to a No Screening monopolist. He chooses a quality q^{NS} and the threshold consumer $\vartheta\left(q^{NS}\right)$ who is indifferent between purchasing the good and not, then sell q^{NS}

[33] That always coincides with the threshold type that is always bunched at the top even with concave concave g.

to types $\left[\vartheta\left(q^{NS}\right), 1\right]$ at price

$$g\left(q^{NS}\right) + \vartheta\left(q^{NS}\right) q^{NS}$$

The exclusion policy $\vartheta : Q \to \Theta$ is found by solving the quality-conditional pricing problem:[34]

$$\Pi^{NS}(q) = \max_{\theta} \left[u(q,\theta)\right](1-\theta)$$

At the acquisition stage, the NS monopolist solves

$$\max_{q} \Pi^{NS}(q) - c(q)$$

The following proposition characterizes the solution in regular environments. We focus on strictly concave g as we already argued that the NS policy is not binding when preferences are linear.

Proposition 2.13. *i) The exclusion policy ϑ is given by*

$$\vartheta(q) = \max\left\{\frac{q-g(q)}{2q}, 0\right\}$$

It holds $\vartheta(q) \leq b(q)$, strictly when $b(q) > 0$. Moreover, $\vartheta^{-1}(0) > \beta(0)$.

ii) Marginal revenues for the NS monopolist are given by

$$\frac{\mathrm{d}}{\mathrm{d}q}\Pi^{NS}(q) = \begin{cases} g'(q) & q \leq g(q) \\ \frac{(q+g(q))(q-g(q)+2qg'(q))}{4q^2} & q > g(q) \end{cases}$$

iii) It holds $q^{NS} \leq q^M$ (strictly whenever $q^M > (g')^{-1}(1)$).

The ϑ and b functions are plotted on the left panel of Figure 2.4. The fact that ϑ has a larger intercept, stated above as $\vartheta^{-1}(0) > \beta(0)$, is key as it implies that the NS monopolist starts excluding some types at a quality level at which the unconstrained monopolist is already actively engaging in inefficient damaging. The quality space is then partitioned in three regions, highlighted in the right panel of Figure 2.4, where marginal revenues for the constrained monopolist are compared with those of the unconstrained monopolist. In Region A, where $q < (g')^{-1}(1)$, both monopolists do full bunching, the constraint is immaterial, and marginal revenue functions coincide. In Region B, where $q < (g')^{-1}(1)$ but $q < g(q)$, the NS monopolist sells to all types (marginal revenue is $g'(q)$) while the unconstrained one does positive damaging. In Region C also the NS monopolist performs his "constrained" screening by excluding a positive mass of low types.

Point *iii)* immediately implies that the NS policy induces a deterioration in the acquisition inefficiency, which becomes strict as soon as the NS constraint becomes binding. As for the screening inefficiencies, we have two competing forces: as $\vartheta(q) < b(q)$, conditional on acquiring the same quality, the NS monopolist would serve it to a larger portion of types. However, types that are excluded in the NS contract receive nothing, while in the unconstrained monopolist contract they have positive consumption (and value). This delivers the immediate welfare implication that the NS policy makes some low types in Region C, i.e. those that are excluded, worse off. It is also clear that the monopolist is always worse off since she is solving a

[34]We keep the assumption $\boldsymbol{\theta} \sim \mathcal{U}[0,1]$. The qualitative properties do not change under generic monotone hazard rate distribution.

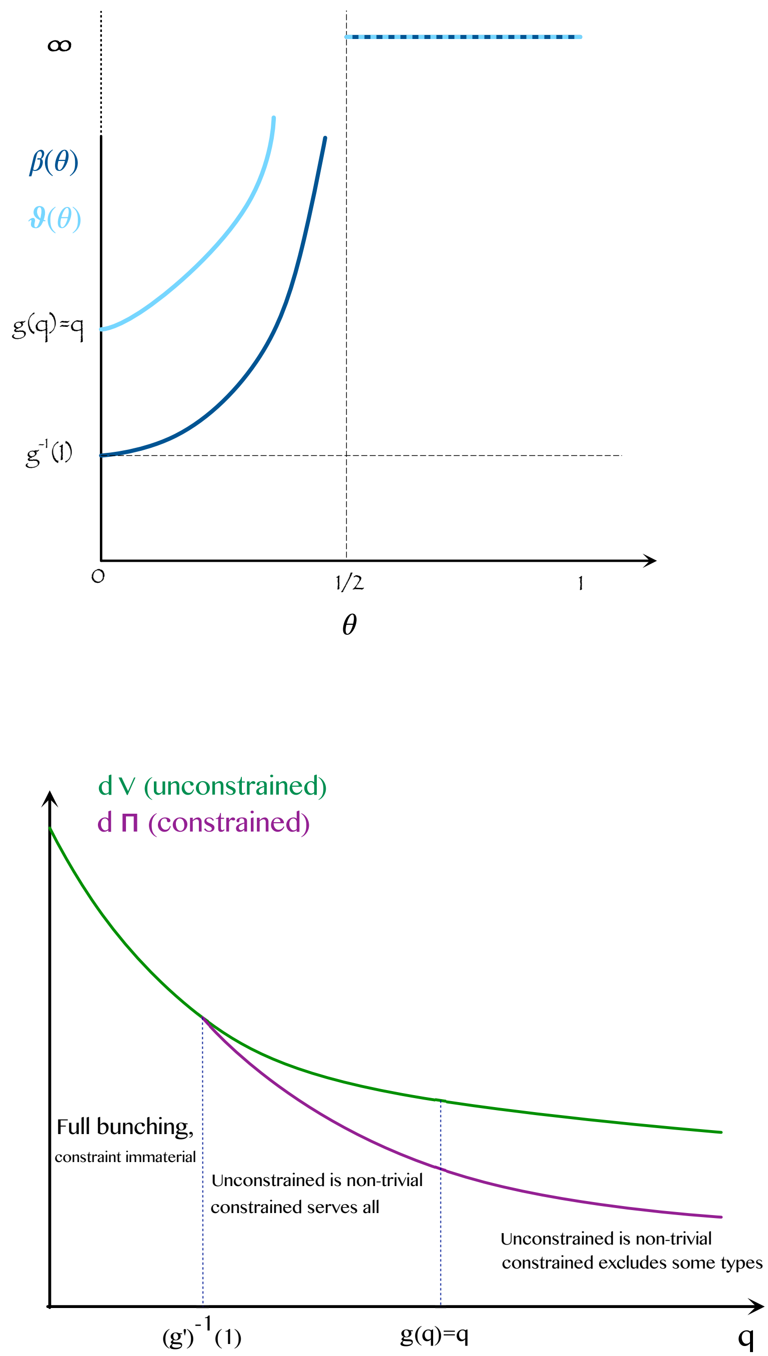

Figure 2.4: Constrained and unconstrained monopolist: full bunching thresholds (top) and marginal revenues (bottom).

constrained version of problem (2.7). We can write foregone profits from the NS policy as

$$V\left(q^{NS}\right)-\Pi\left(q^{NS}\right)+\int_{q^{NS}}^{q^{M}}\left[V'\left(q\right)-c'\left(q\right)\right]\mathrm{d}q$$

both summands are positive and they separate losses for not performing screening and from acquiring a quality that is below the unconstrained optimum. By focusing on cases where the unconstrained monopolist produces in Region B we can derive some less trivial welfare implications of the NS policy.

Proposition 2.14. ***(Welfare impact in Region B)*** *If the monopolist produces in Region B, then*

i) A set of (low) types is better-off under the NS policy.

ii) The net gains from enacting the NS policy can be expressed as

$$\int_{(g')^{-1}(1)}^{q^{NS}} d\left(q\right)\mathrm{d}q-\int_{q^{NS}}^{q^{M}} m\left(q\right)\mathrm{d}q \tag{2.19}$$

where $m\left(q\right)$ is the marginal monopolist surplus (2.17) and $d\left(q\right)$ are the marginal damaging inefficiencies (2.16).

iii) If $c''\left(q^{M}\right)$ is large enough (approaches the fixed cost limit), the NS policy increases total welfare (net gains approach $\int_{(g')^{-1}(1)}^{q^{M}} d\left(q\right)\mathrm{d}q$).

The intuition for $i)$ is as follows: because in Region B there is no exclusion, every type receives welfare $W^{NS}\left(\theta\right)=\theta q^{NS}$; in the unconstrained case low types, those below $b\left(q^{M}\right)$, received $W\left(\theta\right)=\int_{0}^{\theta}\beta\left(\theta'\right)\mathrm{d}\theta'$. As $q^{NS}>\beta\left(0\right)$, it follows the marginal surplus is larger in the NS environment for a set of positive measure, implying the statement. As for the overall impact of the NS policy, it reduces "by brute force" the damaging inefficiencies but it creates two perverse effects: it makes the monopolist worsen underacquisition, and it may force some types to be completely excluded. Focusing on Region B ensures the "complete exclusion" margin in non-existent, so point $ii)$ only trades off the positive impact from undoing damaging in the $\left[(g')^{-1}\left(1\right),q^{NS}\right]$ region, and the negative underacquisition impact. Assuming costs are extremely convex around q^{M} ensures marginal cost cover the $V'-\Pi'$ gap quickly implying $q^{NS}\rightarrow q^{M}$. So in the limit every type receives the quality a monopolist produces (but does not distribute), which increases total welfare by completely undoing damaging inefficiencies without perverse effects. The geometric intuition for the result is given in top panel of Figure 2.5, comparing acquisition and distribution in Region B for different cost functions.

In Region C, portrayed in the bottom panel of Figure 2.5, the convex cost limit would give an ambiguous welfare impact as we would need to take into account the loss of consumer surplus in the region $\left[0,\vartheta\left(q^{M}\right)\right]$: a full welfare comparison needs to take into account of consumer surplus that is lost by inducing complete exclusion in that Region.

3 Competition

This section develops a model of competition in digital goods markets. Following the discussion in the introduction, we specify an extensive form game and study its equilibria.[35] The discussion on stability in Section 3.4 offers a comparison with the equilibrium notion in Rothschild and Stiglitz (1976), which is a

[35] How the setup of the game relates to existing models of competition with screening won't be discussed throughout the exposition, relevant references are given in the dedicated paragraph of the literature review. Section 3.4 compares some properties of equilibria we obtain with those of models that are closer to ours.

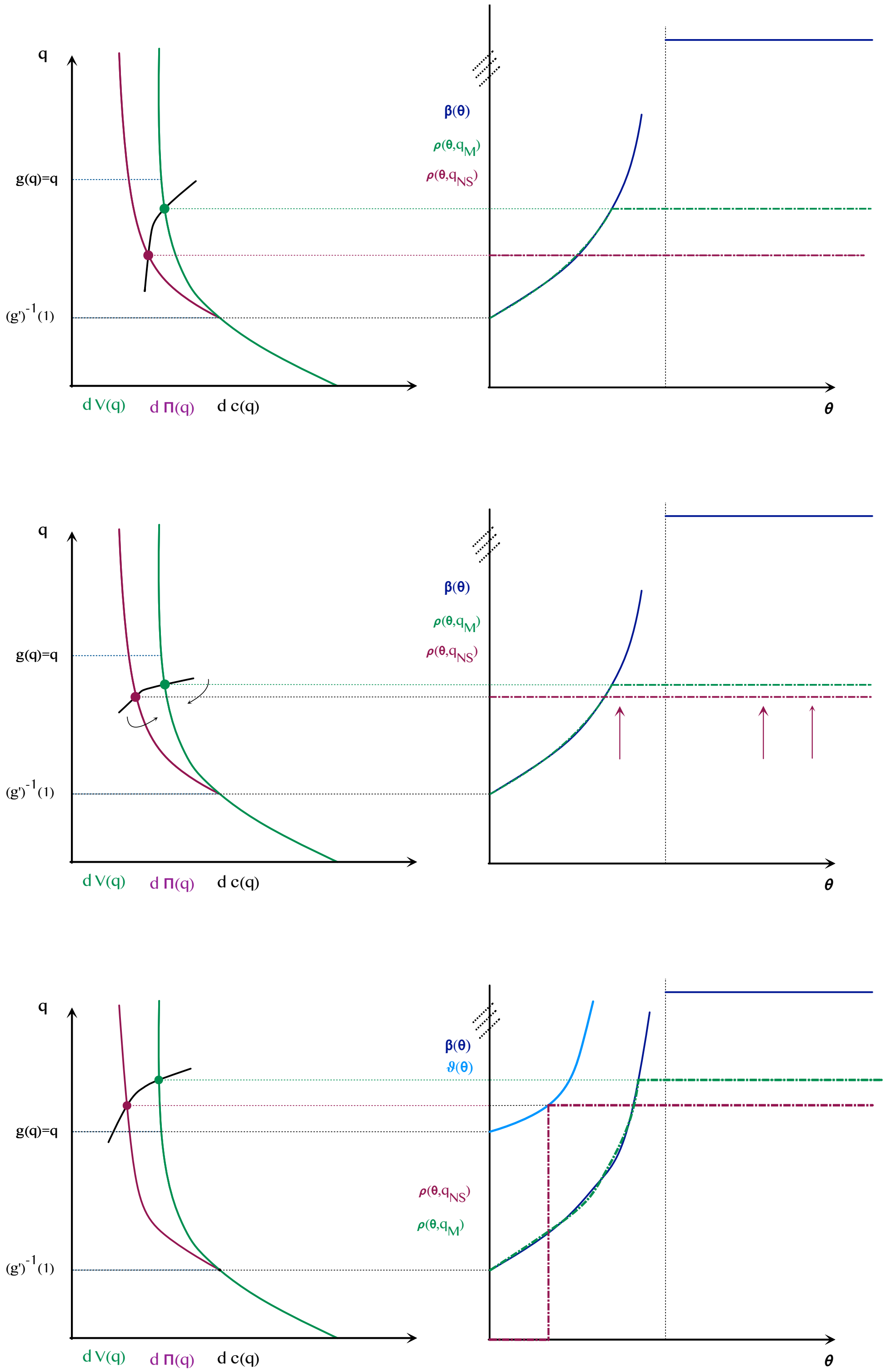

Figure 2.5: (Top panel) Unconstrained and NS allocations in Region B, cost moderately convex. (Medium) Region B, extremely convex cost. (Bottom) Unconstrained and NS allocations in Region C.

standard benchmark in the literature of competition with screening. Before that, we present the competition game, solve for its equilibria by backwards induction (pricing and investment stages), and analyze their welfare properties.

3.1 Primitives and the Extensive Form Game

The production primitives of the model are augmented by adding countably infinite replica of the producer studied in Section 2. The set of firms is denoted $\mathbb{N}$ with typical element i. On the demand side, we assume that firms produce a *homogeneous* (range of) products: the identity of the firm producing a certain quality is immaterial to consumers that will therefore construct their demand by looking at the lower envelope of the pricing functions. Firms play a two stage game of perfect information in which investments in quality (first stage, or acquisition stage) are based upon the belief that the ensuing price decision will constitute a Nash equilibrium in the second (or pricing) stage at which firms quote a feasible pricing function at treat production costs as already been sunk.

3.1.1 Timing and Action Space

At the first stage each firm chooses a mixed strategy over the qualities she acquires, so her action space is

$$A_{1,i} = \Delta\left(Q\right), \forall i$$

Firms pay the cost associated to the realized q in exchange to the right to sell (at the second stage) all qualities below q. At the end of stage one, the vector of realized qualities[36] $\boldsymbol{q} \in Q^{\mathbb{N}}$ becomes common knowledge across competitors. Conditioning on this information, firms then quote a market over the qualities available to them to maximize revenues. The action set

$$A_{i,2}\left(\boldsymbol{q}\right) = \mathbb{R}^{[0,\boldsymbol{q}_i]}$$

is the set of pricing functions on the feasible domain $[0, \boldsymbol{q}_i]$, $\boldsymbol{q}_i$ being the i^{th} entry of vector $\boldsymbol{q}$.[37] Payoffs from this stage are the revenues each firm makes as a consequence of everyone's pricing decision. A formal expression of the payoff function is given in Section 3.2.1.

Firms' strategies in the extensive form game consist then of a distribution over entry qualities and, conditional on each realized quality vector, a *feasible* pricing function. It is assumed that there is no discounting between periods[38] so total payoffs simply add costs incurred in the first stage and revenues earned in the second stage; when those are stochastic, firms behave as expected profit maximizers.

3.2 Equilibrium

This section studies subgame perfect equilibria of the finite horizon game described above by backwards induction.

[36]The bold notation reflects the fact that, from an ex-ante perspective, $\boldsymbol{q}$ is a random vector with distribution parametrized by equilibrium actions.

[37]In principle we can allow for randomization even at the pricing stage so that $A_{i,2}\left(\boldsymbol{q}\right) = \Delta\left(\mathbb{R}^{[0,\boldsymbol{q}_i]}\right)$; randomization will never be optimal in the second stage, so we restrict the action set to pure actions.

[38]As costs are incurred in the first period and profits are possibly earned in the second period. Re-normalizing costs to take discounting into account does not modify the analysis in any substantive way.

3.2.1 The Pricing Stage

The first step consists in specifying revenues as a function of players' actions, i.e. pricing functions. Intuitively, firms make revenues from the set of qualities they offer at the lowest price, individual demands being determined by the market pricing function. The following steps make this intuition formal.

We take as given the set of individual pricing functions $\{p_i\}_{i\in\mathbb{N}}$, $p_i : Q \to \mathbb{R}\cup\{\infty\}$.[39] The market pricing function returns the lower envelope

$$m\left(\{p_i\}_{i\in\mathbb{N}}\right) : Q \to \mathbb{R}\cup\{\infty\}$$

$$m\left(\{p_i\}_{i\in\mathbb{N}}\right)(q) \mapsto \min_i p_i(q)$$

which allows to derive the individual demand correspondence

$$D_{\{p_i\}_{i\in\mathbb{N}}}(\theta) = \arg\max_q u(q,\theta) - m\left(\{p_i\}_{i\in I}\right)(q)$$

To express revenues, we firstly associate each type to the firm he buys from

$$\iota(\theta) : \Theta \to \mathbb{N}$$

$$\iota(\theta) \mapsto \min\left\{\arg\min_i \left\{p_i\left(D_{\{p_i\}_{i\in I}}(\theta)\right)\right\}\right\}$$

Firm j earns revenues that depend on how much consumers demand, which is a function of the whole market, on the quality spectrum it ends up supplying by charging the lowest price:[40]

$$\overline{R}_j\left(\{p_i\}_{i\in I}\right) = \int_{\{\theta:\iota(\theta)=j\}} p_j\left(D_{\{p_i\}_{i\in I}}(\theta)\right) f(\theta)\,\mathrm{d}\theta \tag{3.1}$$

Now suppose $\boldsymbol{q}$ is the vector of realized qualities, and that it is common knowledge across players. As it is standard, the notation $\boldsymbol{q}^{(i)}$ denotes the i^{th} order statistic of vector $\boldsymbol{q}$. The following proposition characterizes the equilibrium of the pricing game.

Proposition 3.1. *If preferences are regular, for each $\boldsymbol{q}$ the second stage game has an essentially unique Nash equilibrium in pure strategies. The induced allocations are*

$$\rho(\boldsymbol{q},\theta) = \begin{cases} \boldsymbol{q}^{(2)} & \text{if } \beta(\theta) < \boldsymbol{q}^{(2)} \\ \beta(\theta) & \text{if } \boldsymbol{q}^{(2)} \leq \beta(\theta) < \boldsymbol{q}^{(1)} \\ \boldsymbol{q}^{(1)} & \text{if } \beta(\theta) \geq \boldsymbol{q}^{(1)} \end{cases} \tag{3.2}$$

Revenues are given by:

$$R_i(\boldsymbol{q}) = \max\left\{V(\boldsymbol{q}_i) - \max_{j\neq i} V\left(\boldsymbol{q}_j\right), 0\right\} \tag{3.3}$$

The intuition for the result is fairly simple: since at the pricing stage costs are sunk and the production realization is common knowledge, Bertrand competition will drive to zero the revenues from versions $\left[0,\boldsymbol{q}^{(2)}\right]$ that can be provided by more than one firm. Therefore, all firms make zero revenues, except for the owner of the highest quality, from now on referred to as "interim monopolist". She enjoys market power on the quality spectrum $\left[\boldsymbol{q}^{(2)},\boldsymbol{q}^{(1)}\right]$, and behaves as a monopolist under the additional constraint that all agents

[39] As before, we identify firm i not offering quality q by writing $p_i(q) = \infty$.
[40] ι is measurable under the assumption that firms quote an increasing function.

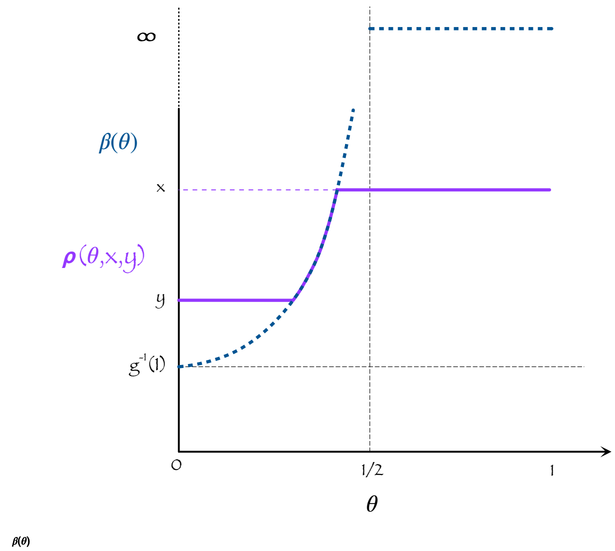

Figure 3.1: Equilibrium allocations in an x, y market.

must receive at least $\boldsymbol{q}^{(2)}$ for free. Given regularity, the solution to this problem is again simple: the β function is now "sliced" both from below and from above: all types below $b\left(\boldsymbol{q}^{(2)}\right)$ receive $\boldsymbol{q}^{(2)}$ for free, the others get the same allocation as under monopolist with quality $\boldsymbol{q}^{(1)}$ - but pay less. Figure 3.1 plots the allocation induced by competition with $\boldsymbol{q}^{(1)} = x$, $\boldsymbol{q}^{(2)} = y$. The equilibrium is only essentially unique because it is not pinned down who between the producer of the first and the second quality (or both of them) ends up distributing $\boldsymbol{q}^{(2)}$.[41]

By (3.3), each firm can compute its revenues as a function of its quality q and the best quality across competitors x, which a payoff sufficient summary of the competitive environment

$$R : Q^2 \to \mathbb{R}$$

$$R\left(q, x\right) \mapsto \max\left\{V\left(q\right) - V\left(x\right), 0\right\} \tag{3.4}$$

Towards the calculation of the first stage equilibrium it will be key that the *marginal* revenues of the interim monopolist only depend on her, and that they own quality and coincide with that of the unconstrained monopolist, namely

$$\frac{\partial}{\partial q} R\left(q, x\right) = \begin{cases} 0 & x > q \\ V'\left(q\right) & x \leq q \end{cases} \tag{3.5}$$

In order to avoid carrying order statistics notation, in the remainder of the paper we will denote with $\boldsymbol{x}$ be the best and $\boldsymbol{y}$ the second quality in the vector of realized entries; x, y will denote particular market realizations and $\rho\left(x, y, \theta\right)$ is the quality assigned to type θ whenever the realized first and second qualities are x, y respectively. The following example begins computation of the equilibrium in the case of linear preferences.

[41] This fact may have an empirical content, since we may observe multiple providers of the free quality, though it clearly has no implications on welfare.

Example 3.2. With linear preferences, the allocation and transfers associated to each pair x, y are given by

$$(\rho, p)(x, y, \theta) = \begin{cases} (y, 0) & \theta \in \left[0, \frac{1}{2}\right] \\ \left(x, \frac{1}{2}(x-y)\right) & \theta \in \left[\frac{1}{2}, 1\right] \end{cases} \tag{3.6}$$

Low types, who were previously excluded, receive the second quality y for free, while high types receive the best quality x and pay price $\frac{1}{2}(x-y)$. The revenues of the interim monopolist are $\frac{1}{4}(x-y)$, and marginal revenues are constant at $\frac{1}{4}$.

3.2.2 The First Stage Game

The pricing game delivers a revenue function (3.4) which is added to the acquisition cost, thus determining the payoff function in the quality investment game. The expected profit associated to quality q is

$$\Pi(q) = \int_Q R(q, x)\,\mathrm{d}H(x) - c(q) \tag{3.7}$$

where H is the CDF of the best quality produced in equilibrium by competitors. The following proposition characterizes equilibria of this game. A firm is called active if it plays an action different from δ_0, that is if it chooses a positive quality with positive probability.

Proposition 3.3. *The first stage game has a unique equilibrium for any number $n \geq 1$ number of active firms. With $n = 1$, the active firm plays δ_{q^M}; this is also the only equilibrium in pure strategies. For each $n \geq 2$, active firms play a mixed symmetric equilibrium*

i) with support $\left[0, q^M\right]$

ii) and continuously differentiable (on the interior of the support) CDF H_n given by

$$H_n(q) = \left[\frac{c'(q)}{V'(q)}\right]^{\frac{1}{n-1}} \tag{3.8}$$

and make zero (expected) profits.

The intuition for the monopolist being the only investment equilibrium in pure strategies is the following. Two firms cannot commit to a positive quality as the owner of the lower quality would profitably deviate by abstaining. All firms abstaining cannot be an equilibrium as well, since everyone would best respond by playing monopolist. So one firm playing monopoly is the only candidate equilibrium in pure strategies. It is indeed an equilibrium: potential entrants by (3.4) do not want to choose a quality below q^M; by deviating above q^M a firm will be interim monopolist and make profits

$$\begin{aligned} \Pi(q) &= V(q) - V\left(q^M\right) - c(q) \\ &= \int_{q^M}^{q} V'(q') - c'(q')\,\mathrm{d}q' - c\left(q^M\right) \end{aligned}$$

Both summands are negative as as $c'(q) > V'(q)$ above q^M. This contrasts with the possibilities of an ex-post deviator in the spirit of Rothschild and Stiglitz (1976), who upon entry can make revenues approximately close to those of an "idle" monopolist.

For equilibria with active competition, i.e. $n \geq 2$, standard arguments from war of attrition games prove that each firm must play an atomless distribution over qualities, and that they must make zero profits. Using

(3.5), the flat profit condition $\Pi'(q) = 0$, necessary for indifference, yields

$$H(q) = \frac{c'(q)}{V'(q)} \tag{3.9}$$

Equation (3.9) pins down the distribution of the maximal quality across competitors. Notice that $H(q) = 1$ right at $q = q^M$, so the support of the maximum across competitors, hence that of each firm, is $\left[0, q^M\right]$. Such a support restriction is implied by the fact that each firm's best response to a realized entry vector belongs to the doubleton set $\left\{0, q^M\right\}$: per (3.3) opponents' revenues are ex-post equivalent to a fixed cost, and only affect the decision of whether to enter, but not the quality upon entry.

Symmetry then implies the formula for H_n; notice the number of active firms is not pinned down. The CDF (3.8) is differentiable since V'' is continuous per Proposition 2.9 *ii*). To check that this is indeed an equilibrium, notice active firms are indifferent by construction on $\left[0, q^M\right]$, meaning that potential entrants would make expected losses by playing in that range (compete against n rather than $n-1$ players). Deviations above q^M are again excluded as profits would be

$$\Pi(q) = \Pi\left(q^M\right) + \int_{q^M}^{q} V'(q') - c'(q') \,\mathrm{d}q'$$

The first summand is zero in expectation by the flat profit condition (negative for a potential entrant), while the second term is negative by the definition of q^M. Again, the intuition is that irreversible investment gives incumbents the commitment to fight and drive expected profits to zero on the common support, so the interim monopolist with quality above q^M adds negative marginal profits to zero.

Combined with Proposition 3.1, the support restriction delivers the following

Corollary 3.4. *Irrespective of the number of active firms, with probability 1 a competitive market distributes*

i) a best quality strictly below q^M*, and*

ii) a strictly positive quality for free

Point *ii*) gives a simple empirical implication of the model: a positive quality is distributed for free if and only if there is active competition. If that is the case, we also know from *i*) that high valuation types receive a quality that is below their second-best allocation. Example 3.2 is expounded upon by computing first stage equilibria for linear preferences under a class of convex cost functions.

Example. (3.2 continued). Consider the class of convex cost functions $c(q) = q^\alpha$, $\alpha > 1$. Constant marginal revenues $V'(q) = \frac{1}{4}$ imply monopolist quality is

$$q^M(\alpha) = \left(\frac{1}{4\alpha}\right)^{\frac{1}{\alpha-1}}$$

Using Proposition 3.3 *i*), each one of n active firms plays in equilibrium the mixed strategy characterized by CDF

$$H_{n,\alpha}(q) = \mathbb{I}\left\{q \in \left[0, \left(\frac{1}{4\alpha}\right)^{\frac{1}{\alpha-1}}\right]\right\} \cdot (4\alpha q)^{\frac{\alpha-1}{n-1}}$$

3.3 Welfare

We now study ex-ante (expected) welfare across different equilibria. One active firm is the monopolist benchmark to which are associated the underacquisition and damaging inefficiencies isolated in expression

(2.15). The positive implications highlighted in Corollary 3.4 give a first idea of the impact that active competition has on welfare (relative to the monopolist equilibrium). Point i) implies that the underacquisition inefficiency will be worsened. Conditional on each realized best quality x two additional forces operate. One is a multiple spending inefficiency: all costs associated to qualities that realize below x are social waste as a planner could achieve the same allocation at no additional cost. This operates in the same direction as the underprovision inefficiency, favoring monopolist. Point ii) of Corollary 3.4 however implies that competition shrinks the image of the allocation function, thus reducing damaging inefficiencies (2.16).

Qualitatively, therefore, comparison between monopolist and active competition equilibria is inconclusive: underacquisition and multiple spending favor the former, undoing distributional inefficiencies favor the latter. Also, notice that the forces into play are similar to those of a NS policy: perturbation of the monopolist environment induces a worsening of the underacquisition inefficiency, though it may have positive distributional effects. A third channel however distinguishes the welfare impact of the NS policy from that of active competition: in the former case we have the "complete exclusion" margin, while in the latter the double spending inefficiency. Also, all competition outcomes are stochastic as firms play mixed equilibria in the first stage.

The next proposition shows that the relative strength of the potential welfare impacts is not unambiguously signed. We show by means of example that depending on the shape of the cost function one can favor either monopoly or competition with two active firms, which uniformly (i.e. type-wise) dominates equilibria with more intense competition. We initially define the surplus for type θ conditional on realized market statistics x, y

$$W(\theta, x, y) = g(y) + \int_0^\theta \max\{y, \min\{x, \beta(\theta')\}\} \, \mathrm{d}\theta' \tag{3.10}$$

and expected welfare of type θ in the n-equilibrium

$$W_n(\theta) := \mathbb{E}_n [W(\theta, \boldsymbol{x}, \boldsymbol{y})]$$

where expectation $\mathbb{E}_n$ integrates market statistics under the distribution induced by the equilibrium (3.8) with n active firms. Since all producers make zero expected profits, total surplus under competition is just

$$W_n = \mathbb{E}_\theta [\mathbb{E}_n [W(\theta)]]$$

Theorem 3.5.

i) Equilibria with active competition are Pareto-ranked, decreasing in n. Moreover, improvement is uniform in types, that is

$$W_n(\theta) \geq W_m(\theta)$$

for all θ and $2 \leq n \leq m$.

ii) If $q^M \leq (g')^{-1}(1)$ (monopolist does full bunching), then competition reduces welfare. Otherwise, we can specify a cost function under which double spending and underprovision inefficiencies vanish and the complete undoing of damaging inefficiencies make duopoly dominate.

We notice that type-dependent welfare (3.10) is an increasing function in both x and y. This is a natural consequence of the fact that larger x gives extra surplus to high types leaving unaffected (allocation and rent of) low types, while larger y increases the allocation of low types and reduces payment for high. For i) it is therefore sufficient to show that the *joint* distribution of $(\boldsymbol{x}, \boldsymbol{y})$ is ordered according to first order stochastic dominance (FOSD) in n. We prove that the distribution of $\boldsymbol{y}$ *conditional on* $\boldsymbol{x} = x$ is independent of n for each x, with (conditional) distribution given by

$$H_{x,n}(y) = \frac{H(y)}{H(x)} \mathbb{I}\{y \in [0, x]\} \tag{3.11}$$

and that the distribution of $\boldsymbol{x}$ is ranked in n according to FOSD, from which sufficiency follows. By point *i*) we can therefore compare 2 active firms[42] with the monopoly equilibrium.

Point *ii*) proves that the qualitative forces highlighted in Corollary 3.4 can have any relative strength and, depending on the shape of the cost function can favor either monopoly or duopoly. The message is delivered by considering two extreme cases: steep cost around zero and the fixed cost limit. We begin with the first case As there is no damaging, monopolist surplus is

$$W_1 = g\left(q^M\right) + \frac{1}{2} q^M - c\left(q^M\right)$$

Also, by Proposition 3.1, as $y \le x \le q^M$ for any market realization everyone will be allocated x at price $g(x) - g(y)$. Per (3.10), type dependent surplus is

$$W(x, y, \theta) = g(x) + \theta x - (g(x) - g(y)) = g(y) + \theta x$$

and total surplus is

$$W(x, y) = \int_{\Theta} W(x, y, \theta)\, \mathrm{d}\theta = g(y) + \frac{1}{2} x$$

By using the expression for the conditional distribution (3.11) and then integrating by parts, we show that welfare under competition $W_n = \mathbb{E}_n[W(\boldsymbol{x}, \boldsymbol{y})]$ can be written as

$$\underbrace{\left[g\left(q^M\right) - c\left(q^M\right) + \frac{1}{2} q^M\right]}_{W_M} - \int_0^q \left[\underbrace{g'(x) - \frac{\mathrm{d}}{\mathrm{d}x}\left(c(x)\frac{g'(x)}{c'(x)}\right)}_{>0} + \frac{1}{2}\right] H_n(x)\, \mathrm{d}x \tag{3.12}$$

First term is monopolist welfare, second term is the integral of a positive function, so the statement is proved.

We prove that in general, i.e. allowing for $q^M > (g')^{-1}(1)$, the difference between duopoly and monopolist welfare can be written as

$$W_2 - W_M = \mathbb{E}_2\left[\underbrace{\int_{(g')^{-1}(1)}^{\boldsymbol{y}} d(q)\, \mathrm{d}q}_{\text{undo screening}} - \underbrace{c(\boldsymbol{y})}_{\text{double spend}}\right] - \mathbb{E}_2\left[\underbrace{\int_{\boldsymbol{x}}^{q^M} m(q)\, \mathrm{d}q}_{\text{underacquisition}}\right] \tag{3.13}$$

where $\mathbb{E}_2$ is distribution of market statistics under duopoly equilibrium, m is marginal monopolist surplus function (2.17), d is marginal damaging inefficiencies (2.16), and the following convention is used: for any real valued function f,

$$\int_a^b f(x)\, \mathrm{d}x = 0 \tag{3.14}$$

[42] We call two active firms duopoly though the term may be misleading, as inactive firms are also key players in equilibrium.

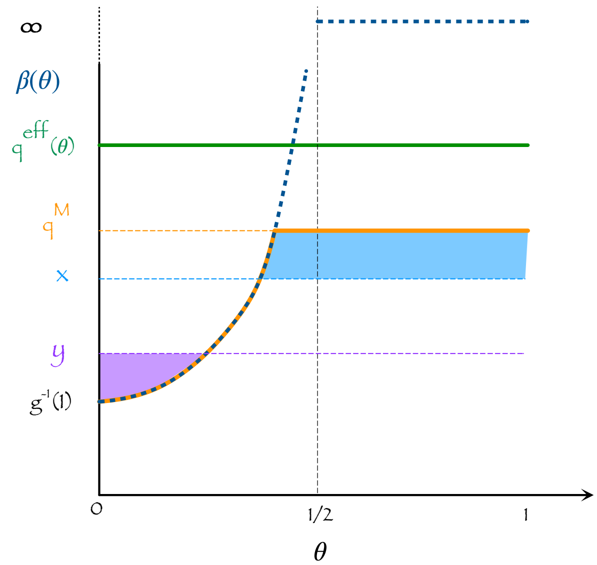

Figure 3.2: Screening undoing (purple) and underprovision (light-blue) welfare impact

whenever $a > b$. As we highlighted in the expression, each summand in (3.13) isolates the impact on one inefficiency. Figure 3.2 below plots for market realizations x, y the resulting welfare impact from undoing screening (purple), and worsened underacquisition (light-blue).

The only positive summand in (3.13) comes from realizations of $\boldsymbol{y}$ above $(g')^{-1}(1)$: unless it shrinks the allocation function compared to an $x-$monopolist, y has no impact on total welfare as it only transfers surplus from the seller to (all) consumers in the form of lower price. The case $q^M < (g')^{-1}(1)$ ensures this happens with probability one.[43] To get a positive impact of competition the following conditions are required. Both $\boldsymbol{x}$ and $\boldsymbol{y}$ should put significant mass on high realizations: the former must be close to q^M to reduce the integration domain of the underprovision inefficiency, while the latter should be well above $(g')^{-1}(1)$ to have significant undoing of damaging inefficiencies; finally, the cost of the second quality must be small to reduce the double spending inefficiency. Notice that high realizations of $\boldsymbol{y}$ have an ambiguous effect on welfare, as they increase both the integration domain for screening inefficiencies and costs.

Expression (3.13) is easily modified to evaluate welfare impact of higher intensity competition: we would need to take expectations under the a different (FOSD dominated) distribution of market statistics and to account for all multiple spending as all realizations below the best quality will increase the amount of wasteful acquisition.

We are left to show that in case monopoly has active damaging inefficiencies, then competition can dominate. To this end, suppose $1 > (g')^{-1}(1)$ and consider the limit of convex cost functions $c(q) = q^\alpha$ as α grows to infinity.[44] Irrespectively of the revenue function monopolist quality will converge to the point at which the cost function explodes, while its cost will converge to zero

$$q^M_\infty = \lim_{\alpha\to\infty} q^M(\alpha) = 1, \quad c^M_\infty = \lim_{\alpha\to\infty} \left(q^M(\alpha)\right)^\alpha = 0$$

[43]Indeed the implication that monopolist dominates duopoly whenever the former does not actively damage could be immediately derived from (3.13). We used the more direct welfare calculations that the subcase allowed to perform.

[44]It is assumed $g'(1) \leq 1$ so the asymptotic monopolist engages in inefficient damaging; otherwise we can target any asymptotic monopoly level $\overline{q}$ by letting $c_\alpha(q) = \left(\frac{q}{\overline{q}}\right)^\alpha$.

also, substituting marginal cost into (3.8) we get that equilibrium strategy converges in probability to q^M_∞, that is

$$H_2(q) \to \begin{cases} 0 & q < q^M_\infty \\ 1 & q = q^M_\infty \end{cases}$$

and $\mathbb{E}_\alpha[c(\boldsymbol{y})] \leq c\left(q^M_\alpha\right) \to 0$. Plugging those results in (3.13) we observe that the limit welfare impact of competition is given by

$$\begin{aligned} &\int_{(g')^{-1}(1)}^{q^M_\infty} d(q)\,\mathrm{d}q - c^M_\infty - \int_{q^M_\infty}^{q^M_\infty} [m(q) - c'(q)]\,\mathrm{d}q \\ &= \int_{(g')^{-1}(1)}^{q^M_\infty} d(q)\,\mathrm{d}q \end{aligned} \tag{3.15}$$

only the impact on screening undoing is active in the limit, and is also "complete": all types receive in the limit the quality that a monopolist would have produced (but not distributed).

We provided an expression for the welfare gains under duopoly and proved it cannot be unambiguously signed. Competition induces positive distributional effects that contrast increased underacquisition and double spending. The cost function can be tailored to shut down either channel. This is the main message of this Section.

We now complete the analysis of equilibria with linear preferences and generic costs started in Example 3.2 by studying their welfare properties and show how results derived in this section apply to a tractable example. Notice that with linear preferences there is no region in which the screening inefficiency is non-existent, as types $\left[0, \frac{1}{2}\right]$ are always (completely) excluded.

Example. (Example 3.2 continued) Using allocation rule (3.6), we can write market type-dependent welfare as

$$W(\theta, x, y) = \begin{cases} \theta y & \theta < \frac{1}{2} \\ \theta x - \frac{1}{2}(x - y) & \theta > \frac{1}{2} \end{cases}$$

So we can derive a closed form expression for market-contingent welfare[45]

$$W(x, y) = \int_0^{\frac{1}{2}} \theta y \mathrm{d}\theta + \int_{\frac{1}{2}}^1 \left[x\theta - (x - y)\frac{1}{2}\right] \mathrm{d}\theta = \frac{1}{8}[x + 3y] \tag{3.16}$$

With quadratic cost $c(q) = \frac{1}{2}q^2$, the monopolist produces $q^M = \frac{1}{4}$, while each firm in equilibrium with n active firms plays distribution

$$H_n(q) = \mathbb{I}\left\{q \in \left[0, \frac{1}{4}\right]\right\} \cdot (4q)^{\frac{1}{n-1}}$$

From which we can calculate expectations of first and second order statistics

$$\mathbb{E}_n[\boldsymbol{x}] = \int_0^{\frac{1}{4}} \left[1 - (4x)^{\frac{n}{n-1}}\right] \mathrm{d}x = \frac{n}{8n - 4} \tag{3.17}$$

$$\mathbb{E}_n[\boldsymbol{y}] = \int_0^{\frac{1}{4}} 4ny \left[1 - [4y]^{\frac{1}{n-1}}\right] \mathrm{d}y = \frac{n}{16n - 8} \tag{3.18}$$

Figure 3.3 plots equilibrium support and mean allocations (sufficient for welfare under linearity) for different competition intensities.

[45] Notice y has disproportionate impact as it increases welfare of low types and reduces payments of others.

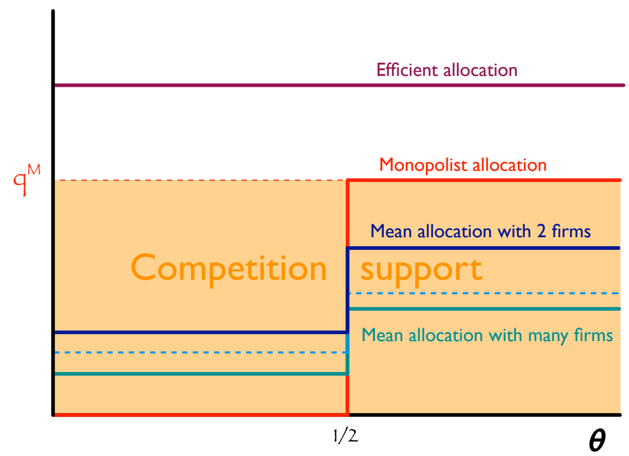

Figure 3.3: Expected Bertrand allocations with different intensity, quadratic cost.

Plugging (3.17) and (3.18) into (3.16) we get that welfare in an equilibrium with n active firms is

$$W_n = \mathbb{E}_n \left[W \left(\boldsymbol{x}, \boldsymbol{y} \right) \right] = \frac{1}{8} \mathbb{E}_n \left[\boldsymbol{x} + 3\boldsymbol{y} \right] = \frac{5}{64} \frac{n}{2n-1}$$

which is a decreasing function in n (Proposition 3.3 i)). Under monopolist, both consumer and producer surpluses are $\frac{1}{32}$, giving total surplus $\frac{1}{16}$. Notice

$$W_1 = \frac{1}{16} > \frac{5}{64} \cdot \frac{2}{3} = W_2 > W_3 > \cdots > \lim_{n \to \infty} W_n = \frac{5}{128} > \frac{1}{32} = CS^M$$

With moderately convex costs monopolist outperforms duopoly and competition of any intensity makes consumers better off than under monopoly.

Now consider generic convex cost $c(q) = q^\alpha$, and focus on the limit case $\alpha \to \infty$. We can check that

$$q^M(\alpha) = \left(\frac{1}{4\alpha} \right)^{\frac{1}{\alpha - 1}} \to 1, \quad c\left(q^M(\alpha) \right) = \left(\frac{1}{4\alpha} \right)^{\frac{\alpha}{\alpha - 1}} \to 0$$

which imply the following limit for monopolist surplus

$$W^M = \left[\int_{\frac{1}{2}}^{1} \left(\theta q^M - \frac{1}{2} q^M \right) \mathrm{d}\theta + \int_{\frac{1}{2}}^{1} \frac{1}{2} q^M \mathrm{d}\theta \right] - c\left(q^M \right) \longrightarrow 1 \int_{\frac{1}{2}}^{1} \theta \mathrm{d}\theta - 0 = \frac{3}{8}$$

Equilibrium play under duopoly is

$$H_{2,\alpha}(q) = \mathbb{I} \left\{ q \in \left[0, \left(\frac{1}{4\alpha} \right)^{\frac{1}{\alpha - 1}} \right] \right\} \cdot (4\alpha q)^{\alpha - 1}$$

which converges in probability to 1, so that

$$W_{2,\alpha} = \frac{1}{8}\mathbb{E}_{2,\alpha}\left[\boldsymbol{x} + 3\boldsymbol{y}\right] \to \frac{1}{8}\left[1+3\right] = \frac{1}{2}$$

Repeating the same steps we can show $W_{n,\alpha} \to \frac{1}{2}$ for each n, any intensity of competition allocates (approximately) quality 1 to all types, delivering surplus $\frac{1}{2}$, which exceeds monopoly surplus by $\frac{1}{8}$. Notice this conforms with equation (3.15), as $\frac{1}{8}$ is exactly the (limit) damaging inefficiency induced by excluding (at all quality levels) types $\left[0, \frac{1}{2}\right]$

$$\int_0^{q_\infty^M} d\left(q\right) \mathrm{d}q = \int_0^1 \left(\int_0^{\frac{1}{2}} \mathrm{d}\theta\right) \mathrm{d}q = \frac{1}{8}$$

3.4 Equilibrium properties and stability

This section has two purposes: *i*) to compare equilibrium outcomes with those obtained in models that have the most similar competition structure, and *ii*) to discuss equilibrium stability as robustness to deviations from idle firms that may unexpectedly occur as the game unfolds.

3.4.1 Similar models of competition

Making producers commit through an ex-ante irreversible investment is a modeling device used since Hotelling (1929): it makes equilibrium existence less problematic (and in our case also guarantees tractability) by granting incumbents some market power at the pricing stage. Separation of the production and pricing stages prevents indeed potential entrants from exploiting profitable deviations that may emerge when the competition environment or its outcome realize.

Two classical papers study similar competition games. Kreps and Scheinkman (1983) have firms commit to a *quantity* level before Bertrand competing (without screening) on the realized investments. The two stage equilibrium yields Cournot competition outcomes. The obvious difference with the setting of this paper is that the social value of aggregate production is obtained by summing quantities but taking the maximum over qualities of an homogenous good. As production along multiple lines is always wasteful, the result that competition may be beneficial is to some degree surprising. Champsaur and Rochet (1989) have the most similar setup as they analyze a MR duopoly where each competitor costlessly commits to a subset of qualities and then chooses a pricing function (paying the distribution costs at this stage). In committing to a quality range firms face a trade-off: they want a broad quality range to discriminate among consumers, but they also want to differentiate their products from those of the competitor as price competition lowers profit margins on neighboring qualities. They show that at a Nash equilibrium where each firm makes positive profits, the quality sets to which firms commit are always disjoint. Our investment game makes, in the language of Champsaur and Rochet (1989), firms commit to a quality range of the type $[0, q]$, so it is technologically impossible that two firms acquire disjoint sets. Indeed in all equilibria only one firm realizes positive revenues (Proposition 3.1), and first stage equilibria with active competition are only mixed (Proposition 3.3 *ii*)).

3.4.2 Stability

We thus far fixed an extensive form game and studied its properties: for sake of tractability we implicitly imposed strong timing and information rigidities, that must be justified by looking at, say, the length and transparency of R&D processes and patenting in the relevant markets, entry regulations etc. Contrary to Champsaur and Rochet (1989) our quality commitment stage is not cheap talk but requires a real and costly investment, though it still excludes plausible production deviations. The perfect information assumption is self-explanatory and so is the realism of its empirical counterparts.[46] We may be interested in investigating equilibria that are induced by different specifications of the competition environment. Unfortunately, no alternative specification delivered tractable results and we can only make an informal discussion of the stability of the equilibria we found. We loosely define stability as robustness to unanticipated deviations from inactive firms: the game is augmented by allowing firms that were idle to take some actions as the game unfolds. Since the game is two stage, two natural notions of stability emerge, depending on the stage at which outsiders are allowed to act.

- *Interim stability*: after the first stage is over a potential entrant observes the realized vector of entry qualities $\boldsymbol{q}$ and chooses whether (and eventually at which quality) to enter and play the second stage against $\boldsymbol{q}$.
- *Ex-post stability:* after the whole game is played a potential entrant observes the realized market pricing function and chooses whether (and eventually with which pricing function) to enter and compete with the realized contract.

Definition 3.6. The degree of interim (ex-post) stability of an equilibrium is the probability that the realized entry vector (market pricing function) does not induce interim (ex-post) entry.

Following the discussion above, it should be noticed that neither the ex-post nor the interim stability notion are associated to the equilibrium of an extended game in which active firms recognize the threat from outsiders: if he anticipates that at a later stage a potential entrant could wipe out his revenues, the interim monopolist would not (in general) choose allocation rule (3.2).[47] Similarly, the expected profit formula for the investment game would be different from (3.7) if active firms anticipated interim (and ex-post) deviations. However, if firms were playing a pure strategy equilibrium in the first stage, then the ex-post stability refinement would collapse to the Rothschild and Stiglitz (1976) equilibrium. Rothschild and Stiglitz (1976) invoke a notion of free-entry to justify the assumption that all firms observe the offered contracts[48] and must have no incentive to deviate, which is exactly what ex-post equilibrium in pure strategies requires.[49]

In this setting a Rothschild and Stiglitz (1976) contract would be a pricing function p_i. The cost of offering contract p_i is

$$\overline{c}\left(p_i\right) = c\left(\sup\left\{q : p_i\left(q\right) < \infty\right\}\right)$$

which, contrary to the original setting, does not depend on competitors' actions and consumers' demand. We associate to each set of pricing functions $\{p_i\}_{i\in\mathbb{N}}$ the firm specific revenue function $\overline{R}_j\left(\{p_i\}_{i\in\mathbb{N}}\right)$ given by (3.1) so we can define

Definition 3.7. An ex-post (Rothschild and Stiglitz (1976)) pure strategy equilibrium is a profile of contracts $\{p_i^\star\}_{i\in\mathbb{N}}$ such that

$$\overline{R}_i\left(p_i^\star, p_{-i}^\star\right) - \overline{c}\left(p_i^\star\right) \geq \overline{R}_i\left(p_i, p_{-i}^\star\right) - \overline{c}\left(p_i\right), \qquad \forall i, p_i$$

The following proposition establishes non-existence of ex-post equilibria and characterizes the degree of stability of monopolist and competitive equilibria.

[46]Otherwise timing is vacuous a competitors equivalently randomize over pricing functions, which would be a perfectly valid but intractable (for us) equilibrium concept.

[47]In particular, he would not alter his pricing decision if and only if this decision induces the potential deviator to abstain.

[48]In their case, a pair of coverage rate and deductible.

[49]Pure strategies make indeed the sequential nature of the game immaterial: as each player knows the realization of the quality vector and the strategy of opponents conditional on each quality vector, he knows the realization of the pricing functions and best responds to them.

Proposition 3.8. *i) There is no ex-post equilibrium in pure strategies.*

ii) The degree of interim stability is always larger than the degree of ex-post stability. The monopolist equilibrium is interim fully stable (degree 1) but ex-post fully unstable (degree 0).

iii) All competitive equilibria feature intermediate degrees of interim and ex-post stability, and both of them are decreasing in n.

The intuition for *i*) is the following: the monopolist pricing function (and all others abstain) is, by the same arguments used in Proposition 3.3, the unique candidate. However this time potential entrants can earn revenues that are ϵ close to $V\left(q^M\right)$ by just granting a small and equal discount to all types. The idle monopolist has no way to fight back, which breaks the candidate equilibrium.

We then notice that interim or ex-post best response implies producing a quality that lies in the doubleton set $\left\{0, q^M\right\}$. As

$$0 = R\left(q^M, q^M\right) < c\left(q^M\right) < R\left(q^M, 0\right) = V\left(q^M\right)$$

and $R\left(q^M, x\right)$ is monotonically decreasing there will be a threshold $m^\star \in \left(0, q^M\right)$ such that $R\left(m^\star, q^M\right) = c\left(q^M\right)$. We notice that at interim stage entry occurs if and only if the best active quality is below $m^\star$, while at an ex-post stage entry occurs if and only if the quality offered for free is below $m^\star$. So the degree of interim stability is the CDF (under equilibrium play) of the best quality evaluated at $m^\star$, and the degree of ex-post stability is the CDF of the second quality evaluated at $m^\star$. From this fact and the FOSD order of equilibrium statistics proved in Proposition 3.5 *i*), all results listed in *ii*) and *iii*) follow.

4 The Market for Information

This section discusses information markets as a potential application of the model. We highlight the implicit assumptions on the demand side and the production technology of information sources that make the framework of this paper more suited than other approaches (which are briefly reviewed) to analyze some phenomena in those markets. The main novelty is that production of information is decentralized and that the technology to convert the factor of production (attention) into state-signal structures must be taken as a primitive. We present a simple but exact microfoundation of the model in which firms observe increments of a common Brownian motion and sell the realization to agents solving a standard location problem. A natural extension in which correlation of the primary information sources induces product heterogeneity is presented since the basic model has the counterfactual implication that only one firm will be active (provide a non-trivial screening contract and make profit) even in the competitive setting. We motivate future extension of the model presented in this paper by discussing limitations imposed by the current simple structure.

4.1 A market for hard information

Many economically relevant situations can be modeled as a decision problem (or a game) preceded by an information acquisition phase in which agents speculate about some characteristics of the environment they will be acting in. To this aim it is necessary to specify what type of information agents can access and at what cost. The model presented in this paper can be applied to information markets with the following characteristics

- Decision makers (consumers) have no way to create their own information structure and must rely exclusively on the opinions sold by a set of profit maximizing sellers. This contrasts with models of unrestricted information acquisition with statistical pricing of information structures (e.g. Shannon entropy) which is the approach taken in (most) rational inattention models.[50]

- Firms are endowed with a technology to produce primary information structures and a (free) technology to Blackwell garble those structures. Both production and damaging occurs along a single dimension, exogenously given.

- Firms make revenues from selling information structures.[51] Price competition may become a second order concern (relative to attention and revenues from advertising) when information means entertainment and ideology drives agents' choice among opinion outlets. Newspapers, cable tv and similia are therefore not examples of information markets whose functioning is likely captured by this paper's model.[52]

- Communication of the signal (of any quality) is a *costless* operation that occurs *without frictions*: sellers have no disutility in repeating statements like "This asset is going to default with a likelihood $p \in A$." to whomever wants to pay for that, and investors have no difficulty in understanding what such statements mean,[53] to do the proper (bayesian) updating and to infer the value of such information from the decision problem they face.

- Reputation issues are also neglected: the seller cannot misreport the precision of the (menu of) signal he sells, and buyers believe the products they buy are indeed draws from the promised signal structure.[54]

Real world examples that approximately fit the description are websites that sell weather forecasts and assessments of the likelihood of default of a fixed income security (credit rating). Although (possibly) of limited interest, weather forecast is an insightful example since both the signal structure (probability of rain, temperatures bounds...), and the cost of acquisition (strengths of instruments, stations installed) have a clear interpretation. The website Accuweather offers basic, premium and professional subscriptions respectively for free, at \$7.95 and at \$19.95 monthly rate. Better packages add to basic service a longer horizon (up to 90 days), finer (hourly) weather forecast, experts opinions and radar images that clearly satisfy the feature of free damaging. The market for financial information is much larger and relevant, though some of its complexities require extensions of the model that we will address later. For the moment we let information production be the effort to evaluate the likelihood that a certain fixed-income security defaults. Credit rating has a clear signal structure (intervals of likelihoods of the default event) and to some extent can be modeled

[50]The rational inattention literature originates from the idea of Sims (1998, 2003) that decision makers are finite capacity information channels, unable to process all the information available. The information acquisition problem is equivalently rewritten by making agents pay an attention cost that is linear in the reduction of Shannon entropy, where the (per-bit) price emerges as the Lagrange multiplier on the attention constraint of the original problem. In this world, information is floating around agents that grasp it costly bit by costly bit; the fact that the "information bill" depends on some statistical property of the joint state-action distribution pays off in terms of a great tractability which accounts for some of the success of this approach.

[51]The vast literature on bayesian persuasion (initiated by Kamenica and Gentzkow (2011)) and, more in general, of information design (see Bergemann and Morris (2017)) study the problem of a principal who knows the state (has already "produced" the best information structure) and optimally transmits it to a set of agents. In both cases that objective is not the maximization of revenues from selling information securities: principal's utility depends on agents' actions which he influences by tailoring the information transmission.

[52]Galperti and Trevino (2017) endogenizes the supply of information as the outcome of competition among potential information sources that choose where to locate on the accuracy/clarity space in a Myatt and Wallace (2011) setting. In a different setting Perego and Yuksel (2018) studies competitive provision and endogenous acquisition of political information with horizontal differentiation of potential consumers. In both those papers firms compete for the *attention* of their consumers, which is justified as many information companies make most revenues from advertising.

[53]Again, this contrasts with a rational inattention setting.

[54]Since posting a menu of prices is essentially cheap talk and the model is static, this is also a strong assumption. Reputational issues in information transmission are studied by Wang (2009) and Ottaviani and Sørensen (2006) among others.

as the costly conjecturing exercise of some experts in the consultancy sector. Damaging is performed by coarsening the rating partition or hiding some parts of the report.

The key primitives that characterize such markets are *i*) the expression for the value of information and *ii*) the set of information structures that can be acquired and obtained through damaging. We proceed and give an example of both.

4.1.1 Information value

In an information market a type $\theta \in \Theta$ generically corresponds to a bayesian decision problem (action set, priors, utility) defined over a common uncertainty space (over which producers construct signal structures). This paper took this object as a primitive which means we can describe the shape of the optimal contract for classes of decision problems/type heterogeneity that induce a value of information having shape (2.1). It is however clear that to study specific phenomena we cannot take the value of available signal structures as exogenous but derive it from the relevant decision problem. Two recent papers that focus on information distribution derive the information value function from either heterogeneity in the decision problems or strategic externalities: Bergemann et al. (2018a) obtain a (piecewise) linear value of information when agents' types are their prior beliefs over a finite dimensional state space. They show (among many other results) that it will never be optimal for the seller to damage information by reducing precision (i.e. a quality dimension along which preferences are linear). This conforms with the "no-haggling" result stated in this paper as Corollary 2.10. In general they show that information is degraded and sold in non-trivial screening packages by revealing only a portion of the available data to the buyer: along this deterioration margin consumers' valuation are not linear.[55] Kastl et al. (2018) study the problem of a monopolist seller that may want to supply imprecise information to competitive firms that are uncertain about the marginal cost type of their contractors;[56] in their setup the state space is binary but information structures are allowed to be asymmetric and characterized by a two dimensional vector (α, β).

We can obtain an information value with concave component as in (2.1) from a decision problem that is standard in the literature: agents choose a location a to minimize the realized euclidean distance from an unknown state

$$u(a, \omega) = -(a - \omega)^2$$

so the value of an information structure $\mathcal{S} : \Omega \to \Delta(S)$ is just the expected (i.e. after observing a draw from $\mathcal{S}$) reduction in the variance

$$g(\mathcal{S}) = -\mathbb{E}_{\mathcal{S}}[\mathbb{V}_{post}] + \mathbb{V}_{prior}$$

If agents have normal prior belief $\omega \sim \mathcal{N}(\mu, \tau_p^{-1})$ and information structures take the form $\mathcal{N}(\omega, q^{-1})$ for $q \in \mathbb{R}$,[57] then such value is given by

$$g(q) = -\frac{1}{q + \tau_p} + \frac{1}{\tau_p} = \frac{q}{\tau_p(q + \tau_p)} \tag{4.1}$$

[55] As a concrete example of damaging through partial revelation occurring in information markets, Bergemann et al. (2018b) point at the "Undisclosed Debt monitoring" packages sold by Equifax in which the data broker offers individual rating reports to financial firms considering application for loans in three different versions differing in the number of "red flags" that the lender receives if the borrowers' history includes some negative events.

[56] They focus on the trade-off of a monopolist that may want to sell imprecise signal in order to limit distortions due to internal agency problems.

[57] Induced by speculation effort (4.2) (next section).

which is a concave function in q. Moreover, notice that

$$g''(q) = -\frac{1}{(q+\tau_p)^2}$$

is decreasing (in absolute value) in τ_p, so larger precision implies "more concave" returns from quality in the sense of footnote 15. To get the specification (2.1) we should therefore assume that agents solve two independent "location problems" and have heterogeneous valuations from guessing right the problem with larger prior variance. In general, we can keep the common location game as inducing the concave component, and take the linear part as some Taylor approximation for an additional type-dependent returns from precision.

4.1.2 Information production

Primary information structures are the state-signal correlations that can be produced by the firms. Suppose $\Omega = \mathbb{R}$ and the choice of primary information corresponds to observing a Brownian motion

$$dX_t = \omega dt + dW_t^j \tag{4.2}$$

for some period of (costly) time. After staring at the Brownian Motion for q_j units of time, firm j "produced" a signal (sufficient for ω)

$$s_j = \frac{1}{\sqrt{q_j}} X_{q_j} \sim \mathcal{N}\left(\omega, \frac{1}{q_j}\right)$$

about the state, which can be (damaged and) distributed to interested parties. Clearly, as every firm $i \in \mathbb{N}$ observes the same Brownian motion but just for different time, whoever stares at it the longer can push out of the market (owns a product that is superior to) everyone else. Every firm j can "quote" a market for any precision that is below q_j since inferior qualities can be obtained by reporting the Brownian motion at a period before q_j or by adding independent normal noise.[58] Notice we have made the implicit assumption that damaged structures must fall in the normal family. However there is no reason for which sellers should have this restriction: they can report whether it lies in a certain region, whether it is closer to point A or B, or commit to any garbling. The restriction to single dimensional quality is therefore substantial especially when it comes to damaging, and becomes untenable when we have a generic state space without parametric restrictions on primary structures.[59]

4.1.3 Correlation as product heterogeneity

Proposition 3.1 gives the implication that only one firm sells positive qualities and makes profits. This seems to be counterfactual at least in the market for financial consultancy where many people make money out of saying something. Product homogeneity in this framework arises from having all firms observe the same Brownian Motion (4.2): the empirical counterpart of this assumption is that all experts look at the same set of evidence (share a common reasoning process), or that meteorological instruments make perfectly correlated

[58] If agents do not care about correlation (e.g. they play independent decision problems) then how the signal is damaged is immaterial. If they were playing a game then also the correlation of the signal would induce a value. In particular, in presence of strategic substitutability agent derive value from being uncorrelated and we would obtain the result that a monopolist may acquire a quality higher than the highest quality distributed, just to be able to damage it in an agent-independent way.

[59] It is somehow natural to have single dimensional production sets: we can say that a weather forecaster can buy a stronger telescope that allows for fixed maps into signal structures, or that financial experts can only determine the intensity of their search and the resulting Markov kernels are model primitives. It is clearly a much stronger assumption to say that such conjecturing effort can be damaged only along the production dimension.

errors.

Such extreme assumption can be relaxed by letting the Wiener process driving observation of firm j be

$$dW_t^j = \rho dW_t + \sqrt{1-\rho^2} dZ_t^j \tag{4.3}$$

where dZ_t^j is a firm-specific process[60] and $\rho \in [0,1]$ parametrizes the correlation of the conjecturing effort of the different firms. Suppose for expositional clarity that there are two firms i, j and that at the investment stage they produced $q_i > q_j$. Despite being of inferior quality (correlation with the state) now X_{q_j} contains information about ω even after conditioning on X_{q_i}. Indeed, it holds

$$Cov\left(X_{q_j}, X_{q_i}\right) = \rho \min\{q_i, q_j\} = \rho q_j$$

Consumers only care about the final precision of the signal they observe (see (4.1)), that is $u(q_i, q_j, \theta)$ admits the aggregator representation

$$u(q_i, q_j, \theta) = u(\Psi(q_i, q_j), \theta)$$

where $\Psi : Q^2 \to Q$, derived by the normal updating formulas is the piecewise convex function given by

$$\Psi(q_i, q_j) = \frac{q_i(q_j(1-2\rho) + q_i)}{q_i - q_j \rho^2}$$

Notice that if $\rho = 0$ (signals are uncorrelated), then $\Psi(q) = q_i + q_j$ and we have a model of additive social value of production as in Kreps and Scheinkman (1983), though the distribution problem is subject to the screening frictions. If $\rho = 1$ then we are back to homogenous products and maximum aggregator $\Psi(q) = q_i$ which has been the subject of study of this paper. Continuity of Ψ implies full deterioration domain, namely that that for each $\{q_i\}_{i\in\mathbb{N}}$ and $q' < \Psi\left(\{q_i\}_{i\in\mathbb{N}}\right)$ we can find $\{q_i'\}_{i\in\mathbb{N}} \leq \{q_i\}_{i\in\mathbb{N}}$ so that $q' = \Psi\left(\{q_i'\}_{i\in\mathbb{N}}\right)$.

We define the cost of aggregate quality as the value function

$$\bar{c}(q) = \min_{\{q_i\}_{i\in\mathbb{N}}} \sum_{i\in\mathbb{N}} c(q_i), \qquad \text{s.t. } \Psi\left(\{q_i\}_{i\in\mathbb{N}}\right) \geq q \tag{4.4}$$

With the full deterioration property we can adapt the techniques used for the homogenous product and characterize first and second best. Second best is equivalent to a monopolist that owns all production sources and distributes damaged qualities only subject to information frictions: he allocates packages $\{q_i(\theta)\}_{i\in\mathbb{N},\theta\in\Theta}$ subject to IR and IC where each type θ can only choose among *profiles* $\{q_i(\theta')\}_{i\in\mathbb{N}}$ (cannot pick $q_i(\theta')$ and $q_j(\theta'')$ for $i \neq j$). Proposition (2.9) applies verbatim to characterize the second best distribution of *aggregate qualities* which are produced at cost (4.4). We cannot however decentralize the second best as a pricing stage equilibrium among competitive firms. The technical complications to solve for a competitive equilibrium with screening and heterogeneous products are illustrated in the Handbook chapter of Stole (2007), and we could not extend the tractability of competitive equilibria for homogenous digital goods to this more general case.

4.2 Limitations and future research

Section 4 had three objectives: *i*) discuss features of information markets that make them fit to be studied under the framework of this paper rather than under alternative approaches, *ii*) use standard building

[60] Meaning dW_t, dZ_t^j, dZ_t^i are uncorrelated Wiener processes.

blocks to provide a microfoundation for an information market that fits exactly the reduced form description given in the paper, and *iii*) suggest correlation in primary structures as a natural way to introduce product heterogeneity in a market that fails one stark empirical implication of the homogenous product model. Section 4.1.3 explored a tractable way to allow for heterogeneous in preferences that admit a quality aggregator. The unsolved technical challenge is the specification of a competition environment that delivers tractable equilibria. We conclude the section exploring other extensions of the model

- Contrary to a growing literature on information acquisition in markets, we do not allow for strategic interaction at the decision stage as this would make an exogenous (i.e. independent of equilibrium play) specification of the value for information untenable.[61] The extension would make a second quality dimension (broadly speaking, correlation with information given to co-players) emerge endogenously. Broadly speaking, this direction points at an heterogeneous agent version of Myatt and Wallace (2011) (or derived setups) with information acquisition as in (4.3) and price competition.

- The quality restrictions become substantive when the state space is large: as investors have interests that are differentiated either geographically or for the type of assets they trade, restricting sellers' marketable products to be single dimensional is unreasonable. We could think of a model with horizontally differentiated agents that care, say, only about some dimension of a large state space. A conjecturing effort produces signal structures about the whole state space but firms may choose to sell what they know about different portions of the state space as different products. We could use this specification to ask under what conditions firms favor production and marketing of signal structures characterized by a large breadth (learn about the approximate location of many states) or depth (focus on one state and identify it more precisely) component.

- The issue of non-excludability is particularly relevant for information markets: beyond prohibiting reselling of opinions, private information may be "leaked" through aggregate variables (this channel is explored in Admati and Pfleiderer (1986)). Many financial information packages include agent specific information (as the rating of potential borrowers in the Equifax example of Bergemann et al. (2018b)), and live prices (Bloomberg vs Reuters), which somehow reduce the concern of failure of non-excludability.

5 Conclusions

In this paper we developed a model of production and distribution of digital goods. The monopolist problem reduces to quality screening where the cost of an allocation is not additively separable but depends solely on one statistic of such allocation (the maximum). Under regularity assumptions on the demand side the optimal allocation is characterized by a bunching at the top threshold that increases with the quality cap, which is then easy to solve for. Market power and asymmetric information induce interdependent inefficiencies in acquisition and distribution of the digital good. Preventing damaging always worsens the acquisition inefficiency and induces exclusion of types that would be served under the unconstrained contract. The mechanical undoing of damaging inefficiencies counters those perverse effects yielding an ambiguous welfare

[61] Several contributions (among which Hellwig and Veldkamp (2009), Myatt and Wallace (2011), and Colombo et al. (2014)) show that in the presence of strategic externalities also the information acquisition game has strategic complementarities (i.e. there is an information acquisition externality). The game structure makes the *value* of information an endogenous object, but the *cost* is still exogenous even in cases where the supply side is a rich set of information sources who are characterized by an accuracy-clarity pair (as in Myatt and Wallace (2011)).

impact of the policy whose sign depends on the cost primitive. We then studied competition in digital goods markets as an extensive form game in which investment in quality is a sunk cost at the pricing stage. Monopoly is the only equilibrium in pure strategies, but there are also equilibria with different levels of competition. Competition induces wasteful double spending and worsens underacquisition since the highest quality distributed by a competitive market is below monopolist level. However, Bertrand forces induce a contraction of the screening domain that reduces distributional inefficiencies. Across equilibria with active competition the duopoly is Pareto dominant and welfare is decreasing in the intensity of competition. The welfare comparison between monopoly and duopoly is ambiguous and we can tailor the cost function to completely shut down the channels that favor either of them. The monopolist equilibrium is always subject to ex-post deviations while the duopoly features the highest degree of ex-post stability across all equilibria. We concluded by discussing how to apply and extend the model to study some phenomena in the market for information from a novel perspective.

References

Anat R Admati and Paul Pfleiderer. A monopolistic market for information. *Journal of Economic Theory*, 39(2):400–438, 1986.

Paul Beaudry and Michel Poitevin. Competitive screening in financial markets when borrowers can recontract. *The Review of Economic Studies*, 62(3):401–423, 1995.

Dirk Bergemann and Stephen Morris. Information design: A unified perspective. 2017.

Dirk Bergemann, Benjamin Brooks, and Stephen Morris. The limits of price discrimination. *American Economic Review*, 105(3):921–57, 2015.

Dirk Bergemann, Alessandro Bonatti, and Alex Smolin. The design and price of information. *American Economic Review*, 108(1):1–48, 2018a.

Dirk Bergemann, Alessandro Bonatti, et al. Markets for information: An introduction. *Cowles Foundation for Research in Economics Discussion paper*, 2018b.

Alberto Bisin and Piero Gottardi. Competitive equilibria with asymmetric information. *Journal of Economic Theory*, 87(1):1–48, 1999.

Alberto Bisin, John D Geanakoplos, Piero Gottardi, Enrico Minelli, and Herakles Polemarchakis. Markets and contracts. *Journal of Mathematical Economics*, 47(3):279–288, 2011.

Paul Champsaur and Jean-Charles Rochet. Multiproduct duopolists. *Econometrica: Journal of the Econometric Society*, pages 533–557, 1989.

Luca Colombo, Gianluca Femminis, and Alessandro Pavan. Information acquisition and welfare. *The Review of Economic Studies*, 81(4):1438–1483, 2014.

Raymond J Deneckere and R Preston McAfee. Damaged goods. *Journal of Economics & Management Strategy*, 5(2):149–174, 1996.

Pradeep Dubey and John Geanakoplos. Competitive pooling: Rothschild-stiglitz reconsidered. *The Quarterly Journal of Economics*, 117(4):1529–1570, 2002.

Bhaskar Dutta and Rajiv Vohra. Incomplete information, credibility and the core. *Mathematical Social Sciences*, 50(2):148–165, 2005.

Simone Galperti and Isabel Trevino. Coordination motives and competition for attention in information markets. Technical report, Mimeo, 2017.

Daniel Garrett, Renato Gomes, and Lucas Maestri. Competitive screening under heterogeneous information. 2014.

Andrew V Goldberg and Jason D Hartline. Envy-free auctions for digital goods. In *Proceedings of the 4th ACM conference on Electronic commerce*, pages 29–35. ACM, 2003.

Andrew V Goldberg, Jason D Hartline, and Andrew Wright. Competitive auctions and digital goods. In *Proceedings of the twelfth annual ACM-SIAM symposium on Discrete algorithms*, pages 735–744. Society for Industrial and Applied Mathematics, 2001.

Andrew V Goldberg, Jason D Hartline, and Andrew K Wright. Systems and methods for pricing and selling digital goods, January 10 2006. US Patent 6,985,885.

Jong-Hee Hahn. Damaged durable goods. *The RAND Journal of Economics*, 37(1):121–133, 2006.

Oliver D Hart and Jean Tirole. Contract renegotiation and coasian dynamics. *The Review of Economic Studies*, 55(4):509–540, 1988.

Jason D Hartline and Tim Roughgarden. Optimal mechanism design and money burning. In *Proceedings of the fortieth annual ACM symposium on Theory of computing*, pages 75–84. ACM, 2008.

Christian Hellwig and Laura Veldkamp. Knowing what others know: Coordination motives in information acquisition. *The Review of Economic Studies*, 76(1):223–251, 2009.

Martin F Hellwig. A note on the specification of interfirm communication in insurance markets with adverse selection. *Journal of Economic Theory*, 46(1):154–163, 1988.

H. Hotelling. Stability in competition. *Economic Journal, 39*, 1929.

Roman Inderst. Durable goods with quality differentiation. *Economics Letters*, 100(2):173–177, 2008.

Gerald D Jaynes. Competitive screening and market segmentation. 2006.

Gerald David Jaynes. Equilibria in monopolistically competitive insurance markets. *Journal of Economic Theory*, 19(2):394–422, 1978.

Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, 2011.

Jakub Kastl, Marco Pagnozzi, and Salvatore Piccolo. Selling information to competitive firms. *The RAND Journal of Economics*, 49(1):254–282, 2018.

David M Kreps and Jose A Scheinkman. Quantity precommitment and bertrand competition yield cournot outcomes. *The Bell Journal of Economics*, pages 326–337, 1983.

Eric Maskin and John Riley. Monopoly with incomplete information. *The RAND Journal of Economics*, 15 (2):171–196, 1984.

R Preston McAfee. Pricing damaged goods. 2007.

Michael Mussa and Sherwin Rosen. Monopoly and product quality. *Journal of Economic theory*, 18(2): 301–317, 1978.

Shigeo Muto. An information good market with symmetric externalities. *Econometrica: Journal of the Econometric Society*, pages 295–312, 1986.

David P Myatt and Chris Wallace. Endogenous information acquisition in coordination games. *The Review of Economic Studies*, 79(1):340–374, 2011.

Nick Netzer and Florian Scheuer. Competitive screening in insurance markets with endogenous wealth heterogeneity. *Economic Theory*, 44(2):187–211, 2010.

Marco Ottaviani and Peter Norman Sørensen. Reputational cheap talk. *The Rand journal of economics*, 37 (1):155–175, 2006.

Martin Peitz and Patrick Waelbroeck. Piracy of digital products: A critical review of the theoretical literature. *Information Economics and Policy*, 18(4):449–476, 2006.

Jacopo Perego and Sevgi Yuksel. Media competition and social disagreement. 2018.

Arnold Polanski. A decentralized model of information pricing in networks. *Journal of Economic Theory*, 136(1):497–512, 2007.

Jean-Charles Rochet and Philippe Choné. Ironing, sweeping, and multidimensional screening. *Econometrica*, pages 783–826, 1998.

Michael Rothschild and Joseph Stiglitz. Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *The quarterly journal of economics*, pages 629–649, 1976.

Christopher A Sims. Stickiness. In *Carnegie-Rochester Conference Series on Public Policy*, volume 49, pages 317–356. Elsevier, 1998.

Christopher A Sims. Implications of rational inattention. *Journal of monetary Economics*, 50(3):665–690, 2003.

Michael Smart. Competitive insurance markets with two unobservables. *International Economic Review*, 41 (1):153–170, 2000.

Padmanabhan Srinagesh and Ralph M Bradburd. Quality distortion by a discriminating monopolist. *The American Economic Review*, pages 96–105, 1989.

Lars A Stole. Price discrimination and competition. *Handbook of industrial organization*, 3:2221–2299, 2007.

Lisa N Takeyama. The welfare implications of unauthorized reproduction of intellectual property in the presence of demand network externalities. *The journal of industrial economics*, pages 155–166, 1994.

Hal R Varian. Buying, sharing and renting information goods. *The Journal of Industrial Economics*, 48(4): 473–488, 2000.

Rajiv Vohra. Incomplete information, incentive compatibility, and the core. *Journal of Economic Theory*, 86 (1):123–147, 1999.

Hefei Wang. Reputation acquisition of underwriter analysts–theory and evidence. *Journal of Applied Economics*, 12(2):331–363, 2009.

Robert Wilson. Information, efficiency, and the core of an economy. *Econometrica: Journal of the Econometric Society*, pages 807–816, 1978.

Robert B Wilson. *Nonlinear pricing.* Oxford University Press on Demand, 1993.

A Proof Appendix

A.1 Proof of Proposition 2.2 and Remark 2.3

As $u_q > 0$ for any type θ, deterioration is inefficient: the planner produces a quality that equates the average marginal valuation to the marginal cost and distributes such quality to each type. (2.5) is the first order condition of problem (2.4) after noticing the efficient allocation is flat, whose sufficiency is immediate.

A seller with perfect information can charge a type-dependent price p_θ. Then the profit maximization problem coincides with the social surplus problem, he produces $q^\star$, distributes it to all types and he extracts all the surplus.

A.2 Proof of Lemma 2.4

Consider the monopolist problem

$$\begin{array}{ll} \max_{\rho,p:\Theta\to Q\times\mathbb{R}} & \int_\Theta p(\theta)\,\mathrm{d}F(\theta) - \overline{c}(\rho) \\ \text{s.t.} & \text{IC,IR} \end{array} \tag{A.1}$$

and define

$$\omega(x) = \{\rho, p : \Theta \to Q \times \mathbb{R} : \text{ IC, IR hold and } \overline{c}(\rho) \le x\}$$

be the set incentive compatible and individually rational allocation and pricing function whose cost does not exceed x. Using this constraint sets, problem (A.1) can be rewritten as

$$\begin{aligned} &\max_{x,\{p:\exists\rho,(p,\rho)\in\omega(x)\}} \int_\Theta p(\theta)\,\mathrm{d}F(\theta) - x \\ &= \max_{x\in\mathbb{R}} \left[\max_{\{p:\exists\rho,(p,\rho)\in\omega(x)\}} \int_\Theta p(\theta)\,\mathrm{d}F(\theta)\right] - x \end{aligned} \tag{A.2}$$

now given the specification of the cost function (2.6)

$$\begin{array}{ccc} \rho\in\omega(x) & \text{only if} & c(\max_\theta \rho(\theta)) \le x \\ & \iff & \max_\theta \rho(\theta) \le c^{-1}(x) \end{array}$$

so

$$\omega(x) = \left\{\rho, p : \Theta \to Q \times \mathbb{R} : \text{ IC, IR hold and } \max_\theta \rho(\theta) \le c^{-1}(x)\right\}$$

and therefore

$$\max_{\{p:\exists\rho,(p,\rho)\in\omega(x)\}} \int_\Theta p(\theta)\,\mathrm{d}F(\theta) = V\left(c^{-1}(x)\right)$$

where V is the value of quality defined in (2.8) as the constraint set of that problem coincides with $\omega\left(c^{-1}(x)\right)$. Redefining the domain of choice to be $Q = c^{-1}(\mathbb{R})$ (which we can do as c strictly increasing), problem (A.2) becomes

$$\max_{q\in Q} V(q) - c(q)$$

as we wanted to show.

A.3 Proof Proposition 2.5

Fix an arbitrary q. The monopolist chooses allocations and transfers $\{q_i, T_i\}_{i \in \{H,L\}}$ to solve

$$V(q) = \max_{\{q_i, T_i\}_{i \in \{H,L\}}} (1-\pi) T_1 + \pi T_2$$

subject to

$$u(q_i, i) - T_i \geq u(q_j, i) - T_j$$

$$u(q_i, i) - T_i \geq 0$$

$$q_i \leq q$$

Combining the incentive constraint for high and low types we get monotonicity of the optimal allocation. Optimality implies the rationality constraint of the low type and the incentive to deviate from high to low types must be binding. With those observations the revenue maximization problem can written as a control problem where we choose allocation to high type and low type, x, y respectively.

$$V(q) = \max_{0 \leq y \leq x \leq q} u(L, y) + \pi (u(H, x) - u(H, y))$$

Maximization with respect to x immediately gives the corner solution

$$x^\star = q$$

Necessary condition for interior y is

$$u'(L, y^\star) - \pi u'(H, y^\star) = 0$$

$$\frac{g'(q) + \theta_L}{g'(q) + \theta_H} = \pi \tag{A.3}$$

Notice that by assumption $u''_H(q) = g''(q) = u''_L$ so

$$\frac{u'(L, x)}{u'(H, x)} = \frac{g'(x) + \theta_L}{g'(x) + \theta_H}$$

is monotonically decreasing, and equation (A.3) has (at most) one solution. and also guarantees that the SOC

$$u''(L, y^\star) - \pi u''(H, y^\star) = (1-\pi) g''(y^\star) < 0$$

is satisfied at the critical point, giving the unique solution to the program.

Now, if $y^\star \leq q$, then also the monotonicity constraint is satisfied and we have a global solution. If the threshold $y^\star$ is below the maximal quality we need to compare the fully pooling and exclusion equilibrium.

By the single crossing property, for all $x < y^\star$ it holds $u'(L,x) > \pi u'(H,x)$, therefore

$$u(q,L) = \int_0^q u'(L,x)\,\mathrm{d}x > \int_0^q \pi u'(H,x)\,\mathrm{d}x = \pi u(q,H)$$

and it is more profitable to serve q to all consumers at price $u(q,L)$. This proves L receives $y^\star$ if $y^\star \le q$ and q otherwise. Given allocation we can infer transfers from the binding constraints and obtain the expression for the revenue (and marginal revenue) function.

A.4 Proof of Lemma 2.8

We check that the virtual valuation is indeed a supermodular function in q, θ.

$$\begin{array}{rcll} \frac{\partial}{\partial q \partial \theta} vv(q,\theta) & = & \frac{\partial}{\partial q \partial \theta} & \left[g(q) + \theta q - q\frac{1-F(\theta)}{f(\theta)}\right] \\ & = & \frac{\partial}{\partial \theta} & g'(q) + \theta - \frac{1-F(\theta)}{f(\theta)} \\ & = & 1 - \underbrace{h'(\theta)}_{<0} & \ge 0 \end{array}$$

Now notice that

$$g(q) + q\left[\theta - \frac{1-F(\theta)}{f(\theta)}\right]$$

is maximized at $q = \infty$ whenever the multiplier on the linear part is greater than zero, since this is a monotonically increasing function of q. On the contrary, when $\theta \le \widetilde{\theta}$, then $\theta - \frac{1-F(\theta)}{f(\theta)} < 0$ and the objective is concave as the difference between a concave function g and a linear part. So maximizer $\beta(\theta)$ is characterized by the first order condition

$$g'(\beta(\theta)) = \frac{1-F(\theta)}{f(\theta)} - \theta$$

As g' is a strictly increasing function and can invert it to get the other branch of the maximizer function

$$\beta(\theta) = \begin{cases} (g')^{-1}(h(\theta) - \theta) & \theta < \widetilde{\theta} \\ \infty & \theta \ge \widetilde{\theta} \end{cases}$$

as we wanted to show.

A.5 Proof of Proposition 2.9

Fix a generic quality cap q. We want to show i), that is $\rho(q,\theta) = \min\{\beta(\theta), q\}$.

The quality constraint $\rho(\theta) \le q$ defining problem (2.8) is inserted in the objective function by subtracting to type dependent revenues the cost

$$c_\infty(q') = \begin{cases} 0 & q' \le q \\ \infty & else \end{cases}$$

So the cap conditional problem equivalently reads

$$\begin{array}{rcc} V(q) & \longmapsto & \max_{\rho,p:\Theta \to Q \times \mathbb{R}} \int_\Theta p(\theta) - c_\infty(\rho(\theta))\,\mathrm{d}F(\theta) \\ & \text{s.t.} & \text{IC , IR} \end{array}$$

Notice $c_\infty(q')$ is not differentiable, but can be approximated by the continuously differentiable convex function

$$c_n(q') = \left(\frac{q'}{q}\right)^n$$

We can now define the sequence of auxiliary problems

$$\begin{array}{lll} V_n(q) & \longmapsto & \max_{\rho,p:\Theta\to Q\times\mathbb{R}} \int_\Theta p(\theta) - c_n(\rho(\theta))\,\mathrm{d}F(\theta) \\ & \text{s.t.} & \text{IC , IR} \end{array}$$

as $\lim_{n\to\infty} c_n(q') = c_\infty(q')$, the objective in V_n converges to the objective in V and as policies and values of the auxiliary problems are bounded, the sequence of solutions (ρ_n, p_n) converges to the solution to the original problem. The auxiliary problem for a generic n is a monopolist screening problem with additively separable cost function, which we solve using standard arguments.

Firstly, the pairwise comparison of incentive constraints implies that the allocation ρ_n is monotonically increasing. Then, using the envelope theorem assuming sufficiency of the first order approach we obtain

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\left[u(\rho_n(\theta),\theta) - p_n(\theta)\right] = u_\theta(\rho_n(\theta'),\theta') = \rho_n(\theta')$$

from which we get the payoff equivalence function

$$u(\rho_n(\theta),\theta) - p_n(\theta) = \int_0^\theta \rho_n(\theta')\,\mathrm{d}\theta'$$

from which we infer prices

$$p_n(\theta) = u(\rho_n(\theta),\theta) - \int_0^\theta \rho_n(\theta')\,\mathrm{d}\theta'$$

and substitute in the objective to get the relaxed problem

$$W_n(q) \longrightarrow \max_{\rho_n(\theta)\ \text{increasing}} \int_\Theta \left[u(\rho_n(\theta),\theta) - c_n(\rho_n(\theta)) - \int_0^\theta \rho_n(\theta')\,\mathrm{d}\theta'\right] f(\theta)\,\mathrm{d}\theta$$

integrating by parts we obtain

$$\max_{\rho_n(\theta)\ \text{increasing}} \int_\Theta \left[u(\rho_n(\theta),\theta) - c_n(\rho_n(\theta)) - \rho_n(\theta)\frac{1-F(\theta)}{f(\theta)}\right] f(\theta)\,\mathrm{d}\theta$$

pointwise maximization of the integrand gives that a candidate $\rho_n(\theta)$ must satisfy

$$g'(\rho_n(\theta)) + \rho_n(\theta) - \frac{1}{h(\theta)} - \frac{n}{q^n}(\rho_n(\theta))^{n-1} = 0$$

that $\rho_n(\theta)$ so defined satisfies monotonicity follows from supermodularity of the profit function

$$r(x,\theta) = g(x) + \theta x - x\frac{1-F(\theta)}{f(\theta)} - \left(\frac{x}{q}\right)^n$$

as subtraction of a (convex) function of x does not change the sign of the cross derivative computed in Lemma 2.8 (neither sufficiency of the first order condition).

Recall that $\beta(\theta)$ solves

$$g'(\beta(\theta)) + \beta(\theta) - \frac{1}{h(\theta)} = 0$$

and as

$$\lim_{n \longrightarrow \infty} \frac{n}{q^n} (x)^{n-1} = \begin{cases} 0 & x < q \\ 1 & x = q \\ \infty & x > q \end{cases}$$

we conclude that the pointwise limit of $\rho_n(\theta)$ is

$$\rho_n(\theta) \to \begin{cases} \beta(\theta) & \text{if } \beta(\theta) < q \\ q & \beta(\theta) > q \end{cases}$$

and so $\rho(q,\theta) = \min\{\beta(\theta), q\}$, which is our desideratum.

ii) Now assume that types are uniformly distributed.[62] Using that, we write

$$V(q) = \int_0^1 u(\rho(\theta,q),\theta) - (1-\theta)\rho(\theta,q)\,\mathrm{d}\theta \tag{A.4}$$

it is convenient to define $b : Q \to \left[0, \frac{1}{2}\right]$ the inverse β function, namely

$$b(q) = \beta^{-1}(q) = \max\left\{0, \frac{1-g'(q)}{2}\right\} \tag{A.5}$$

so that

$$\rho(q,\theta) = \begin{cases} \beta(\theta) & \theta \leq b(q) \\ q & else \end{cases}$$

which can be substituted in A.4 to obtain

$$V(q) = \int_0^{b(q)} u(\beta(\theta),\theta) - (1-\theta)\beta(\theta)\,\mathrm{d}\theta + \int_{b(q)}^1 u(q,\theta) - (1-\theta)\,q\mathrm{d}\theta$$

now we can differentiate it

$$\begin{aligned} V'(q) &= & b'(q)\left[u(q,b(q)) - (1-b(q))\,q\right] \\ & - & b'(q)\left[u(q,b(q)) - (1-b(q))\,q\right] \\ & + & \textstyle\int_{b(q)}^1 u_1(q,\theta) - (1-\theta)\,\mathrm{d}\theta \\ &= & \textstyle\int_{b(q)}^1 g'(q) + (2\theta-1)\,\mathrm{d}\theta \\ &= & (1-b(q))(g'(q)-1) + \left(1-b(q)^2\right) \\ &= & (1-b(q))\left[g'(q) - 1 + (1+b(q))\right] \\ &= & (1-b(q))\left[g'(q) + b(q)\right] \end{aligned}$$

substituting expression (A.5) for b we get:

If $1 < g'(q)$ then $b(q) = 0$ and $V'(q) = g'(q)$

Otherwise,

[62] See discussion in Section

$$V'(q) = \left(\frac{1+g'(q)}{2}\right)^2$$

which is the expression in the Proposition. We are left now left to show that V' is $\mathcal{C}^1$. Continuous differentiability in the two branches is immediate, we need to show they are smoothly pasted.

$$\left(\frac{1+g'(q)}{2}\right)^2\Bigg|_{g'(q)=1} = 1 = g'(q)\Bigg|_{g'(q)=1}$$

proves continuity, while

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}q}\left(\tfrac{1+g'(q)}{2}\right)^2\Bigg|_{g'(q)=1} &= 2\tfrac{1+g'(q)}{2}\tfrac{g''(q)}{2}\Bigg|_{g'(q)=1}\\
&= 2\tfrac{1+1}{2}\tfrac{g''(q)}{2}\\
&= g''(q)\\
&= \tfrac{\mathrm{d}}{\mathrm{d}q}g'(q)
\end{aligned}$$

proves continuous differentiability.

iii) Per Proposition 2.2 the efficient quality $q^\star$ is determined by the first order condition

$$g'(q) + \frac{1}{2} = c'(q)$$

Monopolist quality q^M instead solves

$$V'(q) = c'(q)$$

so it is sufficient to show

$$V'(q) < g'(q) + \frac{1}{2} \tag{A.6}$$

always. Clearly, $g'(q) + \frac{1}{2} > g'(q)$ so in the full bunching region this is true. It is also easy to check that

$$x \leq 1 \implies x + \frac{1}{2} > \left(\frac{1+x}{2}\right)^2$$

proving $V'(q)$ is strictly below efficient marginal surplus even when $g'\left(q^M\right) \leq 1$. As (A.6) always holds, $q^M < q^\star$ and we have inefficient acquisition.

A.6 Proof of Proposition 2.12

We firstly need to compute consumers' surplus. As u_θ is $\rho(\theta, q)$, type θ welfare when the cap is q reads

$$W(\theta, q) = u(0) + \int_0^\theta u_\theta\left(\rho(\theta', q), \theta'\right)\mathrm{d}\theta' = \int_0^\theta \min\{\beta(\theta'), q\}\,\mathrm{d}\theta' \tag{A.7}$$

Integrating over $\Theta = [0,1]$ to get total consumer surplus, we have

$$\begin{array}{rcl} W(q) & = & \int_0^1 W(\theta, q)\,\mathrm{d}\theta \\ & = & \int_0^1 \int_0^\theta \min\{\beta(\theta'), q\}\,\mathrm{d}\theta'\mathrm{d}\theta \\ & = & \int_0^1 \min\{\beta(\theta), q\}(1-\theta)\,\mathrm{d}\theta \\ & = & \int_0^1 \min\left\{(g')^{-1}(2\theta-1), q\right\}(1-\theta)\,\mathrm{d}\theta \\ & = & \int_0^{b(q)} (g')^{-1}(2\theta-1)(1-\theta)\,\mathrm{d}\theta + \int_{b(q)}^1 q(1-\theta)\,\mathrm{d}\theta \end{array}$$

First line is definition, second substitutes (3.3), third is integration by parts, fourth expresses $\beta(\theta)$ and finally breaks the integral in the parts above and below q.

When differentiating, as for the marginal revenues V', the terms in b' will drop so we are left with

$$\begin{array}{rcl} W'(q) & = & \int_{b(q)}^1 \frac{\partial}{\partial q} q(1-\theta)\,\mathrm{d}\theta \\ & = & \int_{b(q)}^1 (1-\theta)\,\mathrm{d}\theta \\ & = & (1-b(q)) - \frac{1}{2}x^2\Big|_{b(q)}^1 \\ & = & (1-b(q)) - \frac{1}{2}\left(1-b(q)^2\right) \\ & = & \frac{1}{2}(1-b(q))^2 \\ & = & \frac{1}{2}\left(\frac{1+g'(q)}{2}\right)^2 = \frac{1}{2}V'(q) \end{array} \tag{A.8}$$

Integrating marginal surplus below $(\frac{1}{2})$, and above (A.8) $(g')^{-1}(1)$ we obtain (2.14).

To obtain the inefficiencies decomposition, notice first best surplus is given by

$$\begin{array}{rcl} S^{FB} & = & \frac{1}{2}q^\star + g(q^\star) - c(q^\star) \\ & = & \int_0^{q^\star} \left(\frac{1}{2} + g'(q) - c'(q)\right)\mathrm{d}q \end{array}$$

Summing marginal revenues (2.12) and consumer surplus (A.8) we get monopolist surplus below $(g')^{-1}(1)$ grows as first best (has no damaging), while above it grows with slope

$$\left(\frac{1+g'(q)}{2}\right)^2 - c'(q) + \frac{1}{2}\left(\frac{1+g'(q)}{2}\right)^2 = \frac{3}{2}\left(\frac{1+g'(q)}{2}\right)^2 - c'(q)$$

Monopolist welfare can therefore be written as

$$W_M = \int_0^{(g')^{-1}(1)} \left(\frac{1}{2} + g'(q) - c'(q)\right)\mathrm{d}q + \int_{(g')^{-1}(1)}^{q^M} \frac{3}{2}\left(\frac{1+g'(q)}{2}\right)^2 - c'(q)\,\mathrm{d}q \tag{A.9}$$

Notice in the region $\left[(g')^{-1}(1), q^M\right]$ we have marginal damaging inefficiencies

$$\begin{array}{rcl} d(q) & = & \underbrace{\frac{1}{2} + g'(q) - c'(q)}_{(S^{FB})'} - \underbrace{\left[\frac{3}{2}\left(\frac{1+g'(q)}{2}\right)^2 - c'(q)\right]}_{W_M'} \\ & = & \frac{1}{8}\left(1 + (2 - 3g'(q))\,g'(q)\right) \end{array}$$

By splitting integrals in the three regions $\left[0, (g')^{-1}(1)\right]$, $\left[(g')^{-1}(1), q^M\right]$ and $\left[q^M, q^\star\right]$ we can write losses relative to first best as

$$\begin{array}{rcc} S^{FB}-W_M & = & \int_0^{(g')^{-1}(1)}\left(\frac{1}{2}+g'(q)-c'(q)\right)\mathrm{d}q-\int_0^{(g')^{-1}(1)}\left(\frac{1}{2}+g'(q)-c'(q)\right)\mathrm{d}q \\ & + & \int_{(g')^{-1}(1)}^{q^M}\left(\frac{1}{2}+g'(q)-c'(q)\right)\mathrm{d}q-\int_{(g')^{-1}(1)}^{q^M}\frac{3}{2}\left(\frac{1+g'(q)}{2}\right)^2-c'(q)\,\mathrm{d}q \\ & + & \int_{q^M}^{q^\star}\left(\frac{1}{2}+g'(q)-c'(q)\right)\mathrm{d}q-0 \\ & = & 0 \\ & + & \int_{(g')^{-1}(1)}^{q^M}d(q)\,\mathrm{d}q \\ & + & \int_{q^M}^{q^\star}\left(\frac{1}{2}+g'(q)-c'(q)\right)\mathrm{d}q \end{array}$$

which is expression (2.15).

A.7 Proof of Proposition 2.13

i) The quality conditional profit reads

$$\Pi^{NS}(q)=\max_\theta\,[g(q)+\theta q](1-\theta)$$

First order condition is

$$\begin{array}{rcc} \theta q-\theta g(q)-\theta^2 q & \geq & 0 \\ \left[-[g(q)+\theta q]+q\right] & \theta\geq & 0 \\ \vartheta(q) & = & \max\left\{\frac{q-g(q)}{2q},0\right\} \end{array}$$

When $\frac{q-g(q)}{2q}=0$, then everyone is sold good q and $\Pi^{NS}(q)=g(q)$. If instead $\frac{q-g(q)}{2q}>0$ then only types $\left[\frac{q-g(q)}{2q},1\right]$ receive quality q and pay the valuation of the marginal consumer.

We need to show that the screening monopolist stops full bunching before the NS monopolist. The threshold for the screening monopolist $\beta(0)$ solves

$$\frac{1-g'(q)}{2}=0\implies g'(q)=1$$

The threshold for the NS monopolist $\vartheta^{-1}(0)$ solves

$$\frac{q-g(q)}{2q}=0\implies g(q)=q$$

As g is continuous and convex, $\beta(0)<\vartheta^{-1}(0)$ is an immediate implication of the Largange Theorem.

iii) The marginal revenue function $\frac{\mathrm{d}}{\mathrm{d}q}\Pi^{NS}(q)$ is equal to $g'(q)$ when $q-g(q)<0$ and we have full bunching (serve everyone at price $g(q)$). If $\vartheta(q)\in(0,1)$, we can use the envelope theorem to obtain

$$\begin{array}{rcc} \Pi^{NS}(q) & = & \max_\theta\,[g(q)+\theta q](1-\theta) \\ \frac{\mathrm{d}}{\mathrm{d}q}\Pi^{NS}(q) & = & [g'(q)+\vartheta(q)](1-\vartheta(q)) \\ & = & \left[g'(q)+\frac{q-g(q)}{2q}\right]\left(1-\frac{q-g(q)}{2q}\right) \\ & = & \frac{(q+g(q))(q-g(q)+2qg'(q))}{4q^2} \end{array}$$

proving the statement.

iii) As in the proof of Proposition 14 *iii*), it will be sufficient to show

$$\frac{\mathrm{d}}{\mathrm{d}q}\Pi^{NS}(q)\leq V'(q)$$

strictly when $g'(q) \leq 1$. We distinguish three cases.

If q is such that $g'(q) > 1$, then the constraint is immaterial and marginal revenues coincide.

If q is such that $g'(q) \leq 1$ but $g(q) > q$, then the monopolist screens so its optimal choice is given by

$$V'(q) = \left(\frac{1+g'(q)}{2}\right)^2 > g'(q) = \frac{\mathrm{d}}{\mathrm{d}q}\Pi^{NS}(q)$$

If also $g(q) < q$, then the derivative of the profit function is given by

$$\begin{array}{rcl} \frac{\mathrm{d}}{\mathrm{d}q}\Pi^{NS}(q) & = & \frac{(q+g(q))(q-g(q)+2qg'(q))}{4q^2} \\ & < & \frac{(q+q)(g(q)-g(q)+2qg'(q))}{4q^2} \\ & = & g'(q) \\ & \leq & \left(\frac{1+g'(q)}{2}\right)^2 \\ & = & V'(q) \end{array}$$

where the first inequality uses $g(q) < q$ (twice). This proves in general that $q^{NS} \leq q^M$, strictly whenever $g'(q) \leq 1$ (no full bunching in the monopolist case).

A.8 Proof of Proposition 2.14

i) Type θ welfare under the NS monopolist is

$$W^{NS}(\theta) = \begin{cases} \left(\theta - \vartheta\left(q^{NS}\right)\right) q^{NS} & \theta > \vartheta\left(q^{NS}\right) \\ 0 & else \end{cases}$$

a linear function (in θ) which is zero at the exclusion threshold and has slope q^{NS}. If q^M is in region B, then $\vartheta\left(q^{NS}\right) = 0$ and

$$q^M > q^{NS} > \beta(0)$$

As β is continuous in types there exists some $\overline{\theta} > 0$ for which

$$q^{NS} > \beta(\theta) = \rho\left(\theta, q^M\right) \qquad \forall\theta \in \left[0, \overline{\theta}\right]$$

Now pick $\theta \in \left(0, \overline{\theta}\right)$ and notice

$$\begin{array}{rcl} W(\theta) & = & \int_0^\theta \beta(\theta')\,\mathrm{d}\theta' \\ & < & \int_0^\theta q^{NS}\mathrm{d}\theta' \\ & = & \theta q^{NS} \\ & = & W^{NS}(\theta) \end{array}$$

and θ is better-off under the NS policy.

ii) Since in Region B no one is excluded by the NS monopolist, total surplus is

$$\begin{array}{rcl} W^{NS} & = & g\left(q^{NS}\right) - c\left(q^{NS}\right) + \frac{1}{2}q^{NS} \\ & = & \int_0^{q^{NS}} g'\left(q^{NS}\right) - c'\left(q^{NS}\right) + \frac{1}{2}\mathrm{d}q \end{array}$$

Subtracting this from monopolist surplus 53 and breaking down the integral in regions $(g')^{-1}(1) < q^{NS} < q^M$

we get

$$\begin{aligned} W^{NS}-W^{M} &= \int_{(g')^{-1}(1)}^{q^{NS}}\left(\tfrac{1}{2}+g'(q)-c'(q)\right)\mathrm{d}q-\int_{(g')^{-1}(1)}^{q^{M}}\tfrac{3}{2}\left(\tfrac{1+g'(q)}{2}\right)^{2}-c'(q)\,\mathrm{d}q \\ &+ 0-\int_{q^{NS}}^{q^{M}}\tfrac{3}{2}\left(\tfrac{1+g'(q)}{2}\right)^{2}-c'(q)\,\mathrm{d}q \\ &= \int_{(g')^{-1}(1)}^{q^{NS}}d(q)\,\mathrm{d}q-\int_{q^{NS}}^{q^{M}}\tfrac{3}{2}\left(\tfrac{1+g'(q)}{2}\right)^{2}-c'(q)\,\mathrm{d}q \end{aligned}$$

first term are welfare gains from undoing damaging, second are losses from underacquisition (compared to monopolist).

iii) We keep fixed q^M. As $c''\left(q^M\right)\to\infty$, then $c'\left(q^M\right)-c'(q)\to\infty$ for each $q>q^M$. As $V'\left(q^M\right)-\Pi'\left(q^M\right)$ is positive but finite this means $q^{NS}\to q^M$ and by *ii*) above

$$W^{NS}-W^{M}\to\int_{(g')^{-1}(1)}^{q^{M}}d(q)\,\mathrm{d}q$$

when convex costs shut down underacquisition, NS policy in Region B has the only (welfare increasing) effect of undoing screening inefficiencies.

A.9 Proof of Proposition 3.1

We avoid order statistics notation and let x,y generic be the realized maximum and second order statistics of the entry vector $\boldsymbol{q}$.

We firstly establish that in any market equilibrium, all qualities $q\leq y$ make zero revenues. Suppose otherwise, that is some firm j makes positive revenues selling a set of qualities bounded above by y. By definition of y at least an active competitor can modify his pricing function to copy the revenue-earner in that quality region, not alter the market pricing function (hence his revenues on other qualities) and share those positive revenues. Without invoking arbitrary tie-breaks, and for future reference, we should notice the competitor can, by quoting a pricing function $p_j(q)-\epsilon$ for ϵ arbitrarily close to 0 appropriate all revenues in the shared region without altering his revenues from other qualities: only marginal types would deviate as $\epsilon\to 0$. As no firm makes revenues on $q\leq y$, it follows $m(q)=0$ for all $q\leq y$, and in particular $m(y)=0$.

We now solve the problem of the interim monopolist. Beyond forcing $p_i(y)=0$ (or, equivalently, the quality allocation to be bounded below by y), competition has no impact on the interim monopolist as feasibility forces competitors $-i$ to quote a price $p_{-i}(q)=\infty$ for all $q>y$.

Using the same steps as for the unconstrained monopolist, we write the problem of the interim monopolist as choosing a type dependent quality allocation rule ρ which is increasing (pairwise comparison of IC) and has image $[y,x]$. The interim monopolist revenues are therefore

$$R_i(x,y)\longrightarrow\max_{\rho(\theta)\in[y,x],\text{ increasing}}\int_\Theta\left[u(\rho(\theta),\theta)-\rho(\theta)\frac{1-F(\theta)}{f(\theta)}\right]f(\theta)\,\mathrm{d}\theta$$

Pointwise maximization of the objective delivers the candidate allocation

$$\widetilde{\rho}(x,y,\theta)=\arg\max_{q\in[x,y]}g(q)+q(\theta-h(\theta))$$

From the concavity of the objective first order condition characterizes the interior optimum, hence $\widetilde{\rho}(x,y,\theta)=\beta(\theta)$ if $\beta(\theta)\in[y,x]$. The objective is instead strictly decreasing (on the relevant domain) in q if $\beta(\theta)<y$,

strictly increasing if $\beta(\theta) > x$. It therefore follows

$$\widetilde{\rho}(x, y, \theta) = \begin{cases} y & y < \beta(\theta) \\ \beta(\theta) & \beta(\theta) \in [y, x] \\ x & x > \beta(\theta) \end{cases}$$

which is a weakly increasing function (in θ), hence the solution of the interim monopolist problem. This proves the allocation rule (3.2).

We now need to compute the revenue function. Per the discussion above all firms but the interim monopolist make zero revenues. The interim monopolist earns

$$R_i(x, y) = R(y, y) + \int_y^x \frac{\partial}{\partial q} R(q, y)\,\mathrm{d}q \tag{A.10}$$

We have $R(y, y) = 0$ and, given allocation function (3.2), marginal revenues for the interim monopolist coincide with those of the unconstrained monopolist: the marginal quality is assigned to types $[b(q), 1]$ at marginal price $g'(q) + b(q)$. So

$$\frac{\partial}{\partial q} R(q, y) = \begin{cases} 0 & y > q \\ V'(q) & y \leq q \end{cases}$$

And we can rewrite (A.10) as

$$R_i(x, y) \quad = \quad \textstyle\int_y^x V'(q)\,\mathrm{d}q \quad = V(x) - V(y)$$

Since x, y was generic this proves expression (3.3) (after the appropriate notation changes are made).

A.10 Proof of Proposition 3.1

We divide the proof in several steps.

Step 1: **Monopolist is the only equilibrium in pure strategies.**

Consider a generic infinite profile of pure strategies $\widetilde{\boldsymbol{q}}$.[63] Suppose there is i, j with $0 < \widetilde{\boldsymbol{q}}_i \leq \widetilde{\boldsymbol{q}}_j$. Then by (3.3), firm i makes zero revenues but pays positive cost, she is better off staying inactive. So no two firms can choose positive qualities in the pure strategies equilibrium. $\mathbf{0}$ cannot be an equilibrium either as everyone's best response is to produce q^M. $\widetilde{\boldsymbol{q}}$ is therefore a candidate equilibrium in pure strategies only if $\widetilde{\boldsymbol{q}}_i = q^M$ for one and only one i, $\widetilde{\boldsymbol{q}}_{-i} = 0$. We need to show this is indeed an equilibrium. The entrant clearly has no incentive to deviate. Other firms do not enter with a quality below q^M as

$$q \leq q^M \implies R\left(q, q^M\right) = 0 < c(q)$$

If $q > q^M$, the deviator becomes interim monopolist and makes profits

$$\begin{aligned} \Pi(q) &= R\left(q, q^M\right) - c(q) \\ &= V(q) - V\left(q^M\right) - c(q) \\ &= \textstyle\int_{q^M}^q \left(V'(q') - c'(q')\right)\mathrm{d}q' - c\left(q^M\right) \end{aligned}$$

[63] Where $\widetilde{\boldsymbol{q}}$ is shorthand notation for each firm i plays $\delta_{\widetilde{\boldsymbol{q}}_i}$.

Both summands are negative as as $c'(q) > V'(q)$ above q^M. Now we look for equilibria where there are at least 2 active firms. By Step 1, those equilibria must be in mixed strategies.

Step 2: **Active firms play an atomless distribution with support including** 0**.**

The support of equilibrium play must contain 0: if the support was bounded below by a strictly positive quality $\underline{q}$, then playing the costly $\underline{q}$ would deliver zero revenues with probability 1, and abstaining is better. Similarly, in equilibrium no firm can choose a quality q with positive probability. If that were the case, all opponents best respond by placing zero probability on an open set including q to get a discrete jump in the probability of winning and (almost) the same profits. But then the firm itself wants to shift the mass away from q, depending on the sign of $V'(q) - c'(q)$.

Step 3: **The distribution of the maximum of opponent's qualities must be** $H(q) = \frac{c'(q)}{V'(q)}$.

We know from (3.3) that the maximum across competitors' realizations x is sufficient to determine firms' revenues. Let $H(x)$ be the distribution of such maximum, which from the previous step we know is continuous on $[0, \overline{q}]$ for some $\overline{q} > 0$. By playing q makes expected profits

$$\Pi(q) = \int_Q R(q,x)\,\mathrm{d}H(x) - c(q)$$

The flat profit condition $\Pi'(q) = 0$, necessary for indifference reads

$$\frac{\partial}{\partial q}\int_Q R(q,x)\,\mathrm{d}H(x) - c'(q) = 0$$

Using Leibnitz rule on an invariant support

$$\frac{\partial}{\partial q}\int_Q R(q,x)\,\mathrm{d}H(x) = \int_Q \frac{\partial}{\partial q} R(q,x)\,\mathrm{d}H(x)$$

Now use again

$$\frac{\partial}{\partial q} R(q,x) = \begin{cases} 0 & x > q \\ V'(q) & x \leq q \end{cases}$$

to write

$$\int_Q \frac{\partial}{\partial q} R(q,x)\,\mathrm{d}H(x) = \int_0^q V'(q)\,\mathrm{d}H(x) = V'(q)\int_0^q \mathrm{d}H(x) = V'(q)\,H(q)$$

Which, substituted in the flat profit condition gives the desideratum

$$H(q) = \frac{c'(q)}{V'(q)}$$

It holds $H(0) = 0$ since, by Proposition 2.9 *ii*), for low q marginal revenues $V'(q)$ is equal to $g'(q)$ approaching ∞ by the Inada condition. H is increasing as c is assumed convex and V is concave. The right extremum of the support is determined by

$$H(q) = 1 \implies c'(q) = V'(q) \implies q = q^M$$

Hence the maximum among $n-1$ competitors is an absolutely continuous random variable with support $\left[0, q^M\right]$.

Step 4: **An equilibrium candidate with** n **active firms is (3.8)**

The CDF H pins down the distribution of the maximal quality among $n-1$ competitors that makes the n^{th} firm indifferent among any quality $q \in \left[0, q^M\right]$. So for each n we have one (and only one) candidate equilibrium which has everyone plays

$$H_n(q) = [H(q)]^{\frac{1}{n-1}}$$

Also absolutely continuous with support $\left[0, q^M\right]$. This proves expression **(3.8)**. Notice CDF H_n admits a positive density

$$h_n(q) = \frac{1}{n-1}[H(q)]^{\frac{2-n}{n-1}} h(q)$$

It is continuous since $h(q) = \frac{\mathrm{d}}{\mathrm{d}q} H(q)$ is continuous in q in light of continuity of V'' established in Proposition 2.9 *ii*).

Step 5: **Sufficiency**

We are left to prove that this is indeed an equilibrium. All active firms are indifferent across all qualities in $\left[0, q^M\right]$ by the flat profit condition, they are indifferent with abstaining as 0 is in the support of the equilibrium. We are left to prove that firms do not want to produce more than q^M. In that case they would be the interim monopolist for sure, making profits

$$\Pi(q) = \Pi\left(q^M\right) + \int_{q^M}^{q} V'(q') - c'(q')\,\mathrm{d}q'$$

The first summand is zero in expectation by the flat profit condition, while the second term is negative by definition of q^M. That inactive firms do not want to produce any positive quantity is immediate: each of the n firms, competing against $n-1$ opponents makes zero profits in expectation and competing competing against n firms increases (in the sense of FOSD) the distribution of the best competitors' quality. Inactive firms are strictly better off abstaining completing the proof that this is an equilibrium.

A.11 Proof of Theorem 3.5

i) We preliminary derive type dependent consumer for each realized competitive environment x, y adding to the utility of the lowest type the integral of allocations (equal to u_θ) characterized in Proposition 3.1

$$\begin{aligned} W(\theta, x, y) &= u(0, x, y) + \textstyle\int_0^\theta u_\theta(\rho(\theta', x, y), \theta')\,\mathrm{d}\theta' \\ &= g(y) + \textstyle\int_0^\theta \max\{y, \min\{x, \beta(\theta')\}\}\,\mathrm{d}\theta' \end{aligned}$$

which is expression (3.10) in the main text (and for $x = q$, $y = 0$ give surplus under monopoly). It is immediate to notice that type dependent welfare is increasing in (x, y)

$$\forall\theta, \quad (x, y) \geq_2 (x', y') \text{ implies } W(\theta, (x, y)) \geq W(\theta, (x, y'))$$

where $\geq_2$ is the standard incomplete order in $\mathbb{R}^2$. Given monotonicity of value conditional on the realized qualities, to establish the result is sufficient to show that the random vector of marketed qualities $\boldsymbol{x}, \boldsymbol{y}$ has distribution ordered according to first order stochastic dominance (FOSD) in the equilibria with active competition.

Using individual firms' equilibrium play (3.8) we derive the distribution of the maximal quality $\boldsymbol{x}$ in equilibria with n active firms

$$\begin{array}{rcl} H_n[x] & = & Pr[\max\{q_1, q_2, \dots q_n\} \leq x] \\ & = & \prod_{i=1}^{n}[q_i \leq x] \\ & = & [H_n(x)]^n \\ & = & [H(x)]^{\frac{n}{n-1}} \end{array}$$

Since

$$s(n) = v^{\frac{n}{n-1}}$$

is a (strictly) increasing function of n for every $v \in [0,1]$, it follows

$$n > m \implies H_n[x] > H_m[x] \ \ \forall x$$

the best quality in equilibria with lower competition first order stochastically dominates the best quality in equilibria with more intense competition. Now we use the following fact[64]

Fact. *Let $X_1, \dots X_n$ be independent observations from a continuous CDF F. Then, the conditional distribution of the second order statistic given $\max_{i\in[n]} X_i = x$ is the same as the unconditional distribution of the maximum in a sample of size $n-1$ from a new distribution, namely the original F truncated at the right at x.*

It follows from the fact that the distribution of each other firm's quality conditional on $\boldsymbol{x} = x$ in an equilibrium with n active firms is

$$H_{x,n}(q) = \left[\frac{H(q)}{H(x)}\right]^{\frac{1}{n-1}} \mathbb{I}\{q \in [0,x]\}$$

so $\boldsymbol{y}\,|x$, the maximum across $n-1$ of them is distributed

$$H_{x,n}(y) = \left[\left[\frac{H(y)}{H(x)}\right]^{\frac{1}{n-1}}\right]^{n-1} \mathbb{I}\{y \in [0,x]\} = \frac{H(y)}{H(x)}\mathbb{I}\{y \in [0,x]\}$$

In particular, is independent of n. As the distribution of $\boldsymbol{x}$ is FOSD ranked across equilibria and the distribution of $\boldsymbol{y}$ given $\boldsymbol{x}$ is invariant across equilibria, it follows the joint of $\boldsymbol{x}, \boldsymbol{y}$ is FOSD across equilibria, which we argued above is sufficient for the statement.

ii) Suppose $y \leq x \leq q^M \leq (g')^{-1}(1)$. As both market statistics realize in the full bunching region, the competitive allocation (3.2) will assign every type the undamaged quality x at price $g(x) - g(y)$. Per (3.10), type dependent surplus is

$$W(x,y,\theta) = g(y) + \theta x$$

and total surplus is

$$W(x,y) = \int_{\Theta} W(x,y,\theta)\,\mathrm{d}\theta = g(y) + \frac{1}{2}x$$

So

$$W_n = \mathbb{E}_n[W(\boldsymbol{x},\boldsymbol{y})] = \mathbb{E}_x\left[\mathbb{E}_{y|x}[g(\boldsymbol{y})] + \frac{1}{2}\boldsymbol{x}\right] \tag{A.11}$$

[64] See theorem 6.7 in http://www.stat.purdue.edu/~dasgupta/orderstats.pdf.

Since

$$H_x(y) = \frac{g'(x)}{c'(x)} \frac{c'(y)}{g'(y)}, \quad y \in [0, x]$$

is the conditional CDF of $\boldsymbol{y}$ given $\boldsymbol{x} = x$ it follows

$$\begin{array}{rcc} \mathbb{E}_{y|x}\left[g(\boldsymbol{y})\right] & = & \int_0^x g(y)\,\mathrm{d}H_x(y) \\ & = & \frac{1}{H(x)}\left(H(y)\,g(y)\Big|_0^x - \int_0^x g'(y)\,H(y)\,\mathrm{d}y\right) \\ & = & \frac{1}{H(x)}\left[(H(x)\,g(x)) - \int_0^x g'(y)\,\frac{c'(y)}{g'(y)}\mathrm{d}y\right] \\ & = & g(x) - \frac{c(x)}{H(x)} \\ & = & g(x) - c(x)\,\frac{g'(x)}{c'(x)} \end{array} \tag{A.12}$$

The function

$$s(x) = g(x) - c(x)\,\frac{g'(x)}{c'(x)}$$

has $s(0) = 0$ and

$$\begin{array}{rcc} s'(x) & = & g'(x) - c'(x)\,\frac{g'(x)}{c'(x)} - c(x)\,\frac{g''(x)c'(x) - g'(x)c''(x)}{[c'](x)^2} \\ & = & -c(x)\,\frac{g''(x)c'(x) - g'(x)c''(x)}{[c'](x)^2} \\ & > & 0 \end{array}$$

so s is positive and monotonically increasing in $\left[0, q^M\right]$, and $s\left(q^M\right) = g\left(q^M\right) - c\left(q^M\right)$

Now, we substitute (A.12) into (A.11) to get

$$\begin{array}{rcc} W_n & = & \mathbb{E}_n\left[g(\boldsymbol{x}) - c(\boldsymbol{x})\,\frac{g'(\boldsymbol{x})}{c'(\boldsymbol{x})} + \frac{1}{2}\boldsymbol{x}\right] \\ & = & \int_0^{q^M}\left[g(x) - c(x)\,\frac{g'(x)}{c'(x)} + \frac{1}{2}x\right]\mathrm{d}H^{\frac{n}{n-1}}(x) \end{array}$$

Integrating by parts the last expression we obtain

$$\left[g\left(q^M\right) - c\left(q^M\right) + \frac{1}{2}q^M\right] - \int_0^{q^M}\left[g'(x) - \frac{\mathrm{d}}{\mathrm{d}x}\left(c(x)\,\frac{g'(x)}{c'(x)}\right) + \frac{1}{2}\right]H^{\frac{n}{n-1}}(x)\,\mathrm{d}x$$

the first term is monopolist profit, while we can notice

$$g'(x) - \frac{\mathrm{d}}{\mathrm{d}x}\left(c(x)\,\frac{g'(x)}{c'(x)}\right) = s'(x) > 0$$

as proved above, so we are subtracting the integral of a positive function, reducing monopolist welfare and proving the statement.

We now prove the general welfare decomposition (3.13). Fix a realization of market statistics x, y. We can write

$$W(x, y) = \int_0^x \frac{\mathrm{d}}{\mathrm{d}q} W(x, y)(q)\,\mathrm{d}q$$

where $\frac{\mathrm{d}}{\mathrm{d}q} W(x, y)(q)$ is the marginal contribution to welfare . It holds

$$\frac{\mathrm{d}}{\mathrm{d}q} W(x, y)(q) = \begin{cases} \left(S^{FB}\right)'(q) - c'(q) & q < y \\ m(q) & q \in [y, x] \end{cases}$$

since for qualities below y welfare from competition grows as first best (since all all goods are distributed to everyone) net of the additional marginal costs (incurred twice). In the quality region $[y, x]$ surplus grows as monopolist since those qualities are assigned to the same types and cost incurred only once. It then follows

$$W(x,y) - W_M = \int_0^x \frac{\mathrm{d}}{\mathrm{d}q} W(x,y)(q)\,\mathrm{d}q - \int_0^{q^M} (W_M)'(q)\,\mathrm{d}q =$$

As monopolist surplus grows as first best below $(g')^{-1}$ and as the second branch of $m(q)$ in $\left[(g')^{-1}, q^M\right]$ we rewrite the difference (assuming $y > (g')^{-1}$)

$$= -c(y) + \underbrace{\int_0^{(g')^{-1}} \left[\left(S^{FB}\right)'(q) - \left(S^{FB}\right)'(q)\right]\mathrm{d}q}_{=0} + \int_{(g')^{-1}}^{y} \underbrace{\left[\left(S^{FB}\right)'(q) - (W_M)'(q)\right]}_{=d(q)}\mathrm{d}q$$

$$+ \underbrace{\int_y^x m(q) - m(q)\,\mathrm{d}q}_{=0} - \int_x^{q^M} m(q)\,\mathrm{d}q$$

$$= \int_{(g')^{-1}}^{y} d(q)\,\mathrm{d}q - c(y) - \int_x^{q^M} m(q)\,\mathrm{d}q$$

Taking expectations under the duopoly equilibrium distribution of market statistics $\boldsymbol{x}, \boldsymbol{y}$,

$$W_2 = \mathbb{E}_2\left[W(\boldsymbol{x}, \boldsymbol{y})\right]$$

we obtain the fundamental welfare decomposition

$$W_2 - W_M = \mathbb{E}_2\left[\int_{(g')^{-1}(1)}^{\boldsymbol{y}} d(q)\,\mathrm{d}q - c(\boldsymbol{y})\right] - \mathbb{E}_2\left[\int_{\boldsymbol{x}}^{q^M} \left[m(q) - c'(q)\right]\mathrm{d}q\right]$$

where we mean

$$\int_{(g')^{-1}(1)}^{y} d(q)\,\mathrm{d}q = 0$$

whenever $y < (g')^{-1}(1)$.

We are now left to show that under convex cost $c(q) = q^\alpha$ as α grows to infinity equilibria with active competition dominate monopoly. As $q_M(\alpha)$ solves $V'(q) = \alpha q^{\alpha-1}$, irrespectively of the revenue function monopolist quality will converge to the point at which the (marginal) cost function explodes, that is

$$q^M_\infty = \lim_{\alpha\to\infty} q^M(\alpha) = 1$$

while its cost will converge to zero

$$c^M_\infty = \lim_{\alpha\to\infty} \left(q^M(\alpha)\right)^\alpha = 0$$

also, since

$$\lim_{\alpha\to\infty} c'_\alpha(x) \to \begin{cases} 0 & x < 1 \\ 1 & x = 1 \\ \infty & x > 1 \end{cases}$$

it follows that

$$H_{2,\alpha}(q)=\begin{cases}\frac{c'(q)}{V'(q)} & q\le q^M(\alpha)\\ 1 & q>q^M(\alpha)\end{cases}\longrightarrow\begin{cases}0 & q<q^M_\infty\\ 1 & q\ge q^M_\infty\end{cases}$$

which means that each firm's equilibrium strategy converges in probability to q^M_∞. Also, notice that as $\boldsymbol{y}\le q^M(\alpha)$ then

$$\mathbb{E}_\alpha\left[c(\boldsymbol{y})\right]\le c\left(q^M_\alpha\right)\to 0$$

Plugging those results in (3.13) we observe that the limit welfare impact of competition is given by

$$\int_{(g')^{-1}(1)}^{q^M_\infty} d(q)\,\mathrm{d}q-c^M_\infty-\int_{q^M_\infty}^{q^M_\infty}\left[m(q)-c'(q)\right]\mathrm{d}q$$

$$=\int_{(g')^{-1}(1)}^{q^M_\infty} d(q)\,\mathrm{d}q \tag{A.13}$$

proving the statement.

A.12 Proof of Proposition 3.8

i) By the same reasoning as in the proof of Proposition 3.3, one firm offering the monopolist pricing function $p^M(\cdot)$ is the only candidate equilibrium: no two firms can quote a non-trivial pricing function but someone must. However in this case, an inactive firm can offer pricing function $p^M(\cdot)-\epsilon$ (so that allocations would be unchanged), pay $c\left(q^M\right)$ and make revenues that are $\epsilon-$close to $V\left(q^M\right)$, a profitable deviation. There is no RS equilibrium.

For the second statement in *ii*) we notice that full interim stability of the monopolist follows from Proposition 3.3 *i*) (it is a Nash equilibrium of the first stage game), while its ex-post instability is proved above.

To prove the remaining part of *ii*) and *iii*) we preliminarily notice that

$$0=R\left(q^M,q^M\right)<c\left(q^M\right)<R\left(q^M,0\right)=V\left(q^M\right)$$

and $R\left(q^M,x\right)$ is monotonically decreasing, so there will be a threshold $m^\star$ at which $R\left(m^\star,q^M\right)=c\left(q^M\right)$. By Proposition 3.3 a potential entrant against best quality x can make revenues

$$\widetilde{R}(x)=\max_q R(q,x)-c(q)$$

As $R(x,q)=V(q)-V(x)$, x induces a sunk cost, upon entry q^M is played and

$$\widetilde{R}(x)=\max\left\{V\left(q^M\right)-V(x)-c\left(q^M\right)\right\}$$

From which it follows

$$\widetilde{R}(x)>0\iff x<m^\star$$

By a similar argument, an ex-post deviation occurs if and only if the second quality is above $m^\star$. A deviator that enters with q^M can indeed offer allocation $\rho\left(y,q^M,\cdot\right)$ and make revenues that are arbitrarily close to $V\left(q^M\right)-V(y)$: to do that he must grant a small discount to all types in $\left[b(y),b(x)\right]$ so to push the interim monopolist out of the market. Whenever y realizes strictly below $m^\star$ therefore an ex-post entrant the profitable deviation of offering the interim-contract $\left[y, q^M\right]$, with prices ϵ-reduced to all types for which $\beta\left(\theta\right)$ is below the level owned by the realized interim monopolist.

It therefore follows that the degree of interim stability of the $n-$equilibrium is $H_{x,n}\left(m^{\star}\right)$ and the degree of ex-post stability is $H_{y,n}\left(m^{\star}\right)$, where $H_{x,n}$, $H_{y,n}$ are the CDF of market statistics calculated under equilibrium play with n active firms. Point *ii*) now follows from $\boldsymbol{x} \geq \boldsymbol{y}$ by definition, while the $n-$ranking in *iii*) follows from the FOSD ranking of market statistics proved in Proposition 3.5 *i*).