

Supply and Demand Law under Limited Information

Yi-Cheng Zhang*

March 2004

ABSTRACT:

We present a model for the supply-demand law with quality and limited information capability. We postulate that imperfect information permeates in almost all economic transactions to varying degrees. Through a simple model we outline a research agenda that re-examines many standard issues in economics. Our analysis shows that whereas imperfect information can be improved, it leads to new uncertainties so that the perfect information limit can never be reached. As a corollary neoclassical perfect equilibrium can never be attained.

1 Introduction

The standard supply and demand law in economics relates price and quantity; equilibrium is achieved when the consumers' downward demand curve and the firm's upward supply curve intersect. When Alfred Marshall more than a century ago first systematically studied the law and its applications, he used it merely as a convenient technical device[1]. Neoclassical economics since has made it the most important theoretical pillar. The core content of the supply-demand law has remained the same over the past century, in which the world economy has undergone dramatic transitions. However, products in the modern economy are much more complex a century later, and hence the capability to ascertain their quality from the consumers' side plays a paramount importance in purchase decisions.

*Department of Physics, University of Fribourg, CH-1700, Fribourg, Switzerland.
yi-cheng.zhang@unifr.ch

In the days of Mill and Ricardo there were already vocal critiques by well known thinkers like Charles Babbage[2] and Cliffe Leslie[3], questioning the assumption of classical economics of perfect information capability. Fast forward to 1970, Akerlof’s work [4] on the ‘Lemons Problem’ clearly pointed out that strong information asymmetry about the product’s quality can lead to ‘market failures’. Mutually beneficial deals between sellers and buyers may not happen, if the buyers cannot reliably determine the quality of the objects under transaction. Despite much current research on many other interesting areas of Information Asymmetry[5][6], the fundamental insight of Akerlof is not yet adequately incorporated into mainstream economics. In standard textbooks one encounters ‘price’ and ‘quantity’, rarely ‘quality’ is adequately modeled.

In this work we shall consider an alternative version of the supply-demand law, with *quality* and *imperfect information* as the key ingredients. We consider a continuously varying degree of imperfect information, with the Akerlof ‘Lemons Problem’ as a special, extreme case. We posit that in all economic transactions, some degree of information imperfection always exists. In the modern economy, especially in affluent societies, the conduct of daily life necessitates a myriad of products and services, which become ever more complex in their visible and invisible features. On the other hand a consumer as a ‘generalist’[18] in consumption cannot possibly spend sufficient time and resources to determine what she is buying each time. Our position is anticipated by Kenneth Arrow[7]: “Market failure is not absolute; it is better to consider a broader category, that of transaction costs, which in general impede and in particular cases block the formation of markets”.

2 Quality and Information

2.1 demand side

Let us denote by Q the quality of a product (to be distinguished from the usual notation quantity—hereafter denoted by q); by $1/\epsilon$ information capability for a consumer to see through the quality. $\epsilon = \infty$ implies she is totally ignorant of the product’s quality; $\epsilon = 0$, she has perfect clairvoyance, i.e. no errors made. The information capability depends on her ability and effort of detecting the quality as well as on the producer’s marketing strategies of either facilitating or hindering such detection. Thus ϵ , a relational parame-

ter, depends on both sides rather than consumers' innate capability alone. A consumer may be good at ascertaining quality for one type of product while being ignorant for many other types. ϵ can also serve as a measure of the severity of information asymmetry.

Note that quality Q cannot be *a priori* measured in monetary terms: it must be priced through market transactions. Price p and quality Q should be treated as independent variables. From daily economic life we can observe that quality and price are always more or less correlated, sometimes to an astonishing degree. This intuition may lead us to believe that quality and price are so well correlated that there is no need to treat quality as an independent variable. On the other hand, we also observe that small but appreciable quality fluctuations do exist for a given price; experts and careless consumers may end up buying products with substantial deviations in the quality-to-price ratio. By treating quality and price independently we shall nevertheless recover their correlation through market transactions; but such a correlation depends on the level of information capability of the consuming public.

The central function for our present work $q(Q, p, \epsilon)$ measures the aggregate quantity consumers will buy a given product. For simplicity we assume they are all of the same information capability level—more interesting inhomogeneous case will be postponed till a later section. The aggregate demand is from N potential consumers

$$q(Q, p, \epsilon) = \sum_{i=1}^N q_i(Q, p, \epsilon). \quad (1)$$

We assume that any product is to some extent substitutable, therefore a potential buyer can decide either to buy, how much, or walk away. For simplicity we consider the product in discrete units. A q_i less than unity, say 0.30, implies the consumer has the probability of 30% to buy; a q_i greater than unity can be directly interpreted as the individual's demand function. Throughout this paper we shall use the aggregate q rather than the individual q_i , unless otherwise specified. All the three variables (q, Q, p) are expressed in unspecified units such that only their relative values are of interest. Thus the quantity q is linearly proportional to the 'likelihood' for a consumer to buy a given product, apart from a constant of the order of $O(N)$.

To highlight the new role of quality Q and to keep the algebra to a minimum, we first consider price p being fixed, and postpone the discussion

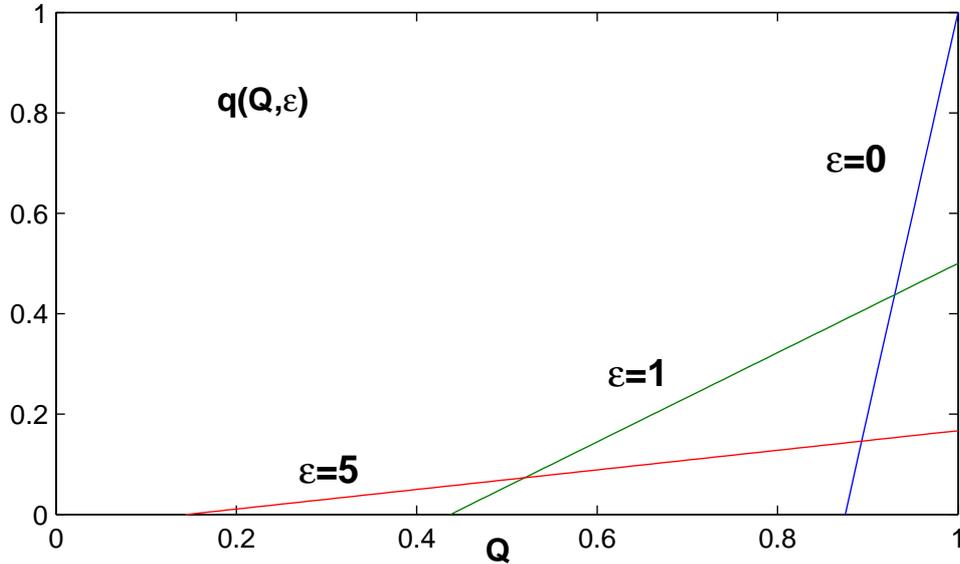


Figure 1: *The demand function $q(Q, \epsilon)$ for three different values of ϵ .*

of the full set of parameters and their relations in the subsequent sections. Hence for our reduced problem we have $q(Q, \epsilon)$. Let us start by considering the hypothetical case of perfect clairvoyance $q(Q, \epsilon = 0)$, in Fig.1. It is a sharply downward linear function as quality Q decreases. Transactions have the highest likelihood when $Q = Q_{\max}$, (for convenience we take this Q_{\max} value to be unity). This likelihood drops linearly to null at $Q_{\min} = AQ_0$, and remains null for $Q < Q_{\min}$. The difference $Q_{\max} - Q_{\min}$ measures the tolerance for the consumer. Even armed with perfect clairvoyance, her buying likelihood is tolerant within a finite range. This range can depend on possible competitions among the producers—a question we shall raise later.

Let us also plot $q(Q, \epsilon)$ for finite ϵ in Fig.1, i.e. limited capability to detect the quality of the product. Our general assumption in this work is that with reduced transparency, the consumer may misperceive the true quality, to the extent parameterized by ϵ . Thus the demand curve is somewhat flatter than the perfect clairvoyant case. In other words, the demand response to the quality incentive is less sensitive.

The overall transaction likelihood should also drop as quality visibility is reduced, in line with Akerlof's 'market failure' observation. The highest likelihood (at $Q = Q_{\max}$) is reduced by a factor $A = 1/(1 + \epsilon)$. This stems

from the consideration that being aware of their handicap of information asymmetry, consumers reduce, somewhat indiscriminately of quality, their overall confidence in what they would purchase. The reduction is in relation to the severity of the information asymmetry. Akerlof ‘Lemons Problem’ is recovered when $\epsilon = \infty$, i.e. the worst degree of information asymmetry, hence the factor A is null and no transaction. In keeping with our main postulate, we insist that all transactions are subject to some degree of information asymmetry, only the degree varies.

The above considerations can be easily coded into a simple formula:

$$q(Q, \epsilon) = A(Q - AQ_0)/(1 - AQ_0), \quad (2)$$

where A is the above mentioned confidence factor. Note that $q(Q_{\max}, \epsilon) = A$, for any ϵ . In proposing the above formula, we try to keep the functions simple and whenever possible, linear functions are preferred. $Q_0 = 7/8$. The choice of parameters seems to be somewhat contrived; this is because the formula is a special example of the full expression $q(Q, p, \epsilon)$, to be discussed later.

2.2 supply side

We assume that the firm knows only its profit margins per unit product if sold at a given price. At the fixed constant price as we have assumed before, his per-unit-profit (revenue less cost) is taken as a linearly decreasing function of Q ,

$$g(Q) = 1 - Q. \quad (3)$$

Q has the general range $[0,1]$ but we shall further limit its range so that the profit $g(Q)$ is non-negative. We assume that the firm knows perfectly the quality of its product, thus no ϵ -dependence for $g(Q)$. We plot $g(Q)$ in Fig.2. Note that the above $g(Q)$ does not depend on quantity to be produced. It’s straightforward to relax this assumption to consider quantity-dependence: a custom-made item has a different per-unit profit than mass-produced items. We assume that the firm’s objective is to maximize its total profit, rather than merely the per-unit profit. If the firm can decide its product’s price and quality, then the quantity must be decided by the market.

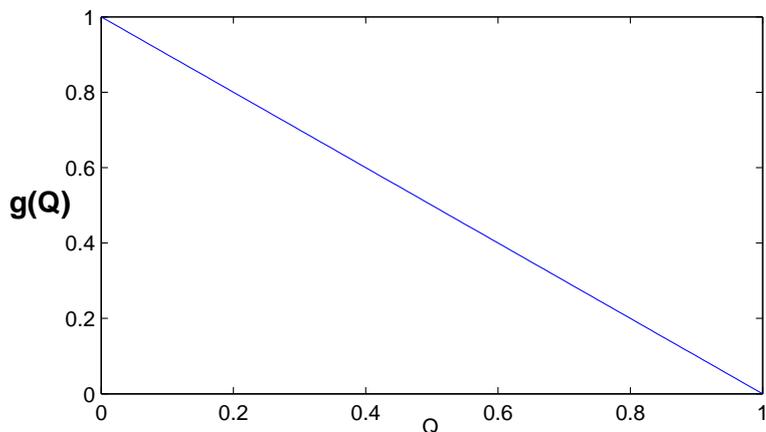


Figure 2: *The per-unit profit for the firm.*

2.3 market

We are ready to take the demand and supply sides to meet in the market. The firm's total profit from a given type of product is defined as:

$$G(Q, \epsilon) = q(Q, \epsilon)g(Q). \quad (4)$$

Whereas the total profit, or total gain to the firm is G , we can denote by $B = B(q, \epsilon)$ the total benefit that accrues to consumers. In the simplest form B is proportional to the quality of the product bought, conditional on the buying probability. In general there can be a constant factor β in the definition, which we choose to be $1/2$ for convenience.

$$B(Q, \epsilon) = \beta Qq(Q, \epsilon). \quad (5)$$

Why is B not just Q ? This is because the good quality of a product is not enough for consumers: if they have a lousy ϵ hence negligible quantity q actually bought. Therefore both conditions must simultaneously be satisfied: the product being bought (large q) and with good quality (large Q), in order to be enjoyed. Other types of B can be proposed, however our main conclusions will not be affected. Whereas G depends on Q in a non trivial way, B is a monotonically increasing function of Q . In fact we have $B \sim Q^2$. For our understanding of how this market works let us point out that the total economic pie is the sum $G + B$, though the firm only consider

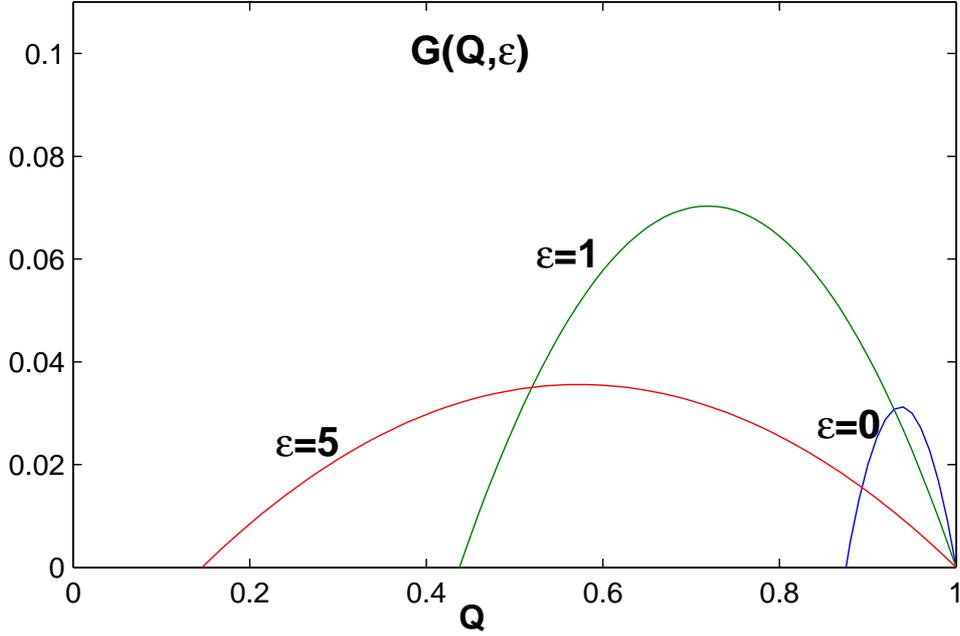


Figure 3: *The firm's total profit G as a function of Q for three values of ϵ .*

G -maximization as its motivation. In the current model, the firm has the production initiatives.

We plot the total gain $G(Q, \epsilon)$ versus Q for various values of ϵ in Fig.3. It is a quadratic function in Q with downwards parabolic shapes. A rational firm would attempt to maximize its total gain $G(Q, \epsilon)$. If consumers have some capability of seeing the underlying quality, the firm will not be too 'greedy' to sell a product with too low a quality, lest too few products be sold. On the other hand, it would not 'charitably' sell a product with too high quality since the (slightly) increased output cannot offset the diminished per-unit profit margins. Therefore facing a given ϵ , which characterizes the current information relation, the firm would choose a suitable Q to maximize $G(Q, \epsilon)$. Since $G(Q, \epsilon)$ is a quadratic function in Q , there is a unique maximizing solution, denoted by $Q^*(\epsilon)$ and it turns out to be $(1 + AQ_0)/2$, as a function of ϵ . Let us assume that the firm alone decides at which Q to make the product and the rational firm desires to produce exactly at the maximizing point $Q^*(\epsilon)$, at whatever ϵ as an external condition.

We find $q^*(\epsilon) \equiv q(Q^*(\epsilon), \epsilon) = A/2$ and together with $Q^*(\epsilon) = (1 + AQ_0)/2$

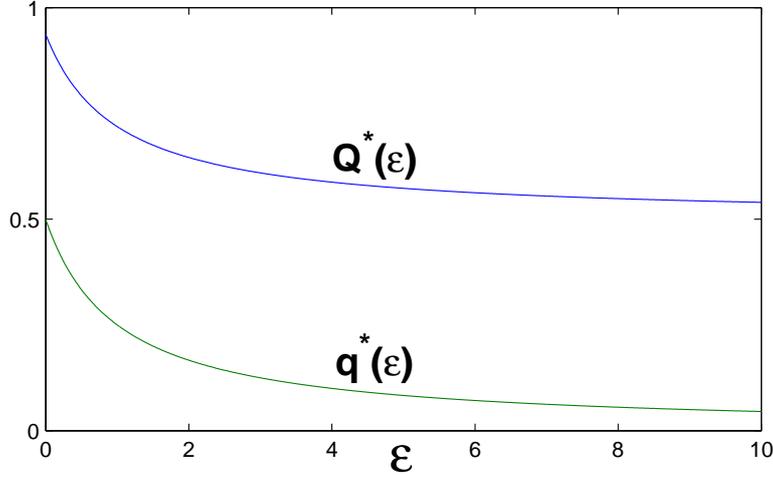


Figure 4: *Both quantity and quality decrease as information capability deteriorates (ϵ increases).*

both are monotonically decreasing functions with ϵ , as plotted in Fig.4. The corresponding total gain $G^*(\epsilon) \equiv G(Q^*(\epsilon), \epsilon)$ turns out to be $A(1 - AQ_0)/4$, as plotted in Fig.5. As information capability increases (i.e. ϵ diminishes), $G^*(\epsilon)$ initially increases. There is an inflection point at $\epsilon^* = 2Q_0 - 1$ ($G^*(\epsilon^*) = 1/(16Q_0)$). Reducing further ϵ beyond this point makes $G^*(\epsilon)$ decrease rapidly. This is a remarkable effect that a very high level of information capability leads to the firm's reduced profit; whereas when very low it is also detrimental to the firm.

2.4 information capability

So far we have treated ϵ , the central parameter of our model, as if it were a god-given, external parameter. In the real world, both the producer and the consumer have some influence in its value. Even though by definition $1/\epsilon$ represents information capability for consumers to discern a product's quality, a producer certainly has options to either facilitate or hinder the task. Both demand and supply sides can make ϵ change. The firm can market its products more or less transparently; consumers with different experience and search effort can also impact its value. The simplest assumption posits that the general ϵ is the product of two factors. Neither Producers nor consumers

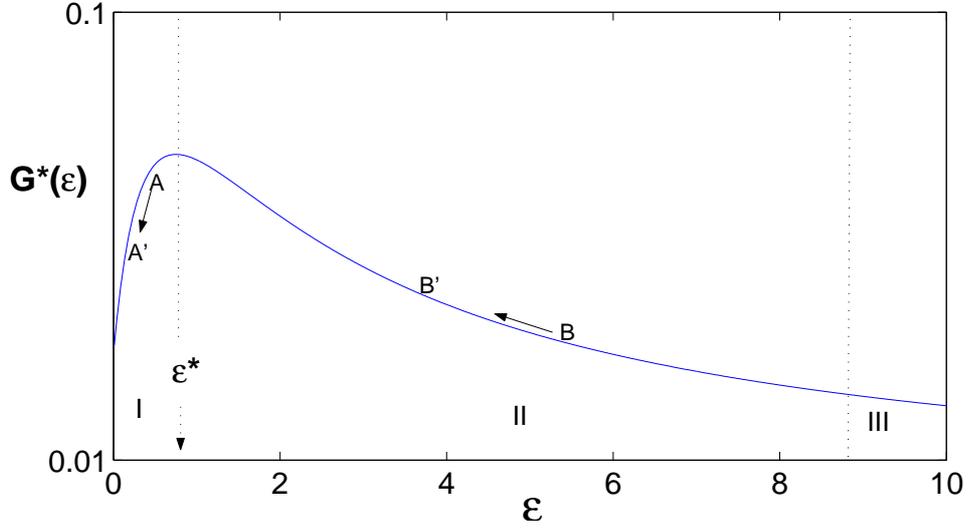


Figure 5: Given ϵ , the optimized $G^*(\epsilon)$ can be divided into three regions: I–defensive; II–collaborative; III–innovative. An overall increase in information capability, as described in (7) will move product A to A', B to B'.

have monopolistic control of it, but rather both exert a mutual control.

$$\epsilon = \epsilon_{\text{firm}} \cdot \epsilon_{\text{consumer}}. \quad (6)$$

Given that both sides can influence ϵ , let us examine who would like to move it in which direction, if they can. Let us consider what the producer and consumers would do to influence ϵ . In region II of Fig.5. reducing ϵ is to the interests of both the producer and consumers. Each side would use whatever leeway available to them to reduce ϵ_{firm} or $\epsilon_{\text{consumer}}$. Look at region I of Fig.6. $G^*(\epsilon)$ decreases with ϵ while $B^* = B(Q^*(\epsilon), \epsilon)$ increases at the same time. We conclude that while in region II the interests of both sides overlap; in region I they conflict.

Let us examine the implications of the above conclusion. When a new product comes to market, in the beginning phase we should expect the corresponding information capability to be relatively low. Thus in region II the producer in his own interest would make effort to reduce the difficulty of detecting quality, i.e. to educate and help consumers. In region I, on the other hand, it is rational for the firm to exercise its available marketing tactics to make information about the product's quality more opaque, to induce errors

by consumers. In this region the two sides' interests diverge. The firm's behavior in region I can be termed defensive.

The above observation has serious consequences on the current thinking of marketing practices, as it is in serious conflict with the position held by many mainstream economists. Among the most important marketing tools is advertising. In the economic literature there are two major categories of advertisements recognized: informative and persuasive[8]. Deceitful tactics like disinformation, various marketing gambits like loss leaders, and other questionable marketing tactics are rarely accounted for in the academic literature. This is because in neoclassical economics the role of information is downplayed. For example in the famous paper 'De gustibus non est disputandum'[9] by Stigler and Becker, advertising is seen as communicating information about a product from the firm to consumers, and both sides seem to have their interests always aligned and never was there a question about the consumer information capability. The possibility of distorting information is not accounted for. Our above results show, however, that whereas in some regions both sides' interests do converge, in other regions their interests may be opposed to each other. Thus one side may spend effort and/or money to penetrate the foggy information veil; the other side may spend to hinder the former's effort. Such mutually destructive forces may have profound implications in understanding the modern economy.

The convergence or divergence of these interests is not a black-and-white story, the degree of severity is the key. In Fig.6 we depict together $G^*(\epsilon)$ and $B^*(\epsilon)$. A more detailed picture can be obtained by considering the derivatives, which show the intensity with which the profit G^* and benefit B^* change while reducing ϵ —and hence the incentives for these changes. $-\partial G^*/\partial\epsilon = A^2(1 - 2AQ_0)/4$ and $-\partial B^*/\partial\epsilon = \beta A^2(1 + 2AQ_0)/4$. The minus sign is because we are interested in the reduction of ϵ . Note that $-\partial B/\partial\epsilon$ is larger and grows faster than $-\partial G/\partial\epsilon$, implying that the 'marginal gain' for consumers with a given amount of ϵ reduction is larger than that for the firm. Nevertheless the producer is cooperative in region II; the two sides would work towards the same goal, even though improved information capability accrues more benefit to consumers and the firm faces diminishing returns. In region I the two sides strive to annihilate each other's efforts. This represents a negative-sum-game embedded in a larger positive-sum-game.

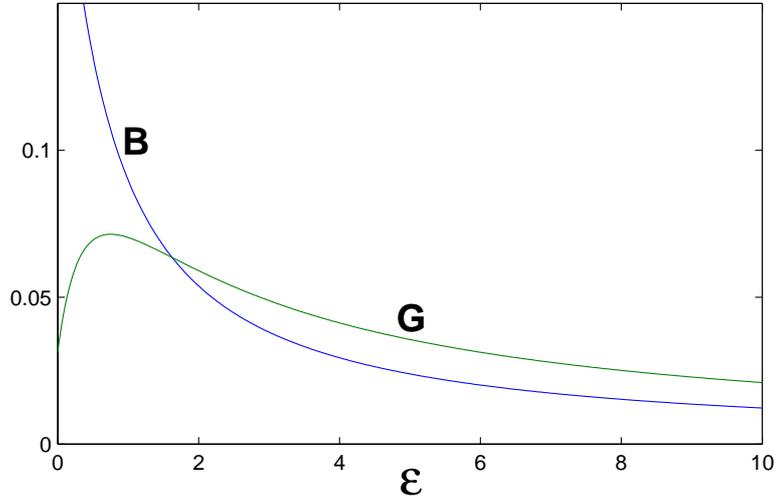


Figure 6: *The consumer benefit $B(\epsilon)$ and the firm's total profit $G(\epsilon)$ can be said collaborative in region II whereas conflictory in region I.*

2.5 magic pie

It is instructive to visualize the above results using pie-charts. They give us an instructive picture of what is at stake. We define the total economic pie $\Sigma = G + B$ and let us see who gets what share (G versus B), for various values of ϵ .

In Fig.7. we show three pie-charts, one for region I, then one at ϵ^* , and one for region II. The total economic pie size Σ is represented by the area of the entire pie. At some point in region II when the pie is small since ϵ is large, the firm would try to increase his pie-slice which is his real motivation. In so doing he also promotes the global welfare, i.e. the total economic pie. As a matter of fact from Fig.6. we learn that his efforts of promoting his own welfare would benefit more the other side, consumers. He is less enthusiastic in further reducing ϵ when approaching nearer the inflection point ϵ^* . At this point he has the highest total profit and if possible he would like to remain there forever. However, the total welfare is only maximized at the ideal limit $\epsilon = 0$. Further information improvement is against his interest as in this region his pie-slice becomes smaller, even though the total pie size is larger. He would not hesitate to use whatever defensive options are available to him to make information opaque. The most important lesson to retain is that

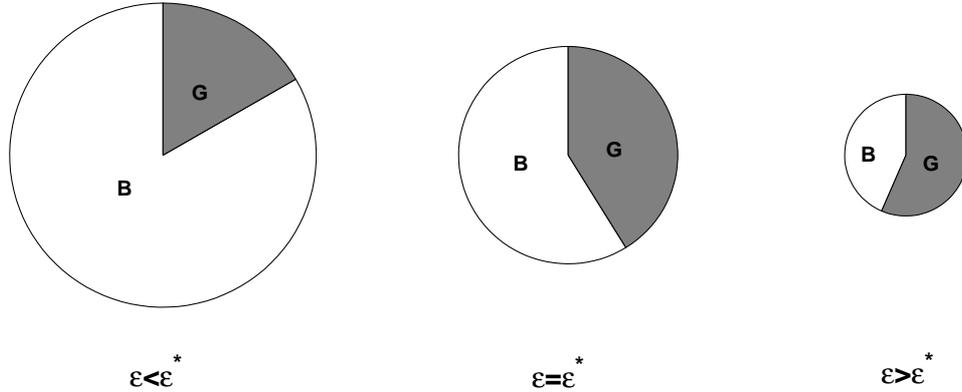


Figure 7: $B^*(\epsilon)$ (white) and $G^*(\epsilon)$ (shade) as shares of a pie-chart, for three different values of ϵ . The total area represents the total economic pie, which depends on how the pie is divided. With low information capability, the firm's relative share is large but the total pie is small; with very high information capability the total pie is much larger but the firm's share can be also small. the optimum sits somewhere in the middle. At a fixed ϵ value and for non-optimized $G(Q, \epsilon)$, we have also a larger pie but a small G -share for $Q > Q^*(\epsilon)$; a smaller pie but a large G -share for $Q < Q^*(\epsilon)$

the total welfare of the both sides cannot be maximized: it depends on the information capability and the degree of information asymmetry, despite of the fact that the both sides are as rational and selfish as they can. To improve the total welfare we need better market institutions as well as advances in Information Technologies.

On Fig.5. there is also region III, representing the production and marketing possibilities that are not yet viable due to the scarce understanding by the consuming public. These possibilities represent mostly innovations and may contribute towards a firm's future profit. However, due to very limited information capability they are not yet marketable. The truncation line between regions II and III depends on many details; but in principle a producer knows where he should start the pursuit of novelty and innovation.

Our above approach may seem to lead to an optimal state akin to neo-classical General Equilibrium. We shall see that since our system is not a closed one, there are always potentially viable innovations which are not yet exploited. Information capability should not be treated as a fixed parameter, many factors can make it vary hence any seemingly optimal state is precari-

ous. We shall discuss two main such factors: one is the evolution of general information technology, the other human cognition.

2.6 evolution

First we consider the impact of the information technology evolution may have on the above supply and demand relationship. When a new technology is developed which can significantly increase consumers' information capability, we may expect that there is a global (downward) shift on all ϵ_i , where $i = 1, 2, \dots$ is a label for the i th product. For this purpose it is convenient to consider that the general shift for all products, occurs in the manner

$$\epsilon'_i = \epsilon_{\text{gen}} \cdot \epsilon_i, \quad (7)$$

where $\epsilon_{\text{gen}} < 1$ denotes the general downwards shift. If this happens, a product A in region I would now be shifted to A' in Fig.5. Products in this region are normally mature, relatively simple, and generic. The producer who spent resources defending (distorting information) it before now would face a still tougher battle. The general tendency is against his defensive attempts. Facing diminishing profit margins for the products in region I, the producer may recoup with the products in region II, where he would have an easier job promoting, educating consumers now. This is shown by the shift from B to B' . There is a small fraction of products that would be pushed over the other side of the hill, i.e. from region II to region I. The producer who facilitated product-quality-transparency now would switch tactics and engage in camouflage marketing gambits for these cross-over products.

Most interesting are the potential products in region III. The products previously not viable now may turn out to be profitable opportunities for the firm, with the general information capability enhancing trend. This region represents inexhaustible future innovations and novelties, and keeps our economy open-ended. This is in sharp contrast with the mechanical paradigm adopted by neoclassical economics, which typically defines a problem with well defined boundaries: factors are either immutably included or excluded. There is no place for unknown or barely known factors. With the open-ended region III of unlimited future possibilities and their concomitant uncertainties, our system can never reach static General Equilibrium. What about infinite information capability that the General Equilibrium approach requires? We see that as $\epsilon \rightarrow 0$ there is no well defined static state: mature

products for which ϵ is much reduced and the squeezed profits make the producer abandon it; the innovative varieties previously not viable come to the fore with concomitant larger ϵ . Therefore we always have information huddles to overcome; we must deal with a dynamic process, such an economy will never end up in an equilibrium.

The next factor we consider is human cognition. Newer products always put strain on consumers' information capability. Consumers gradually get acquainted with the product through accumulated experience via repeated use, learning, etc. Thus during the life-time of a product-line, i.e. from its inception to its final demise, we expect the parameter ϵ evolves gradually from large to small. The rate can vary from product to product, and consumer to consumer, but ϵ cannot stay constant in general. This cognitive evolution can be either assisted or hindered by the firm, giving the latter an additional lever in its rational pursuit of profit. In fact many firms use this observation to plot their long term profit strategy: instead of maximizing the profit at every instant during the product-line's lifetime, they find it more advantageous to market it in the initial phase with little or even negative profit. By so doing they accelerate the ϵ reduction it will allow the firm to cash in later, more than enough to offset the initial losses. This amounts to maximizing the total profit over the whole lifetime of the product-line. It is well known to the business community that the long term maximization goal is often at odds with the short term goal. We may propose the following function:

$$G_{\text{lifetime}} = \int_0^T dt G[Q(t), p(t), \epsilon(t)], \quad (8)$$

where G is a function over Q , p , and ϵ (which are yet to be specified functions of time t). For simplicity questions like interest rates, depreciation, etc are temporarily suppressed. The profit objective for the firm is to choose appropriate functions $Q(t)$ and $p(t)$, given the cognitive evolution course of $\epsilon(t)$, to maximize G_{lifetime} . The previous condition of g being non-negative should be relaxed, as temporary loss can be admitted. Note that this is not maximization in the usual sense, as now the choices are in the strategy space: functions, not parameters, are scanned over. The firm must choose appropriate functions $Q(t)$ and $p(t)$ to achieve the maximal long term result, a task far from being obvious even for a rational and competitive firm. Satisficing *à la* Herbert Simon [10] instead of maximizing is probably the way out.

The products and services in the modern economy are often complex compositions of many factors, as Lancaster has advocated in a similar context[11].

A modern product generally contains many such factors which separately demand varying degrees of information capability. In such situation we expect that the producer's marketing tactics to be a mixed bag: genuinely educating consumers about some novel factors of the product, while at the same time distorting information about other factors, all about a single product.

To conclude this section let us emphasize again the need of treating quality as an independent variable. We may have general impression that our information is roughly adequate for the current economy, we are not cheated outright on most economic transactions, trifling as well as vital. After all, even if we were empowered with perfect information capability, as a consequence improvement to our welfare may not amount to much. This impression may justify regarding neoclassical General Equilibrium as a more or less adequate approximation to reality. The fallacy in this reasoning lies in the fact that current economic offerings are the results of our limited information capability; the demand and supply sides have co-evolved to a state in which we find our current economy. This implies that we won't find our information capability outrageously inadequate of the task: such offerings are simply not proposed. Improved information may seem to reach only finite extra gains, but maximization is illusory. New frontiers are constantly being opened; what seems to be small now will turn out to dominate in the future. Indeed, information and cognitive capabilities are able to drive the economy evolving with no fixed objectives in sight.

3 Supply side uncertainties

3.1 price vs quantity

As we mentioned in the above section, so far we have limited ourselves to the simplest case where price is held constant. Thus the new variable Q and its reliance on limited information capability can be studied with a minimal algebra. In this section we shall instead fix quality Q and let price p vary. The full complexity of the most interesting case with both Q and p present, and especially the relationship between the two is postponed until the next section. The price dependence can be easily extended as follows,

$$q(Q, p, \epsilon) = 2(1 - p)A(Q - Q_0)/(1 - Q_0), \text{ and } g(Q, p) = (1 + \delta)p - Q. \quad (9)$$

The previously discussed $q(Q, \epsilon)$ and $G(Q)$ obtain for $p = 1/2$ and $\delta = 1$. δ measures the per-unit profit when the product is sold at $p_{\max} = 1$ and

$Q_{\max} = 1$. This profit margin can be set to zero in the extreme case. We shall discuss its role when competition is present in a later section. The ranges for both Q and p are calibrated to be within $[0,1]$, and they are further constrained by the requirement that both $q(Q, p, \epsilon)$ and $g(Q, p)$ being non-negative.

Now we want to fix Q to a constant which is not explicitly expressed. Up to a constant the above q becomes simply $q(p, \epsilon) = A(1 - p)$. Even without the explicit dependence on Q , q still depends on information capability. This is because that consumers with finite information capability are not sure of the product's quality, albeit it is a fixed constant. The firm's aim is still to maximize the total gain $G(p, \epsilon) = q(p, \epsilon)g(p)$, which can be done in steps parallel to those of the last section which we shall not spell out explicitly here.

3.2 supply curve

To make the comparison between our new approach to the standard supply-demand law in the economics literature, we need to consider the relation between price and quantity, the central parameters in the demand and supply curves. The decreasing function $q(p)$ of p can be easily identified with the demand curve in economics. But where is the supply curve? Let us recall that ever since Marshall, there have been controversies regarding the supply curve. A sample of various misgivings among economists can be found in ref[12].

Let us examine the supply curve question in detail within our new formalism. We start with the basic question: can a firm have demand curve knowledge, and to what precision? If a firm knew exactly its customers response function q , then since by definition it knows its own per-unit-product profit, there would be no need at all for the supply curve: its maximizing imperatives supersede any other theoretical constructs. Market equilibrium would be established as well: the firm must weigh tradeoffs between a higher per-unit profit (price) and a higher sale volume (lower price). Thus the supply curve is superfluous. It seems that we may do away with the supply curve.

But we shall see below that a supply curve can indeed play an important role and it emerges naturally for a completely different reason than that assumed in neoclassical economics. Suppose that the firm does not know exactly what its customers' responses would be for a new product. From its

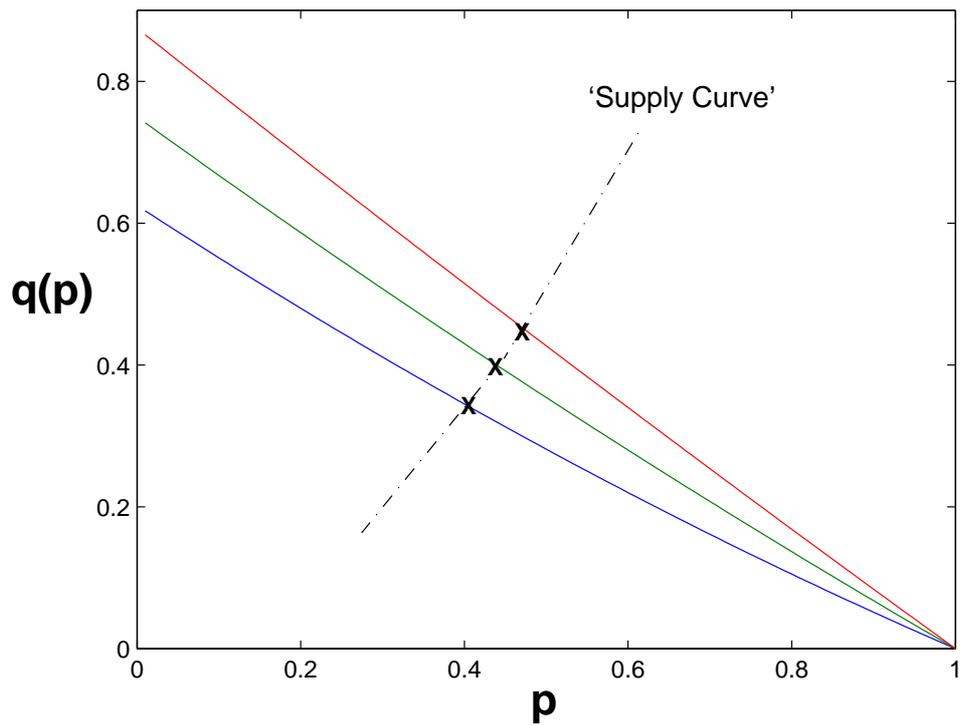


Figure 8: A firm has some assumptions about the demand curve, plotted are three examples. The family of the would-be optimal responses represents a supply curve.

marketing research and previous experience it has some expectations of the function $q(Q, p, \epsilon)$, or the restrictive form $q(p, \epsilon)$. Suppose further that this uncertainty is represented by a family of demand curves, as the few examples depicted in Fig.8. There can be a large number of demand curves in a limited region, each curve does not need to carry the same weight, as the firm may consider some curves more likely than others. But from past experiences the firm decides to be prudent, to hedge with enough contingencies just in case its own production commitment, which might arise from consensus among its executives, goes wrong. Since we assume that the firm knows its own per-unit-profit function, at least better than the demand curves of other people, it must face the maximization task with the above uncertainty about its customers' reactions, *aka* the demand curves. For each of the hypothetical demand curves, there corresponds an optimal price. If the family of the demand curves is taken to be continuous, then these hypothetical prices would form a continuous curve—the supply curve! Of course a continuum of demand curves is a fiction, more likely there is a finite number of hypothetical demand curves, which would yield a supply curve with a dotted line, as depicted in Fig.8.

3.3 firm's uncertainties

Suppose that after the production commitment, the market reveals that one of the hypothetical demand curves turns out to be realized; suppose further that the firm's consensus for the expected demand curve was somewhat different, then as depicted in Fig.9. The firm must take steps to correct the discrepancy. In the example of Fig.9 the firm was over-optimistic in expecting the demand curve and already produced (or committed to produce) the quantity q_f , hoping to fetch a price p_f . It now has some options in finding the best solution out of the less-than-optimal situation (over-production). Suppose the optimum turns out to be (p^*, q^*) , and $q_f > q^*$. The firm may decide to sell all the product quantity q_f in any case to the market, but so doing it cannot expect to be able to price the product at the optimal value p^* ; in fact when the product is sold till clearance, it fetches a price $p_1 < p^*$, so that the pair (p_1, q_f) lies on the real demand curve in Fig.9. However the pair (p^*, q^*) dominates over the pair (p_1, q_f) on the real demand curve by definition of the optimum. This implies that destroying the excess production, assuming such disposal is costless, is a more profitable option than insisting to sell all that is (over-) produced. An alternative option is to insist on selling

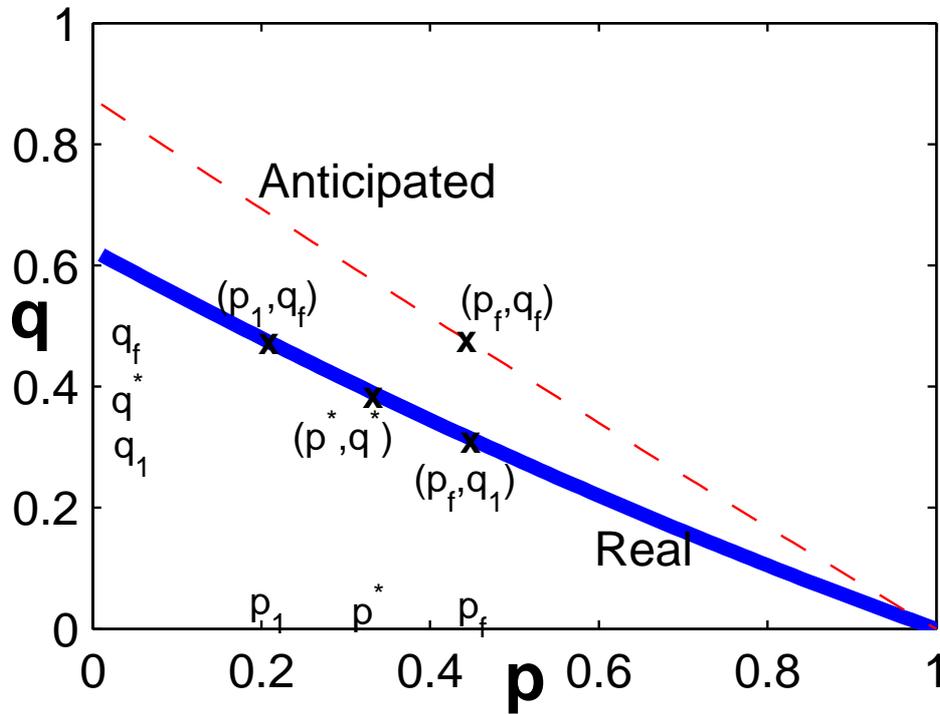


Figure 9: *The consensus-anticipated and the real demand curves. Overproduction calls for corrective measures.*

at the price p_f of the consensus-anticipated demand. This would not help to clear the production, but land us on the pair (p_f, q_1) , with $p_f > p^*$ and $q_1 < q^*$. This is a worse option since the firm needs to dispose of a larger amount of unsold inventory, and the resulting pair (p_f, q_1) is worse than the optimal pair. In sum, for this simple example of discrepancy between the anticipated and real demand curves, the best way out for the firm is to dispose some of the excessive inventory, and then to reap profits at the optimal quantity and price determined jointly by the real demand curve and its own profit function. There are a plethora of other possible scenarios that one can contemplate, depending on the type of mis-expectations and whether the disposal is costly or not. Equally well one can envisage a scenario when the firm under-estimated the real demand curve. In this case there would be missed opportunities.

Let us call the above adjustments in quantity and price $\Delta q = q - q^*$ and

$\Delta p = p - p^*$. To correct the consequences from overproduction, the above firm faces three options: 1) $\Delta q = 0$ but $\Delta p < 0$; 2) $\Delta q < 0$ and $\Delta p < 0$; 3) $\Delta q < 0$, $\Delta p = 0$. We see that the most profitable option for the firm is 2), simultaneously correcting both the quantity and price. Some people might see an amusing similarity with Heisenberg ‘uncertainty principle’ in Quantum Mechanics, where uncertainties over a pair of variables cannot be simultaneously eliminated. But there is no fundamental reason beyond this coincidence.

Besides the unavoidable uncertainty for the firm to estimate the real demand curve, thus necessitating a suite of contingencies that we can call the supply curve. We note that consumers’ tastes and price-sensitivity are always changing as the market evolves; thus the rational firm is always well advised to prepare enough contingencies to cope with the unexpected. The firm must spend both in effort and in financial resources to do substantial marketing research, striving to know better approximations to the real demand curve. However no matter how much it spends, the exact demand curve, which is shifting constantly in any case, cannot be known and clearance does not happen in general. Any marketing research should involve actual tests by selling some amount of the product; this helps to reduce uncertainty. We should expect that there are always either wasted products or missed opportunities—the necessary extra costs due to uncertainties on the firm’s side.

From the above discussion we see that the supply curve arises naturally if the firm’s knowledge about the demand curve cannot be known exactly. The rational firm must take its own lack of perfect knowledge into account by preparing a contingency plan—hence our supply curve. Clearance is in general not guaranteed and additional corrective measures are necessary. Some options are less costly than others but extra costs are inevitable, either in the form of physical costs or lost opportunities, or both. This approach introduces naturally time as a new dimension: as the firm strives to anticipate the demand curve, which itself can evolve, a dynamic sequence should follow. One may note the similarity with Lucas’ Critique and rational expectation theory, where the dynamic sequence plays an important role in macro-economics[13].

In neoclassical economics clearance is assumed. There appears to be two criteria for the market to clear: 1) the traditional demand and supply curves intersect at the market quantity and price; 2) the firm’s supply motivation in the presence of the demand curve and its own costs-profit considerations.

Our above maximizing total gain by the firm is essentially the second type in disguise. This poses a potential logic problem for microeconomics: two criteria may not always be consistent. Of course such inconsistency can be easily avoided by insisting on clearance, and then the two criteria amount to the same thing. Our drastic proposition is to discard the primary supply curve entirely, without affecting the supply-meets-demand law. The trouble we get from this logic is no more guarantee of clearance and subsequent corrective measures are necessary. Isn't it a feature of reality? The secondary, or derived supply curve plays a new and important role as a device to cope with uncertainties about the other side, consumers' demand curve.

By dropping one of the two blades of Marshall's scissors, we lose the mechanical certainty that supply meets demand on fixed tracks. We must take into account non-clearance and non-zero profit margins, questions such as these become the subjects of study, rather than being taken for granted assumptions. Our new approach requires the firm to evaluate a finite numbers of options, none of them optimal, to find the better one. Since information is always imperfect and innovations and uncertainties in unlimited supply, there is no room for mechanical optimization.

Given that firms face also uncertainty about consumers demand, the total profit maximization so far described can be at best approximative. As a matter of fact, the G-function is rather flat near its peak value, the firm will probably choose some value not too far from the inflection point, in reality.

4 Quality Pricing

4.1 the full equation

We finally stand to examine the full fledged relation among the three fundamental variables Q , p , and q . The fact we represent q as a function of other two is just for convenience, we could as well choose any one as a function of other two variables. We shall see that our multiple variables approach brings interesting new features into the supply-demand law; notably, it allows for the first time to price quality. Such pricing is not unique as assumed in neo-classical theory; rather, it depends on information capability, among other factors.

In the economics literature, one rarely examines quality pricing mechanism. Even in the work on Information Asymmetry the underlying quality is

represented by its ‘true price’. This still begs the question: who has priced correctly the used cars in the first place? As if some super-human-being with infinite information capability has determined the true price, with which all other prices constrained on imperfect information should be compared. Take for example Grossman and Stiglitz in ref.[14], for the case of pricing financial assets. To correctly price a stock, these authors recognize that investors may not have sufficient information for the task, so a stock’s price could be away from its fundamental value. Less informed investors may choose to buy the requisite information. In our view this just shifts the pricing problem from one place (stock) to the other (information). How can one be sure the purchased information is reliable, and to what extent? Without completely weaning from neoclassical doctrines one is easily trapped in a circular argument.

We consider the fully generalized quantity function

$$q(Q, p, \epsilon) = \frac{1}{\epsilon}(1 - p)(Q - Q_{\min}(p)), \quad (10)$$

where

$$Q_{\min}(p) = A[(1 - Q_0)p + Q_0]. \quad (11)$$

The factor $1 - p$ is easily interpreted as in the last section and $Q_0 = 3/4$. The last factor needs some explanation: the motivation is that for a higher price the nonzero range of q in Q should be narrower, as consumers are more demanding for more expensive products than they are for cheaper ones. For the firm’s profit we have $g(Q, p) = p - Q$, which is the same as in the last section, except we consider the simplest case $\delta = 0$. Similarly to the simpler case Fig.3, it is instructive to plot the total profit function $G(Q, p, \epsilon) = g(Q, p)q(Q, p, \epsilon)$ for a few ϵ -values in Fig.10. Instead of plotting 3D figures we choose to use gray-scale contours to indicate G -value. For our choice of $\delta = 0$, $\epsilon \rightarrow 0$ is a special limit: the overlap between the non-vanishing regions of q and g vanishes also. q decreases with Q , when the other two variables are held constant, and it vanishes at the borderline. As ϵ increases, the borderline moves leftwards and its dependence on Q becomes flatter. The above $q(Q, p, \epsilon)$ function can be reduced to the simpler $q(Q, \epsilon)$ or $q(p, \epsilon)$ in the last two sections when either p or Q is held constant, apart from an ϵ -dependent factor. Note that the last factor of $q(Q, p, \epsilon)$ has both Q and p dependence. Note that the features in the above construction of q are quite general, other functions can be treated as long as they satisfy the following

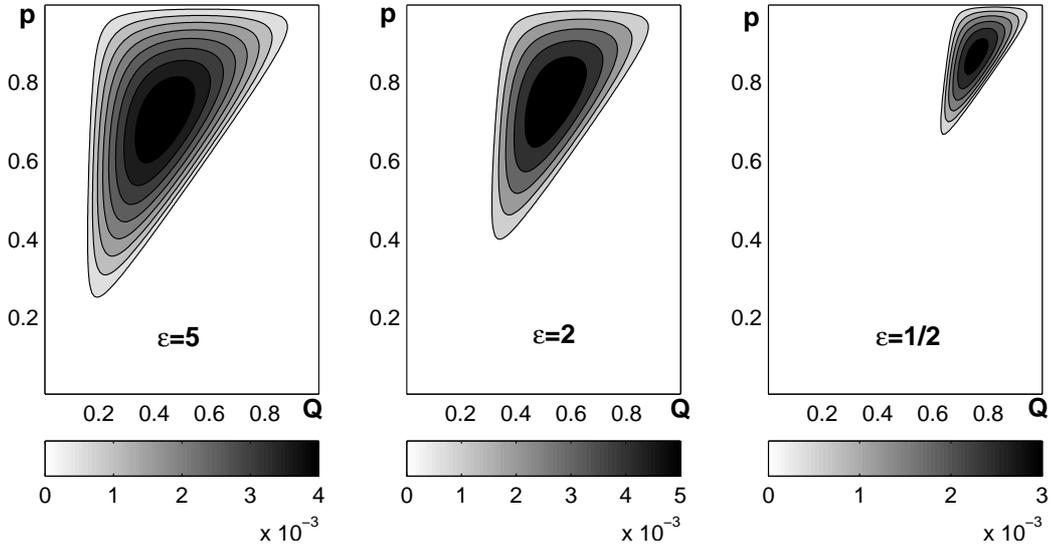


Figure 10: *The total profit $G(Q, p, \epsilon)$ as a function of both Q and p , for given values of ϵ . For high information capability (right), G is concentrated in a small area, with the (implicit) tight correlation between Q and p . For low information capability (left), G is very dispersed and the Qp correlation is very loose. For any given ϵ , the peak value of G gives off the coordinates $(Q^*(\epsilon), p^*(\epsilon))$, which is the solution (13).*

criteria: q is an increasing function in Q , decreasing in p and the borderline Q_{\min} delimiting the range of q is an increasing function of p . The qualitative results are indeed general.

$$G(Q, p, \epsilon) = q(Q, p, \epsilon)g(Q, p). \quad (12)$$

4.2 the solution

Now our firm can choose both Q and p simultaneously to maximize its total profit:

Recall that the range of Q and p is in $[0, 1]$, plus the additional constraints that only q and g non-negative values are admitted. Given ϵ , we can find $Q^*(\epsilon)$ and $p^*(\epsilon)$, by requiring the two partial derivatives $\partial G/\partial Q = 0$ and

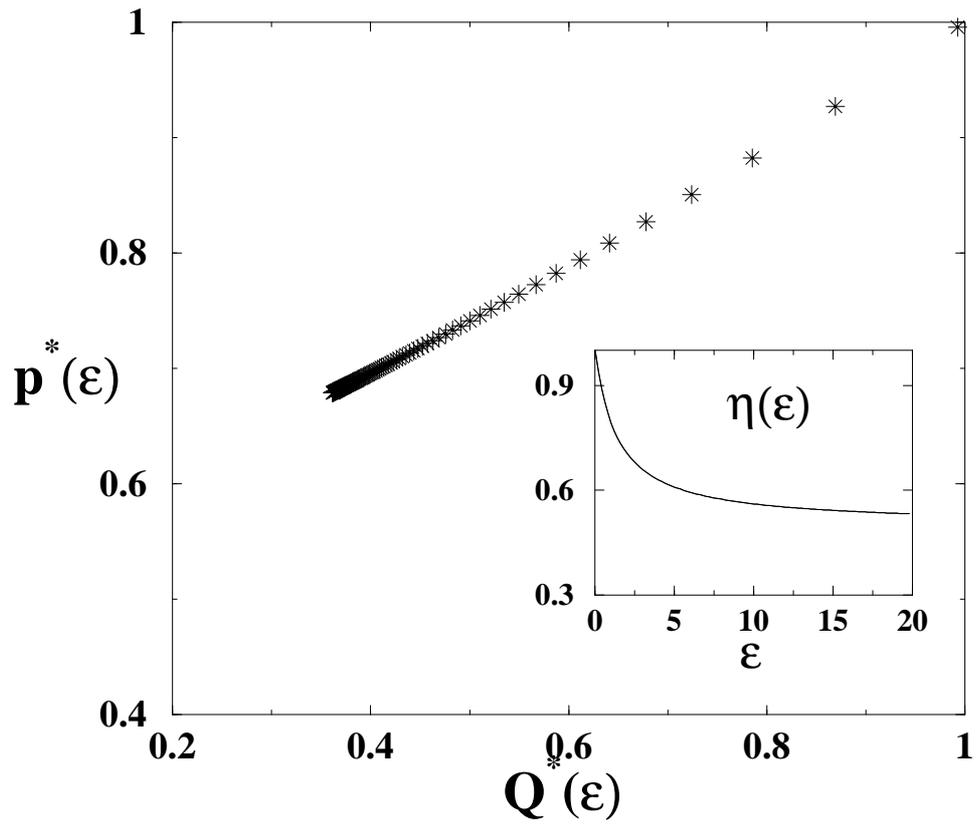


Figure 11: *The solution (13) for ϵ from 0 to 20, with constant intervals of 0.2. In the inset we show the quality-to-price ratio $\eta(\epsilon)$*

$\partial G/\partial p = 0$ being simultaneously zero. The solutions are readily found to be:

$$Q^*(\epsilon) = \frac{1 + 2AQ_0 - A^2 + A^2Q_0}{3(1 - A + AQ_0)}, \quad p^*(\epsilon) = \frac{2 - 2A + 3AQ_0}{3(1 - A + AQ_0)}. \quad (13)$$

To better appreciate the solutions, we plot them in Fig.11 with varying ϵ . We can define the quality-to-price ratio $\eta(\epsilon) = Q^*(\epsilon)/p^*(\epsilon)$, which is also plotted in Fig.11. We see with increasing information capability, (i.e. with an ϵ from a very large value gradually reducing to zero), the ratio η improves from $1/2$ gradually to 1 , the maximal value. Not surprisingly, both price and quality increase with improved information capability, but the latter faster. The myth of quality being uniquely priced has little basis. The ideal limit $\epsilon = 0$ can never be attained. Thus the price-quality correlation necessarily depends on the varying degree of ϵ .

The results similar to those in the previous sections about total profit is readily obtained; substituting the above solutions $(Q^*(\epsilon), p^*(\epsilon))$ back to G , we find

$$G^*(\epsilon) = \frac{(1 - A)^3}{27\epsilon(1 - A + AQ_0)}. \quad (14)$$

Similarly, the quantity actually transacted $q^*(\epsilon) \equiv q(Q^*(\epsilon), p^*(\epsilon), \epsilon)$ can be also found. We plot both $G^*(\epsilon)$ and $q^*(\epsilon)$ in Fig.12.

In fact the asymptotic behavior is of interest: when $\epsilon \rightarrow 0$, to the leading order we have $Q^* \rightarrow 1 - \epsilon(1 + Q_0)/(3Q_0)$, $p^* \rightarrow 1 - \epsilon/(3Q_0)$, $\eta \rightarrow 1 - \epsilon/3$, $q^* \rightarrow \epsilon/(9Q_0)$, and $G^* \rightarrow \epsilon^2/(27Q_0)$. Both q^* and G^* vanish with ϵ , a feature related to the fact that we have chosen the margin profit $\delta = 0$. With competition under limited information capability we expect $\delta \neq 0$, rather G^* approaches a small, positive constant as $\epsilon \rightarrow 0$. With such a constant the algebra can be extended straightforwardly, though the solutions are less compact.

The scenario that we have discussed in the first section can be transplanted here; both sides can influence the ϵ -evolution one way or the other, depending if their interests conflict or not. From the above $G^*(\epsilon)$ we can find the inflection point ϵ^* , and regions I, II, III can be similarly designated. The quantity q^* can be either a monotonical function, or with a peak for a finite ϵ value. The above solutions are simplest when the excessive profit margin $\delta = 0$. For finite and positive δ values, total profit G^* near the point $\epsilon = 0$ is higher than that for $\epsilon = 0$. For sufficiently large δ the inflection point ϵ^* can disappear, so that G^* becomes a monotonical function, in this case the ‘defensive’ region I may disappear.

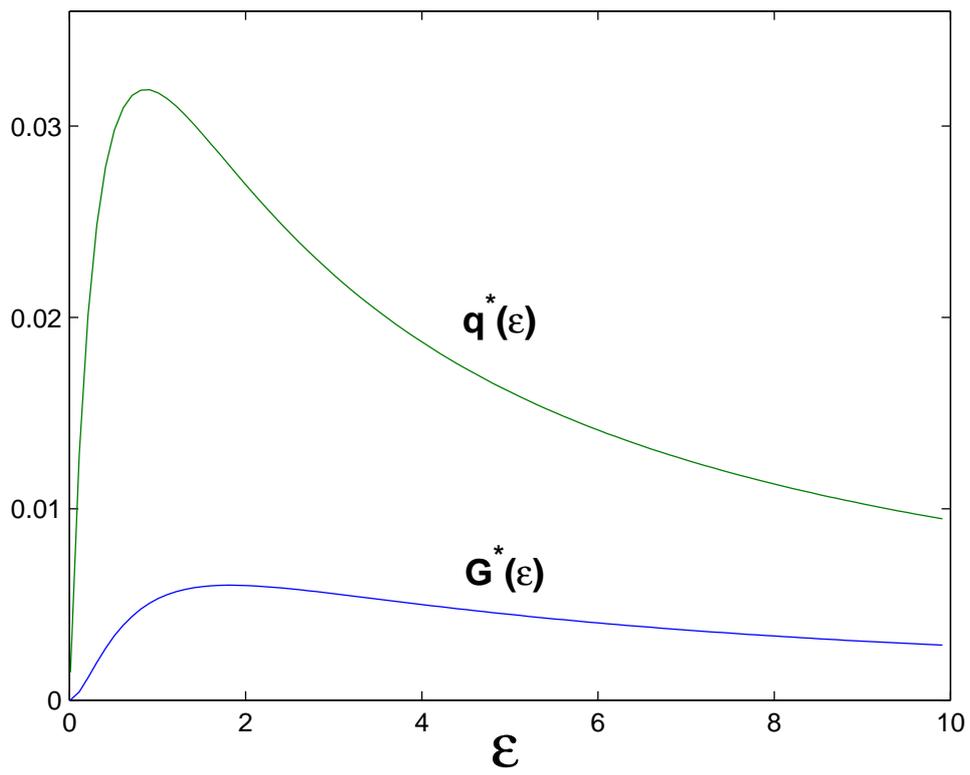


Figure 12: *The total profit $G(\epsilon)$ and quantity $q^*(\epsilon)$ with changing ϵ .*

4.3 partial freedom

Often a producer does not have the same freedom in choosing both variables: in some cases he has his Q more fixed than his p ; in other cases p more fixed than Q . One example of the former case can be that the producer already has his product made, thus Q is fixed, leaving him to vary price while Q is treated as given. In somewhat less fixed Q -cases, his machines for making high-end products already purchased or he has learned skills for high-end products, such long term investments would be wasted if switching to low-end products. So relatively speaking, he has more flexibility in price than in quality, and again we can treat Q as approximately as given.

On the other hand, for the fixed p -case, we may imagine a competitive market where price is well established. A new entrant or just any participant firm would not change price, but has some leeway at tinkering with the less visible factor, quality. The famous ‘sticky’ price puzzle[15] provides another example: often ‘menu costs’ are cited as a reason. But in our view a firm often tries to keep the price invariant as it has many other factors less visible to consumers that it can constantly fine-tune. There is no clear cut criterion as to which variable should be treated as given and which flexible—one must bear the above possibilities in mind when studying real applications.

The above two types of partial freedom can be easily addressed within our formalism. If quality Q is treated as given, then the condition of the derivative $\partial G(Q, p, \epsilon)/\partial p_Q = 0$ yields a maximizing function $p = p(Q)$, for any given Q . If on the other hand, price p is fixed and Q can vary, a similar condition $\partial G(Q, p, \epsilon)/\partial Q_p = 0$ gives rise to a maximizing function $Q = Q(p)$. It is important to note that these two functions are in general not inverse of each other, except at the intersection points. In Fig.13 we show both functions for a few ϵ values.

First we observe that for a given value of p , the corresponding $Q(p)$ is larger for higher information capability (ϵ smaller). On the other hand, for a given Q , the corresponding $p(Q)$ is smaller for improved information capability. For an high-end product ($Q \rightarrow 1$), the price is hardly sensitive to Q , implying the firm has little pricing room here. This hides however an important difference: q is highly sensitive to Q , even near the Q_{\max} limit.

The partial freedom in choosing either p or Q implies some subtle consequences. In Fig.13 we see three pairs of the functions $p(Q)$ and $Q(p)$ for three different values of ϵ . All the intersection points form the solutions (Q^*, p^*) plotted in Fig.11. Let us concentrate (fig.14) on one pair of these functions

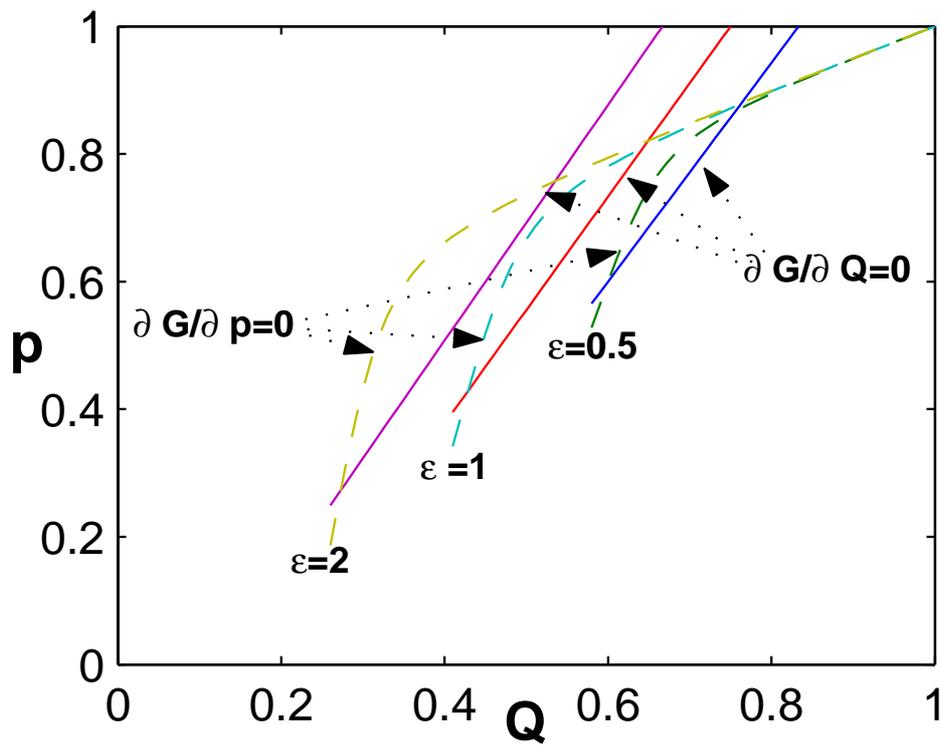


Figure 13: Three sets of $p(Q)$ and $Q(p)$ for three given values of ϵ . The intersection points form the solution (13), as shown in fig.11.

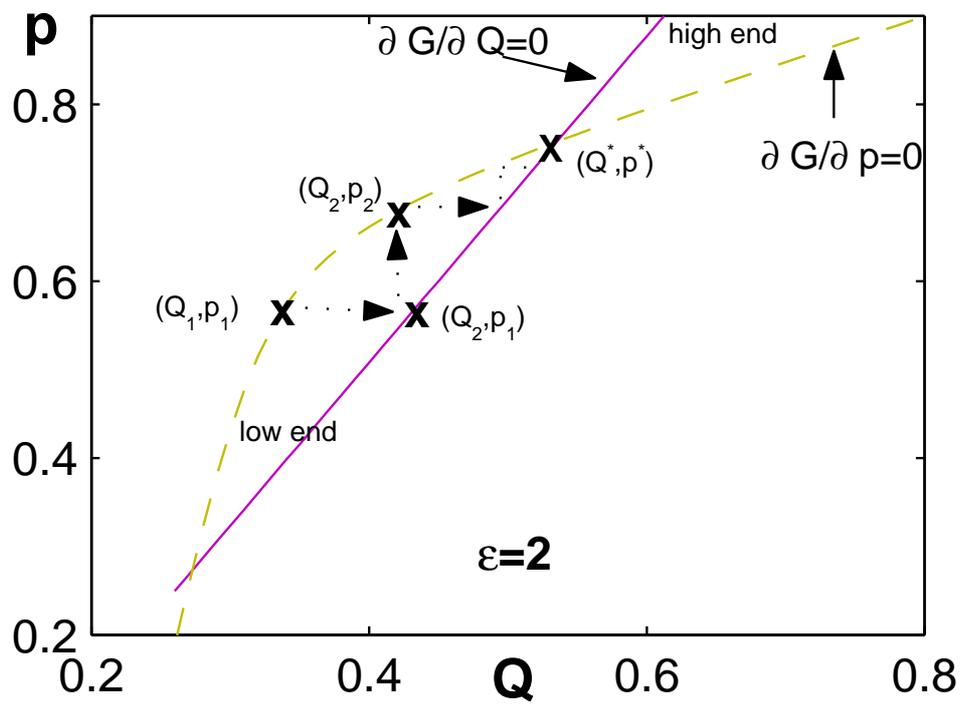


Figure 14: A detailed part of Figure 13, for $\epsilon = 2$.

for $\epsilon = 2$ and call the region below their intersection point as ‘lower end’ products and above ‘higher end’ products. We have singled out three points on these two function at the lower end region, denoted by (Q_1, p_1) , (Q_2, p_1) , and (Q_2, p_2) . Compare now the production possibilities at price p_1 . We see at the same price p_1 , $Q_2 > Q_1$. This implies that if the firm has the freedom to choose the quality Q , it will choose a higher quality to produce for its own profit rationale.

Now compare the production possibilities at the given quality Q_2 . Since $p_2 > p_1$, we deduce that the rational firm will mark-up the price for the same quality, if it has the pricing freedom. This argument can be repeated: if the firm somehow fixes its price at p_2 and is flexible in tinkering with Q , it would choose a still higher quality. We see such iterations lead to the intersection point (Q^*, p^*) at $\epsilon = 2$, as by definition it is the stable fixed point. The above argument is exactly inverted when we consider the higher end region above the cross point. We can draw the conclusion: Facing the handicap of partial freedom, the firm will tend to shift whichever variable is free upwards in the lower end region, to either mark-up price when quality is given or to offer better quality when price being fixed. In the upper end region, the rational firm will tend to do the opposite: to either mark-down price or to cheat on quality. All this is due to the overall imperative to maximize total profit with partial freedom to do so.

The above correlations $p(Q)$ and $Q(p)$ are preordained by the condition that total profit is to be maximized, while one of the two variables is held constant. This is not the only way to obtain correlations between Q and p . We may imagine in certain circumstances the total output (i.e. quantity $q(Q, p, \epsilon)$), is to be held constant; this imposes another type of correlation $p(Q)_q$ or $Q(p)_q$, the subscript denotes the fact that the relation arises conditional on q being held constant. The two functions now are inverses of each other, in contrast to the correlations induced by the max- G condition. Mathematically, we may likewise induce other correlations by holding one of the three variables q , Q , and p constant; we would obtain a function of a single variable and it is invertible. Whereas some of these conditions may be relevant under special market circumstances, in general we do not expect these correlations to play a very significant role in the real world, as the max-profit remains the main concern for the firm. Therefore most attention should be on the results from the max- G condition.

5 Competition leads to efficiency?

So far we have analyzed our model as if there is just a monopoly. Competition must be addressed now and we shall see imperfect information on consumers' part demands a careful re-examination of the traditional competition analysis; for finite ϵ the competition cannot reduce the profit margins to null. Let us consider now what competition would do to the profit margins $g^*(\epsilon) = (1 + \delta)p^* - Q^*$.

5.1 cournot competition

We consider the simplest case where two competing firms produce at the same price and same quality (the more general case when firms compete with different prices and qualities will be relegated to a full discussion elsewhere). Denote by q_1 and q_2 quantities produced by firms 1 and 2, the market composed of the two firms has total output of $q = q_1 + q_2$. Similar to Cournot competition, from $q = q_1 + q_2 = q(Q, p)$ one can solve for $p = p(Q, q_1 + q_2, \epsilon)$. Firm 1 decides its q_1 assuming q_2 is given. The total profit for firm 1 is $G_1 = q_1 g(p(Q, q_1 + q_2, \epsilon), Q)$, which is to be maximized varying q_1 and Q , with ϵ , and q_2 as given and p as the above function. We need to solve two conditions: $\partial G_1 / \partial q_1 = 0$, and $\partial G_1 / \partial Q = 0$. Likewise we can solve for G_2 with respect to q_2 and Q . The calculation is straightforward and we shall not bother to present all the expressions, except to state the conclusion. The resulting $q^{**} \equiv q_1^* = q_2^*$, Q^{**} and consequently $p^{**} \equiv p(Q^{**}, q^{**} + q^{**}, \epsilon)$ should be compared with the single firm monopoly case Q^* , p^* , then the quality-to-price ratio $\eta^* = Q^*/p^*$. The competition of two firms with the same Q and p leads to enhanced market efficiency: in general we obtain the ratio $\eta^{**} = Q^{**}/p^{**} > \eta^*$, confirming the general expectation that competition leads to better efficiency. Much more space is needed to address fully the complex relationship when a finite number of firms competing with each freely choosing their q 's and Q 's, as well as the cases analogous to Bertrand price competition.

We may naturally ask the question: if an infinite number of firms compete in a given product niche, would the profit margin be driven to null, as neoclassical economics affirms? Granted that an infinite number is a fiction, so consider a weaker version of the above position: would a large enough number of competitors make the profit margin negligibly small? We shall see that the question is far from settled, and the answer is actually nega-

tive: no matter how large is the number of competitors (including literally infinity), the profit margin cannot be driven to null, as long as information capability remains limited.

The above analysis of straightforwardly generalizing Cournot competition contains a subtle fallacy, which emerges naturally when information capability is limited, i.e. $\epsilon > 0$. The reason goes as follows. A consumer with a certain level of information capability parameterized by $\epsilon \neq 0$ faces two or more competing firms touting similar products. Her information capability is a combination of both her cognitive abilities as well as effort she must spend each time she inspects a product. Facing two or more products her information capability must be split among competitions. This splitting of her already limited information capability need not be linear, but in general her per-unit information capability is worse than that when there was only a single product. The above analysis about the generalized Cournot competition seems to indicate that her overall welfare is always improved; and when the number of competitions is infinite the firms' profits are driven to null. But we must include the negative effects due to worsening information capability.

5.2 information capability spread

Suppose that the number of competitions is very large. A consumer must then decide whether to spread her already limited information capability very thin on all the competing products, before making a choice, or to simply inspect only a small fraction out of all the available products. In a previous work [16] we have developed a model and obtained a general result that facing the dilemma: either to spread information capability very thin, or to concentrate on very few or even a single item with relatively large information capability. It turns out that the best strategy is to consider a finite number of competing products, neither too large nor too small. The analysis was done for considering investment opportunities but the same mechanism can be readily transplanted here for competition in consumer products. With an optimal, finite number of competitions to inspect, her welfare is best served. Even if the competition number becomes infinite, she would still stick to a finite number of them. They can be picked randomly from the large set of all competitions, or from other factors like geographic locations, marketing gambits, etc. Competition for the firms is both a bliss and curse: bliss because thinly spread information capability make the defensive mea-

sures more effective; curse because usually competitions drive down profit margins. Competition under Luckily for them, consumers cannot take full advantage of competitions for lack of perfect information for detection. We can now draw the general conclusion: since only a finite number of competitions is relevant for a rational consumer with limited information capability, the profit margin can never be driven to null, even when there are infinite number of competitors.

6 Heterogeneous consumers

Most interesting for us is to consider consumers as heterogeneous in their information capability. We thus relax the simplifying assumption that all consumers have the same ϵ , consistent with our working agenda of opening the *Pandora's box* gradually but incessantly. Though we do not associate a monetary value to a consumer's information capability, in general she must spend resources and effort to keep her information capability high. She may be an expert in a particular product niche, able to see through all the marketing gambits by the producers, but if she could get good quality effortlessly, she would prefer to be lazy and careless. Inhomogeneity also arises because of individuals' disposition of abilities and interests can be widely different.

6.1 average vs outstanding consumer

An individual consumer maybe careless in making a purchase decision, but she may expect that *all* other consumers collectively must be doing their due diligence. She may take comfort in thinking that since the firm targets consumers as a whole, the price hence the underlying quality must be right. The firm can fool one or two, but it would be impossible to fool all. We need to inspect in detail such assumptions. First of all, information capability of non-communicating consumers doesn't add up: twice the number of consumers for a given product doesn't double average information capability. Among a large number of consumers certainly there are some experts but their information capability doesn't automatically translate into improvement to the average. However, this can drastically change upon intervention of information matchmakers that can leverage individuals' information capability, especially in the internet age, an issue we shall explore elsewhere. Therefore relevant information capability is on the per-person, per-product

basis, which cannot be improved by the sheer large number of consumers. Besides, each consumers spread her information capability among a myriad of products and service. Firms must always deal with a consuming public with the finite average information capability.

Let us examine what incentives for an individual consumer to depart from the average information capability, if she has such options. First let us consider the simplest case with N consumers having the same level of ϵ_{pack} , except one out of the pack with her $\epsilon' \neq \epsilon_{\text{pack}}$. Facing the split consumers, the producers would target average information capability, i.e. $\bar{\epsilon} \neq \epsilon_{\text{pack}}$, this difference is very small for N large, proportional to $1/N$. Elsewhere we shall show that if consumers' information capability is not much different from each other, the rational response from the producer is still a single product variant, targeting the average. However, as the heterogeneity in consumers' information capability increases over a threshold, the most profitable strategy for the producer is to have two variants: one variant has higher quality-to-price ratio η than the other. If the heterogeneity is still greater, more than two product variants would be needed to maximize the producer's profit. The diversification emerges in a step-wise manner, the full analysis on exactly when the steps happen however is relegated to elsewhere.

Let us check the result of one consumer being ahead of the pack. Her action will raise the average consumers' information capability a little higher, on the order of $1/N$. The net benefit to consumers is the same for everybody, though not much, as a consequence of the improved $1/\epsilon$ value. Thus there are a lot free riders and just one driver: the outstanding consumer contributes to the 'public good' without having any additional benefits to compensate for her more-than-average effort. We know that such public good does not carry sufficient incentives for individuals to contribute, and its fate is like many other public goods—rather precarious without additional institutional settings. However, some consumers may be spontaneously willing to be active 'drivers', i.e. by striving to have smaller ϵ . This is because not all consumers have the same ability, leisure, or personal pleasure of finding better bargains, even without extra compensation. Likewise, we may imagine the opposite case where all consumers having a rather low ϵ value, i.e. all diligent drivers; except one clever but lazy free rider. His reasoning: if I stop driving hard, the average ϵ wouldn't deteriorate much (proportional to $1/N$), but in so doing I'd save a lot of effort whereas the tiny damage resulted from my shirking would be spread on everybody, so it's still advantageous to do. As one can easily recognize such circumstances in the well-keeping of 'public goods', if

everybody in the pack thinks like the above opportunist, the community's information capability will deteriorate.

Things become very different if consumers are sufficiently heterogeneous that the rational firm would find it more profitable to have different variants in quality and price. Let us consider a simple but instructive case: consumers are divided into two groups, one with low ϵ_{low} , the other with high ϵ_{high} . We consider those in the low group as drivers, and those in the high group as free riders, relatively speaking. The firm provides two product variants, one with higher quality-to-price ratio and the other lower. For the producer it is convenient to mix the two variants, since selling them separately would not be a good strategy: even losers won't go to 'Suckers Mall' all the time. If he could he would target the two groups separately, with a separate (ϵ -dependent) optimal strategy for each group. But our two groups of consumers do not live in two segregated locations. The best the producer can do is to mix the two variants. His reasoning: since there are some suckers (free riders) but I don't know who they are, if I sell only the high η -ratio variant these free riders would really get a free lunch. On the other hand, if I sell only the lower η -ratio variant, the more informed consumers would shun my product—a less than optimal strategy. By mixing the two variants I can hope that the less informed consumers who would be relatively blind to the distinction, will have equal probability to stumble on the lesser product; whereas the more informed consumers who can relatively easily differentiate the quality-price difference and preferentially single out the better variant, will not be alienated.

6.2 doing well by doing good

Let us consider one individual consumer in one of the two groups. Suppose she is in the group of the less information capable. Suppose she has some possibility to become an informed consumer if she wants to spend a certain effort (again no need to price effort here). By striving she manages to switch to the more information capable group, thus the average ϵ for the population is improved a little ($1/N$), therefore like before there is some benefit accrued to all consumers. The most important observation is that she does get an additional reward that goes with her striving! By switching groups she has now an appreciable gain by getting a better quality-to-price ratio; at the same time she did a social service by improving average global information capability. Her own additional benefit is some constant independent of N ,

which is the difference between two consumers belonging to the two different groups. So she gets a relatively big benefit (of the order $O(1)$), while doing a small service to each fellow consumer in the population (of the order of $1/N$, but the population as a whole receives a finite contribution of the order $O(1)$). If more people likewise switch from the ‘passive’ mode to the ‘active’ mode, each would get an appreciable compensation for their effort, while contributing at the same time to the population. We see by this logic there is reason to expect the ‘Invisible Hand’ can motivate driving: whoever in her own judgement can play a contributing role will get compensated as well as inadvertently promoting a public good. If information capability levels are in region II (Fig.5), even the firm benefits from her action.

The above analysis can be readily borne out by following our mathematical prescriptions, but the qualitative conclusion is quite general. The above problem was first examined by Tibor Scitovsky, in his well known book *Joyless Economy*[17]. He considers the dilemma faced by consumers, to be a free rider or a driver? His conclusion is rather pessimistic: consumers choose to be free riders as there is no individual reward for their striving. His solution is to appeal to government agencies for safeguarding the vital public good—information. Our conclusion is opposite: sufficiently-inhomogeneous consumers call for the rational producer to respond with multiple variants, which in turn would make individuals’ striving rewarded with extra benefits, thus the welfare of consumers may be taken care of by market forces also, consistent with the basic motto in the capitalist economy: ‘doing well by doing good’. Facing product diversity in quality-to-price ratios, more informed consumers start to be discriminating, picking better ones for themselves, albeit through some extra effort. In other words they don’t believe the price is very communicative. On the contrary, neoclassical economics advocates that price is perfectly communicative, and nobody can make extra gains and there is no ‘driving’ worth doing. This fosters indolence, and if consumers were really convinced of such doctrines they would stop driving even if they could, with the end result of denying a vital driving force behind the capitalist economy.

7 Asymmetric Costs of Failure

7.1 generalist vs specialist

Since we consider all economic transactions to be subject to asymmetrical information of varying degrees, we need to examine its origin. We shall see that such asymmetry is actually the result of another type: asymmetric costs of failure. Our approach to the law of supply and demand treats buyers and sellers in an asymmetrical manner: consumers have a buying predisposition depending on quality and price; firms strive to make maximal profit from such given demand functions. On the other hand the standard neoclassical supply-demand law is symmetrical for both parties, both sides with fixed propensities and their intersection point yields market price and quantity. Is there any fundamental reason to treat the demand and supply sides asymmetrically, or is it just a convenient device? We need to address the question in this section. For this we must take a detour to examine the causes of information asymmetry.

It is convenient to portray the more informed side as more calculating, and the less informed side less calculating. Though the intrinsic cognitive capability can be the same for both producers and consumers, the producers need to concentrate on a narrow niche whereas consumers must spread thinly their information capability among many niches. In another work by Scitovsky[18], he systematically portrays consumers as ‘generalists’, whereas firms as ‘specialists’. It is not surprising that consumers, in their conduct of daily life, need a myriad of products and services; there is just not enough information capability for them to be too calculating on every single niche. Firms, on the other hand, must be more focused on a narrow range of products and services. The decision for a consumer to buy or not buy a given product is not as important for her as for the firm producing/selling it—the latter’s core business. The firm cannot afford to be as careless and sloppy as consumers in all the details regarding a product. A firm’s stake is much higher than that of consumers for a given product. A consumer can fail to buy a good product, or be cheated through another purchase, it is only a small part in her overall consumption life. Therefore we must face the fact of ‘asymmetric costs of failure’ of the two sides.

The principle of asymmetric costs of failure is a well known concept in biology. Dawkins in his book [19] *Selfish Gene* explains this principle vividly using the dinner/life metaphor: when a fox hunts a rabbit, the rabbit usually

outruns the fox. The rabbit runs faster in general since its stake is higher than that of the fox. The fox can permit itself sometimes to be a bit sloppy in hunting, as a missed dinner for the fox is less fatal than a missed escape for the rabbit. Like the rabbit, a firm is obliged to investigate more, since it can ill afford failures. This asymmetry dictates that the firm necessarily must know more about its products than consumers do, to tinker with quality factors and to price carefully. Hence we conclude that the generalized information asymmetry we advocate in this paper has its ultimate roots in the asymmetric costs of failures.

7.2 price taker vs price maker

In the perfect equilibrium all participants are price takers; none has enough clout to influence the price. In the presence of general information asymmetry the price-taking assumption of neoclassical doctrines must be re-examined. If both sides have perfect information and are rational, buyers and sellers can reach the equilibrium price by a ‘Walrasian process’. Indeed there are places in modern markets where professional buyers and sellers find out prices in approximate agreement with a Walrasian scenario.

But most transactions vital to the economy are done by the seller’s posting a price and buyers decide if the deal is good enough to take the bait. Why are sellers most often the price maker? The simplest price making is to post a price; he must make it somewhat acceptable to the other side. Proposing a price inevitably leads to a wait-and-see process, to see if the buyer accepts or not the deal. The asymmetric costs of failure force the seller to have superior information and to be more calculating. Price-making is a complex and more calculating procedure: First you need to figure out your side and the other side’s gain and loss and the likelihood of acceptance, then post a price, and still have to wait to see if the deal goes well according to your expectations or not. If not, either quality and price or both have to be modified to maximize realistic profits. A consumer, being less calculating, simply judges if a deal is good enough, to the extent her information capability allows; she decides either to buy it or to walk away. She bases her satisfaction criterion on a limited number of alternative offers that she happens to have checked out, as well as her limited attention/capability of knowing the real quality behind these checks.

There is another asymmetry between consumers’ wants and possible supplies the firms can make. Relatively speaking, the wants are less flexible

than the production possibilities. Whereas there are only a limited number of wants, there are virtually unlimited possible suppliers, if need be. Apparently when a market clears the wants are met by the supplies—there seems to be a one-to-one correspondence. This apparent symmetry hides an asymmetry, since what is actually supplied is only a tiny fraction of what could be possible. There are much more would-be suppliers than there are would-be consumers. With a large number of would-be suppliers lurking, the supply side normally does the wooing. Wooing leads to proposing, hence price making. The asymmetric costs of failure are also behind this new type of asymmetry, since the failed supplier would be replaced by newcomers in no time. In the real economy, there are littered countless failed suppliers, not that many failed consumers. In some special circumstances the sellers are small and many, the buyers big and few (think of the small banana farms vs multinational banana importers), in this case the role of the asymmetry is reversed and the monosony problem can be similarly analyzed.

7.3 horizontal vs vertical search

Our model of information capability so far can be said to be about *vertical* search for a product: a consumer knows what she wants to buy, but needs digging to ascertain the product's quality. Let us mention the problem of *horizontal* search when a consumer is aware only partially of all the similar products, i.e. before ascertaining a product's quality there is a need to first match the consumer wants to many possible alternatives. To motivate our problem, we consider a simple model: there are N production possibilities (combinations of attributes as well as a price), which are possible products but not necessarily produced. For simplicity we assume that the producer is aware of all of them and a consumer only a small subset of all, $K < N$, chosen randomly. The producer has certain preferences over all the N possibilities, he ranks them from his top choice down to the least preferred one. We consider only the viable products, even the least preferred choice is still meagerly profitable to the firm. In general the interests of consumers are not perfectly aligned with that of the producer, but they are not necessarily opposed to each other. The simplest assumption is that they are 'orthogonal' to each other: the ranking of a consumer is a random reshuffling of the firm's ranking. Two distinct consumers may differ in their rankings about the same set of production possibilities to reflect idiosyncratic tastes, but they are not random and normally are strongly correlated with each other. To a

first order approximation, we treat consumers' ranking as homogeneous, and concentrate on the difference between consumers' preferences and that of the firm. More detailed study should examine the cases when the two sides' interests partially overlap or partially in conflict, as well as inhomogeneous consumer tastes.

Thus we have N distinct production possibilities which are ranked differently by the two sides. Since the consumer knows explicitly only a small fraction of the total, how can she rank the possibilities that she doesn't know? If a production possibility is not explicitly known to the consumer, and if a firm proposes it to her, she is able to recognize it and find a place in her rank list. Thus those not explicitly known to her cannot be proposed by her, but can be passively judged by her if proposed. The above orthogonal relation implies that the top choice on the ranking by the one side occupies only a random place on the other's ranking of preferences. Let us say that the top choice for the producer may not be the consumers' top choice, probabilistically speaking. In the perfect information symmetrical case ($K = N$) where both sides know all the N production possibilities, we can show that both sides must find a compromise part way down their ranking lists, at a distance of \sqrt{N} from their top choices, respectively. Though both sides cannot achieve their top choices, the compromise point is still far better than their average ranking position ($\sqrt{N} \ll N/2$). One can easily show that the above compromise point is the best possible mutually acceptable solution, statistically speaking. When the acceptable solutions are found, sometimes one is proposer sometimes acceptor, with equal probability for both sides. The calculation is similar to that of the "Stable Marriage Problem", a combinatorial mathematical problem [20] well known to game theorists and recently there is renewed interest among physicists [21]. Most interesting is when the consumer side knows only a fraction of the total possibilities; this limit represents a new sort of information asymmetry, hence we label it 'horizontal', to be distinguished from that for the above 'vertical' search parameterized by ϵ . It can be shown that with the asymmetrical information, i.e. consumers are only aware of K of the total possible production possibilities N ($K \leq N$), the compromise point acceptable to both sides tilts toward the gains of producers. The analytical solution, using the so-called mean field method, turns out to be exactly the same obtained previously for SMP with partial information [22]. Let X and Y denote the distances from their top

choice for the producer and consumers, respectively, then we have

$$X = \sqrt{K}, Y = N/\sqrt{K}. \quad (15)$$

(This solution obtains from the more general case when there are total N production possibilities, both producer and consumers know only a random fraction of the total, such that $K < M < N$. We still have information asymmetry since the producer knows more (M) than consumers do (K). Following the reasonings in ref[22] we readily find:

$$X = N\sqrt{K}/\sqrt{(M + K - MK/N)K}, Y = N\sqrt{M}/\sqrt{(M + K - MK/N)M}. \quad (16)$$

When substituting M for N , we obtain the simpler solution (15).)

Most interesting is that the horizontal search model allows to calculate probabilities of active proposing or passive accepting. In the above simplest case, the mutually beneficial transactions happen sometimes from the producer's initiatives, sometimes with consumers' initiatives, the respective probabilities turn out to be $\frac{N}{N+K}$ and $\frac{K}{N+K}$. Thus the ratio between the two is $N/K < 1$, i.e. more likely the producer actively proposes and consumers passively accept. Since price is one of the factors in production combinations, we conclude that facing information asymmetry of the horizontal type, transactions happen with higher probability for the producer to make price than the other way around. The above model is just the simplest example to illustrate the qualitative behavior of the horizontal search problem. In general we should always expect that in a symbiotic relationship that is mutually beneficial to both producers and consumers, their interests will not be perfectly aligned. The above model uses a random, neutral relation between the producer and consumers. In general we can have a continuous spectrum of relations: some are more aligned, some less. If the one side suffers from lack of total available information, the compromise solutions would favor the other side with more information. Most such relations are positive-sum-games, some can happen to be zero-sum-game, still others can be negative-sum-games.

8 Conclusion

In this paper we have presented a simple model for the law of supply-demand in the presence of imperfect information. Compared to the traditional neo-classical law we must introduce extra variables: quality and information

capability of detecting it. Information is not an independent objective variable, rather it resides in the market relation of the two parties. Through this simple model we have outlined a research program that holds promise to be able to address a broad spectrum of economic issues. Our approach can be straightforwardly generalized to other areas where information asymmetry is believed to exist to some extent: like labor market, financial market etc.

Whereas many results of our new approach may not seem to deviate too much from neoclassical ones, the fundamental assumption is different. We consider imperfect information as the crucial starting point and perfection can never be attained. The most relevant issue for our research agenda is not about what is the perfect equilibrium which we posit non-existent, but rather how to find better ones among imperfect situations. Many fundamental concepts so far taken for granted need to be re-examined. For instance, we must face the serious consequence that much cherished Pareto efficiency is no longer relevant in our supply-demand relations in the presence of imperfect information. This is true both for vertical as well as for horizontal search models. The maximal welfare cannot be reached. From the discussion about the pie-charts (Fig.7) one may wonder, why is the maximal welfare solution not preferred? Wouldn't it be rational to realize the maximal pie, then to split the extra gains between the two sides? The short answer is that if the producer and consumers are left to themselves, as in most markets, such utopian dreams are not realizable. To split the extra gains requires another market transaction, which the imperfect information impedes. Current information capability only allows that much pie to be realized as the best possible compromise, according our model. The extra pie is beyond reach by the two sides alone. This is similar to Akerlof's Lemons Problem: the seller of a good used car and his potential buyer would end up in a mutually beneficial transaction, but the asymmetrical information prevents this transaction from happening.

What can one do to improve upon the reality brought about by the ubiquitous asymmetrical information that permeates almost every aspect of our economic life? Even if we could improve information capability, the perfect equilibrium is never in sight: better information capabilities, while improving current transactions, also open up new possibilities that come with innovations. We are led to a new paradigm in which an economy will never settle in equilibrium; new and old opportunities come and go with information technology and improved information institutions.

In the original paper by Akerlof, as well as by other economists, gov-

ernment intervention is generally called for. Licensing, certifying quality by consumers protection agencies may help to mitigate somewhat the imbalance of information. We believe there is also a new type of intervention through market force itself. The big difference between what is possible and what is realizable presents a new type of attractive business niche that entrepreneur may fill for their own profits. This business is the so-called infomediaries that can serve as a matchmaker between consumers and producers. Some prototypes have already become household names, with the advent of the Web: Ebay, Bizrate, Amazon et al. The working mechanism behind these informediaries is to tap into collective information capability of consumers, so that as if an individual consumer is suddenly empowered with much higher information capability. Elsewhere[23,24] we shall explore in detail how these matchmakers work to improve the information capability and the consequences.

Acknowledgements: For helpful discussions and comments I thank Mauro Gallegati, Paolo Laureti and Hassan Masum.

Reference:

[1]A. Marshall, (1920) *Principles of Economics* (Eighth Edition), London: Macmillan Press.

[2]C. Babbage, (1846)*On the Economy of Machinery and Manufactures* (Fourth Edition), London: John Murray, quoted in G.M. Hodgson,(1993) *Economics and Evolution*, p.115, Ann Arbor: Univ. of Michigan Press.

[3]C. Leslie, (1879) ‘The Known and the Unknown in the Economic World’, originally published in *Fortnightly Review*. Reprinted in *Essays in Political Economy*.(1888), p221-242 London: Longmans, Green, & Co.

[4]G. Akerlof, (1970) ‘The Market for ”Lemons”’: Quality Uncertainty and Market Mechanism’, *Quarterly Journal of Economics* **84**(3),488-500.

[5]A.M. Spencer (1977) *Market Signalling: Information Transfer in Hiring and Related Processes*, Cambridge: Harvard Univ. Press.

[6]J. Stiglitz (1977) ‘Monopoly, Nonlinear Pricing, and Imperfect Information: The Insurance Market’, *Review of Economic Studies* **44**,407-430; ‘The Causes and Consequences of The Dependence of Quality on Price’ *J. Econ. Lit.* **25**,1-48.

[7]K. Arrow, (1987) *The PPB system*. Vol. 1. U.S. Joint Economic Committee, 91st Congress, 1st session. Washington D.C.: U.S. Government Printing Office, pp 59-63. 1987; still earlier anticipation can be found in Tibor Scitovsky’s *Welfare and Competition* (1952), London: George Allen and Unwin, where general asymmetric relations in consumer,labor markets

are discussed in details.

[8]K. Bagwell (2001) *The Economics of Advertising* (International Library of Critical Writings in Economics Series) (Ed.)Edward Elgar Publishing.

[9]G.J. Stigler and G.S. Becker, (1977) ‘De Gustibus non est Disputandum’, *American Economics Review* **67**,76-90.

[10]H.A. Simon, (1957) *Models of Man: Social and Rational*, New York: John Wiley.

[11]K. Lancaster, (1971) *Consumer Demand, A New Approach*, New York: Columbia Univ. Press.

[12] P.Sraffa (1926)‘The laws of returns under competitive conditions’,*Economic Journal*36,535; D.Colander (1995), ‘The stories we tell: a reconsideration of AD/AS analysis’, *Journal of Econ. Perspectives*, 9(3), 169; S. Keen, (2001) *Debunking Economics*, London: Zed Books; A. Cohen (1983), *Eastern Economic Journal* 9(3), 213.

[13] R.E.Lucas Jr.(1976)‘Econometric policy evaluation: a critique’, Eds. K.Brunner and A.Meltzer. Amsterdam: North Holland.

[14] S.J.Grossman, and J.E.Stiglitz (1980) ‘On the Impossibility of Informationally Efficient Markets’ *The American Economic Review*, 70(3),393.

[15] N.G.Mankiw, and R.Reis (2002), ‘Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve’, *Quarterly Journal of Economics*117, 1295.

[16]A.Capocci and Y.-C.Zhang (2001) ‘Market ecology of active and passive investors’, *Physica A*298, 488.

[17]T.Scitovsky (1976) *The Joyless Economy: An Enquiry into Human Satisfaction and Consumer Dissatisfaction*, Oxford: Clarendon Press.

[18]T.Scitovsky (1986) *Human Desire and Economic Satisfaction: Essays on the Frontiers of Economics*, Brighton: Wheatsheaf.

[19]R.Dawkins (1976) *The Selfish Gene*, Oxford: Oxford Univ. Press.

[20]D. Gale, D. and L.S. Shapley, (1962) ‘College Admissions and the Stability of Marriage’ *American Math. Monthly* **69**, 9-14.

[21]M.J. Omero, M. Dzierzawa, M. Marsili, and Y.-C. Zhang, (1997) ‘Scaling Behavior in the Stable Marriage Problem’ *Journal de Physique I*, **7**(12),1723.

[22] Y.-C.Zhang, (2001) ‘Happier world with more information’*Physica A*299, 228; P.Laureti and Y.-C.Zhang, (2003) ‘Matching games with partial information’*Physica A*324, 49.

[23] Y.-C.Zhang, (2005) *The Information Economy*, to be published.

[24] Hassan Masum and Y-C Zhang, (2004) 'Manifesto for the Reputation Society', *First Monday*, www.FirstMonday.org Issue July2004.