Available online at www.sciencedirect.com
ScienceDirect
International Journal of Industrial Organization
26 (2008) 998 – 1014
International Journal of Industrial Organization
www.elsevier.com/locate/econbase

# Queues, not just mediocrity: Inefficiency in decentralized markets with vertical differentiation

Clara Ponsatí [a,*], József Sákovics [b]

[a] Institut d'Anàlisi Econòmica — CSIC, Barcelona, Spain
[b] University of Edinburgh, United Kingdom

Received 6 November 2006; received in revised form 23 April 2007; accepted 16 September 2007
Available online 26 September 2007

**Abstract**

We analyze a dynamic, decentralized market with endogenous entry, where in each period the active sellers supply one unit of an indivisible service at varying degrees of quality. The customers that have entered the market are randomly matched with the active sellers and prices are set by (complete information) pair-wise bargaining. In its unique steady state, the market leads to an excess diversity of quality and customers may have to suffer costly delays. Notably, efficiency is not regained as per period delay costs disappear. We also show that setting minimal quality standards, such as licensing rules by a professional college, will improve welfare (and even Consumer Surplus), relative to the free market, whenever the inefficiency is caused by a large enough excess supply.


*JEL classification:* C78; D43; D62; L44; L84

*Keywords:* Liberal professionals; Quality standard; Queuing externality

## 1. Introduction

We model markets where sellers with limited capacity supply goods, the quality of which is heterogeneous and observable only upon entering the store/office. In these markets reputation is not a major factor.[1] This is typically the case for services provided at small establishments over dispersed locations like road restaurants, bed-and-breakfast accommodation, or car repair shops. It is also the case for tradesmen, like carpenters, small builders or plumbers; or middle-of-the-road professional service providers that are hired infrequently, like architects, accountants, lawyers and other liberal professionals. While small individually, these activities play a major role in the performance of advanced industrialized economies. They are quantitatively important in the aggregate[2] and, being an intermediate input for many other activities, improvements in their quality and competitiveness have major spill over effects across the economy.[3]

These markets are often characterized by a very high level of regulation. While specific regulations vary across countries, trades and professions, quality standards, advertising constraints, price limits or other controls are

* Corresponding author.
*E-mail address:* clara.ponsati@iae.csic.es (C. Ponsatí).
[1] Otherwise, quality could be predicted.

[2] According to recent estimates, professional services employ about 6.4% of the EU15 work-force and represent 10% of overall high skill employment; in 2001 the sector had turnover of around 980 billion Euros and created around 500 billion Euros of total value added. See the report on Competition in Professional services COM (2004) 83.

[3] For example, according to a recent estimate by the Italian Antitrust Authority, 6% of the costs in Italian exporting firms are due to professional services.

0167-7187/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijindorg.2007.09.003

extensively imposed either by State regulation or in the form of self-regulation by professional associations. The risk of regulatory capture has been a major concern since Stigler (1971). It has been argued that these arrangements raise prices and limit entry but fail to assure the appropriate quality of service, opening an intense debate among policy-makers on (de-)regulation policies.[4] In fact, since the landmark ruling of the US Supreme Court in 1975 on *Goldfarb v. Virginia State Bar* there is an ever increasingly active application of antitrust laws in the professional services sector in all OECD countries.

The conventional wisdom argument about markets with quality heterogeneity is based on asymmetric information: if at the time of contracting the quality is not observable by the customer, it actually may be the high quality services that are driven out of the market.[5] Consequently, the main focus of policy debates is the correction of market failures due to informational asymmetries. Arguments are often based on the presumption that in the absence of asymmetric information a free market would perform efficiently. The present paper shows that this presumption is not correct; we contend that even if customers could ascertain the quality of service before purchase, efficiency would not obtain. Asymmetric information is an important problem indeed. However, arguments about (de-)regulation policies deserve further scrutiny; understanding the source of inefficiencies abstracting from asymmetric information considerations is a necessary first step.

We develop our arguments in the context of decentralized trade model based on a dynamic random matching market. Matching is random to reflect that the customers cannot tell the quality offered by the different sellers before they choose who to approach. In order to focus on the issues beyond asymmetric information, we assume that customers learn the quality of the service upon meeting a service provider. In short, we model service as a search good. We capture limited capacity by assuming that each supplier can serve only one customer at a time (making agreements on prices also bilateral). We model the resulting two-person bargaining with complete information using a simple strategic model, which can also be interpreted as the — possibly asymmetric — Nash Bargaining Solution.

We characterize the unique steady state of the market, and show that it leads to an inefficient outcome: there is an excess of active sellers leading to a sub-optimal average/marginal quality and in addition, for low enough delay costs, a queue will form and customers will suffer costly delay. A suboptimal provision of quality and delays are standard equilibrium features in markets with costly search. The novelty of our analysis is that we provide a simple model where the removal of the per-period delay costs does not dissipate both these inefficiencies. While the supply distribution does converge to the efficient one, as the delay cost shrinks there is an ever larger queue of customers waiting to be served and the aggregate waiting cost is increasing.

The unique market equilibrium admits a straight-forward characterization, thus comparative statics and policy issues are easily addressed. We highlight that regulatory policies using a single policy tool (such as setting minimal quality standards) can be helpful but they only lead to second-best outcomes. Finally, we show that when the sellers are allowed to self-regulate, welfare (and even Consumer Surplus) is increased relative to the a free market, whenever the inefficiency was caused by too many low quality providers in the market.

Helmut Bester's (1988) model of a decentralized market with quality heterogeneity and bargaining[6] is a direct forbearer of our analysis. Under positive search costs, market equilibria are similar in his formulation and ours. In his model comparative statics are problematic because of the multiplicity of equilibria, so he did not address issues of (self-)regulation. The most important modelling difference is that Bester allows a seller to sell to several buyers simultaneously, with no externalities between trades. As a result, there is no rationing — and, therefore, queuing — in his set-up. Thus, unlike us, he concludes that under negligible search costs any market equilibrium approaches the efficient, competitive, outcome. Chaim Fershtman and Arthur Fishman (1994) examine the effects of a minimum wage in a labor market with costly search. They find that worker search activity may decrease as a result of a higher minimum wage, leading to higher firm bargaining power paradoxically leading to lower wages at ultramarginal firms. While this exercise is somewhat similar to the introduction of minimum quality standards, both the intuition and the effects are rather different. In the more recent literature, Max

[4] See the OECD DAFFE/CLP(2000) 2 report on "Competition in professional services" for a host of evidence on how anti-trust authorities have approached the issue in twelve countries in addition to the EU. See also Paterson et al. (2003).

[5] Uncertainty about service quality (Arrow, 1963) and the resulting adverse selection (Leland, 1979; Stiglitz, 1979; Wilson, 1979, 1980) and moral hazard (Shapiro, 1986) were modelled and understood some time ago.

[6] Bester (1993) extends this model to endogenous quality choice and analyzes how trading rules affect this choice.

Blouin (2003) provides a simple model with experience goods and no new entry after the first period. For the latter reason, waiting is again not an issue for negligible waiting costs. Rachel Kranton (2003) stresses the importance of competition between suppliers — and in general externalities among them — as an important factor determining the choice of the quality produced. She shows that competition may lower the incentive to produce high quality and thus finds that self-regulation my serve as a way of guaranteeing product quality. Steven Davis (2001) looks at a labor market application with investments in — job and worker — quality, within a search context. However, he assumes one-time entry and a single round of search, so again he does not encounter the queuing problem.

The rest of this article continues as follows. In Section 2, we lay out the details of our model, provide the tools to evaluate welfare in the free market and under the different regulatory regimes and discuss the first best allocation. Section 3 contains the analysis of the market equilibrium and addresses comparative statics (illustrated with numerical examples), measuring the impact of changes in the level of frictions and in bargaining power. We analyze the effect of minimum quality standards (licensing) in Section 4. Section 5 concludes. Appendix A discusses the interpretation of the concept of competitive equilibrium in our market. The demonstrations omitted from the text are in Appendix B.

## 2. The model

We consider a market for a single commodity of heterogeneous quality, composed of a set of service providers (sellers) and a set of customers (buyers). The market operates over (discrete) time. Agents are risk neutral and maximize their (undiscounted) expected utility. Trade is carried out by decentralized agreements between buyer–seller pairs that meet randomly and negotiate the price to trade one unit of an indivisible good.

There is a continuum of buyers and sellers. The population of potential sellers is constant and is given by the unit interval, where each seller is uniquely indexed by her type, $\theta \in [0, 1]$. Seller $\theta$ can produce a good of quality $q(\theta)$, where $q(.):[0, 1] \mapsto [0, 1]$ is assumed to be strictly increasing.[7] For simplicity, we also assume that $q(.)$ is differentiable, $q(0)=0$ and $q(1)=1$. Note that $q(.)$ can also be thought of as the inverse of the probability distribution function of quality. We denote the average quality in the market when only goods above quality $q(\theta^*)$ are produced by $\bar{q}(\theta^*) = \frac{1}{1-\theta^*}\int_{\theta^*}^{1} q(x)\mathrm{d}x$.

Each seller can produce a single unit — which cannot be stored — in each period. The cost of production is independent of the quality[8] and it is normalized to zero. We also assume, for simplicity, that the sellers' human capital is fully specific to the market, and thus their opportunity cost of being in the market is zero. Sellers who (rationally) expect not to be able to sell at a positive price are supposed not to be present in the (steady state) market. Since a higher quality seller who trades can always do better than a lower quality one, the *active* sellers form an interval: $(\theta^*, 1]$.

In every period a measure 1 of new potential buyers appear. They are heterogeneous in terms of their outside opportunities,[9] which are distributed according to the (strictly increasing and differentiable) probability distribution function $E(.):[-\infty, 1] \rightarrow [0, 1]$.[10] They must decide whether to enter the market. If they don't, they take up their outside option and exit the model. As a result, the flow of buyers entering the market in each period is $E(.)$ evaluated at the expected profit of a buyer upon entering the market (conditional on the measure of buyers entering simultaneously). In order to inform our intuition, it will be useful as we proceed to consider the homogeneous case of completely inelastic entry as well, where a measure $e<1$ of buyers enter each period (and suffer no opportunity cost). While in the market, all buyers are identical: they wish to purchase a single unit of the good, which they value at its quality, $q(\theta)$. In every period that a buyer is in the market but does not get served he suffers a cost, $c \in (0, 1]$. Following an unsuccessful period in the market, buyers can choose whether to continue in the market next period or to return to their outside option.

In every period, traders seeking a match meet at random. We denote the probabilities of finding a trading counterpart by $\pi_b$ and $\pi_s$ for buyers and sellers, respectively. These probabilities depend (only) on the state of

[7] Note that it would be without loss of generality to assume that it is non-decreasing. We assume strict monotonicity to simplify the analysis.

[8] Since one of our claims is that too much of mediocre quality is produced in equilibrium, by not giving mediocre producers a cost advantage, we actually strengthen our result.

[9] Alternatively, travel costs, either literally or in the sense of Hotelling.

[10] The presence of a positive measure of buyers with (very) negative opportunity costs is not necessary. However, otherwise for some parameter values a steady state equilibrium would fail to exist (see Section 3.1). We assume that $-\int_{-\infty}^{0} x\mathrm{d}E(x) < \infty$, that is, the aggregate "opportunity benefit" of the buyers entering in a given period is finite. Otherwise, welfare would be infinite, independent of the market outcome.

the market, denoted by ($\theta^*$, $b^*$). $\theta^*$ is the marginal seller (so that $1-\theta^*$ is the measure of active sellers in the market), while $b^*$ is the measure of buyers that seek a seller. There is efficient,[11] uniform random matching among the active market participants: sellers receive a customer with probability $\pi_s = \min\left\{1, \frac{b^*}{1-\theta^*}\right\}$ and buyers find an active seller with probability $\pi_b = \min\left\{1, \frac{1-\theta^*}{b^*}\right\}$.

Once a buyer finds himself in a store, he learns the quality of the product offered by the seller. Next, they start bargaining over the price, $p$, where the buyer maximizes $q-p$ while the seller maximizes $p$. Bargaining proceeds as follows. One of the parties is randomly selected to make a proposal. The probability that the buyer (the seller) is selected is $\lambda \in (0, 1)$ (respectively, $1-\lambda$). If the responder accepts the proposed price, the transaction is consummated. The buyer then leaves the market, while the seller can serve a new customer in the following period. If the responder rejects, then they break up negotiations and both traders search for a new match in the following period (or the buyer may decide to take up her outside opportunity).[12]

We are interested in characterizing the market in its steady state, that is, when the composition of active traders in equilibrium is constant over time. We will refer to such a situation as the *market equilibrium*.

### 2.1. Welfare

In this subsection we develop the (utilitarian) welfare function, which will enable us to evaluate the efficiency of our market. We also establish the aggregate profit function, which we will use to derive the optimal decisions of a (self-)regulatory authority.

In equilibrium,[13] the measure of active sellers in each period is $1-\theta^*$. Given the matching technology, everyone on the short side of the market is matched in each period, so — since, as we will see shortly, in equilibrium all matches end in trade — the measure of buyers *entering* the market in its steady state is the same as the measure of trade, $t^* = \min\{b^*, 1-\theta^*\}$ in each period.

The aggregate opportunity cost incurred by the buyers entering in a given period is thus

$$\int_{-\infty}^{E^{-1}(t^*)} x \mathrm{d}E(x) = \int_0^{t^*} E^{-1}(x)\mathrm{d}x.$$

Welfare per period is the realized surplus (quality weighted measure of trade) minus the opportunity and waiting costs:[14]

$$W = t^* \overline{q}(\theta^*) - \int_0^{t^*} E^{-1}(x)\mathrm{d}x - (b^* - t^*)c. \qquad (1)$$

When entry is inelastic, there is no opportunity cost to worry about:

$$W(e) = e\overline{q}(\theta^*) - (b^* - e)c.$$

A profit-maximizing, self-regulating professional college's objective function can be derived along the same lines as the welfare function. Denote by $p(\theta)$ the expected price that prevails in the $q(\theta)$ bargaining pair. Then we have that the aggregate profits are given by

$$\Pi = \frac{t^*}{1-\theta^*} \int_0^1 p(x)\mathrm{d}x = t^* \overline{p}(\theta^*).$$

### 2.2. The efficient outcome

As a benchmark, it is useful to characterize the welfare-maximizing allocation — attainable under centralized trade with complete information — in our market.

**Proposition 1.** *The unconstrained welfare maximizing allocation results in a balanced market. The marginal trader, $\theta^c$, is the solution of $E(q(\theta^c)) = 1-\theta^c$, and $1-\theta^c$ buyers enter in each period. For exogenous entry, the marginal trader in the optimal allocation is given by $\theta^c(e) = 1-e$.*

**Proof.** See Appendix B. □

Quite intuitively, the first best results in no idle traders, and the marginal consumer having an opportunity cost which equals the quality (surplus) provided by the marginal seller. See Fig. 1. Note that unlike in the usual graph, the vertical axis measures buyer's surplus,

[11] Considering the frictionless limit is for simplicity; our analysis can be easily extended to environments where the matching technology displays search frictions.

[12] Note that this bargaining procedure is an equilibrium outcome in a much more elaborate model, where following the random choice of the first proposer, there is an alternating offer bargaining game with either player being able to leave the negotiating table following a rejection. See Ponsatí and Sákovics (1998).

[13] For ease of exposition, we define the welfare function directly for a market equilibrium.

[14] Note that the expected number of periods a new entrant has to wait before getting served is $\frac{1-\pi_b}{\pi_b}$. Since $\pi_b = \frac{t^*}{b^*}$, we have that per entrant the expected waiting cost is $\frac{b^*-t^*}{t^*} - c$. Since there are $t^*$ new entrants, the aggregate cost of waiting is $(b^*-t^*)\,c$. This is the same as the cost of waiting incurred by the currently waiting buyers in any given period.

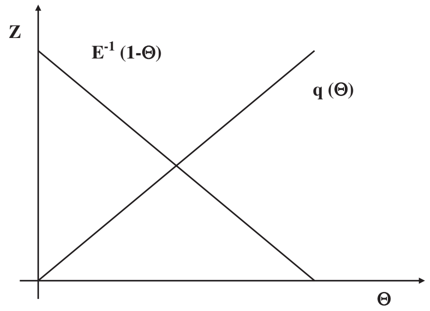


Fig. 1. Determining the buyer's surplus in the competitive equilibrium.

$z := q(\theta) - p(\theta)$, rather than price. For a further elaboration on this, and on how we can derive this "competitive" allocation from first principles, see Appendix A.

## 3. The market equilibrium

The characterization of the market equilibrium requires the simultaneous determination of the price distribution, $p(.)$, the marginal quality, $q(\theta^*)$, and the measure of buyers active in the market, $b^*$.

Consider a market equilibrium. Denote the expected surplus of a buyer upon entry by $x_b$. This variable is central to the analysis, as it impacts on the outcome of bargaining and the entry decisions of traders on both sides of the market. The first result we want to establish is that no buyer ever leaves the market empty handed and that every match will result in a trade. Note that, despite the fact that only sellers who (rationally) expect to be able to trade (eventually) participate in the matching, the latter claim is not obvious as we could have a situation where some buyers only agree to trade if they are chosen to make the first offer — and otherwise they prefer to wait for a new match.

**Lemma 1.** *In equilibrium every match results in trade and no (unmatched) buyer leaves the market. Moreover, the marginal quality is given by*

$$q(\theta^*) = \max\{x_b - c, 0\}. \quad (2)$$

**Proof.** By stationarity, if a buyer expected more in the market than her outside opportunity in the previous period, this will still be the case in the current period. Thus, no buyer ever leaves the market empty handed. By the above (and stationarity) the buyers' continuation value upon disagreement is the same as their expected surplus upon entry, but "discounted" by one period: $x_b - c$. This determines the marginal quality that is supplied in equilibrium, $q(\theta^*)$, in a straightforward way. Note that no buyer will ever purchase from a seller that offers a quality of $q(\theta) < x_b - c$, since even trade at zero price — the lowest price a seller would accept — would yield less than waiting for a rematch. Thus, potential sellers of quality below $x_b - c$ will not be active in a market equilibrium. On the other hand, a situation where the marginal quality offered in the market is $q(\theta^*) > \max\{x_b - c, 0\}$ cannot happen in a market equilibrium either. Otherwise, potential sellers of quality $q(\theta)$, $\max\{x_b - c, 0\} < q(\theta) < q(\theta^*)$, would enter the market and — if matched and chosen to make an offer, which would happen with positive probability — make positive profits by trading at a price slightly above $q(\theta) - (x_b - c) > 0$, since any matched buyer is strictly better off trading at such price than not trading. Hence, in a market equilibrium $q(\theta^*) = \max\{x_b - c, 0\}$. This, in turn, implies that all matches end in trade, since for $\theta' > \theta^*$, both traders are happy to trade for any $p \in (0, q(\theta') - q(\theta^*))$. □

The endogenous outside option of a buyer (not to be confused with the exogenous outside opportunity when she does not enter the market) is equal to the expected profits from a future match, net of the expected cost of waiting for that match. For a buyer the expected value attainable in future matches is independent of the quality of the good that he is currently bargaining for. Instead, it is a function of the distribution of quality effectively supplied in the market, and it coincides with the "discounted" value of the expected payoff of a buyer upon entry, $x_b - c$. Turning to the sellers, note that their outside option is zero, since the future sales are independent of what happens in this period, and upon disagreement with the buyer they are currently matched with, they have no further opportunity to sell in the current period.

In order to derive the (expected) outcome of bargaining, observe that in equilibrium all proposers offer the price that leaves the responder indifferent to his/her outside option. Therefore, the expected equilibrium price (as a function of quality) is given by

$$p(\theta) = (1 - \lambda)(q(\theta) - x_b + c). \quad (3)$$

Next, note that the expected surplus of a buyer upon entry satisfies

$$x_b = \pi_b \int_{\theta^*}^{1} \frac{q(\theta) - p(\theta)}{1 - \theta^*} \mathrm{d}\theta + (1 - \pi_b)(x_b - c). \quad (4)$$

To understand this equality, note that the buyer will either be matched and then obtains the expected difference between the quality of the good and the price she will pay for it, or, she will be unmatched and will have to wait until next period to find herself in the same situation as now.

Substituting the price Eq. (3) into Eq. (4) and rearranging, we obtain[15]

$$x_b = \overline{q}(\theta^*) - \frac{c(1-\lambda\pi_b)}{\lambda\pi_b}. \quad (5)$$

Therefore, putting Eqs. (5) and (2) together, we have identified our first equilibrium condition:

$$q(\theta^*) = \max\left\{\overline{q}(\theta^*) - \frac{c}{\lambda\pi_b}, 0\right\}. \quad (6)$$

Turning to the buyers' entry decision, recall that in a steady state the measure of buyers entering and leaving (trading) must be equal. By Lemma 1, all matches lead to an exit, so the flow of buyers through the market equals $t^*$. Thus, in general, $E(x_b)=t^*$. Substituting in from Eq. (2), we obtain our second equilibrium condition

$$q(\theta^*) = \max\{E^{-1}(t^*) - c, 0\}. \quad (7)$$

In order to simplify the analysis, we will assume that the distribution of quality is well-behaved: We say that the quality distribution is *regular* if $\overline{q}'(\theta)<q'(\theta)$ for all $\theta\in[0, 1]$. A sufficient[16] condition for regularity is that $q(.)$ is concave. Since $q(.)$ is the inverse of the distribution function of quality, this roughly says that we should not have too many low quality potential sellers.

Let $\theta_b(\lambda, c)$ be the unique (guaranteed by regularity) solution to our first equilibrium condition (6) for a balanced market:

$$\lambda[\overline{q}(\theta) - q(\theta)] = c, \quad (8)$$

when it exists (that is, when $\overline{q}(0)\geq c/\lambda$), and zero otherwise. This equation determines the marginal seller if the sellers' entry decision turns out to be the binding constraint. Note that in a buyers' market (where buyers are the short side and, therefore, $\pi_b=1$) the left-hand side is the expected gain to a buyer matched with the marginal seller from refusing to trade and waiting for a rematch, while the right-hand side is the cost of switching. If the gain were larger, then no buyer would be willing to trade with this seller, so she would not enter the market. Similarly, define $\theta_s(c)$ as the unique solution to[17] the second equilibrium condition (7) for a balanced market:

$$E(q(\theta) + c) = 1 - \theta. \quad (9)$$

This equation determines the marginal seller if the buyers' entry decision turns out to be the binding constraint. The left-hand side gives us the measure of buyers willing to enter as a function of the marginal quality and the waiting cost (c.f. Eq. (7)). The right-hand side is the measure of active sellers, which in a sellers' market has to equal the measure of new buyers entering the market.

We are now ready to state our characterization theorem:

**Theorem 1.** *Assume that q(.) is regular. Then there exists a unique market equilibrium.*

*i) When $\theta_b(\lambda, c)\leq\theta_s(c)$, in equilibrium buyers are the short side, the marginal seller is $\theta_b(\lambda, c)$ and the measure of buyers in the market is $b_b = E\left(\overline{q}(\theta_b) - \frac{c(1-\lambda)}{\lambda}\right)$.*

*ii) When $\theta_b(\lambda, c)\geq\theta_s(c)$, in equilibrium sellers are the short side, the marginal seller is $\theta_s(c)$ and the measure of buyers in the market is $b_s = [\overline{q}(\theta_s - q(\theta_s))]\frac{\lambda(1-\theta_s)}{c}$.*

**Proof.** See Appendix B. □

Similar arguments lead to the characterization of the market equilibrium when entry is exogenous:

**Corollary 1.** *Let q(.) be regular and assume that buyers enter at a constant rate e. Then there is a unique market equilibrium.*

1. *For small entry rates, $e\in(0, 1-\theta_b]$, buyers are the short side ($b=e$) and the marginal seller is $\theta_b(\lambda, c)$.*
2. *For larger entry rates, $e\in(1-\theta_b, 1)$, sellers are the short side ($\theta^*=1-e$) and the measure of buyers in the market is $b^* = (\overline{q}(1-e) - q(1-e))\frac{\lambda e}{c}$.*

### 3.1. The price distribution

In order to complete our description of the market equilibrium we need to establish how the surplus

[15] A more intuitive way of writing Eq. (5) is $x_b = \overline{q}(\theta^*) - \frac{c(1-\lambda)}{\lambda\pi_b} - \frac{c(1-\pi_b)}{\pi_b}$, where the second term on the right-hand side is the average price and the last term is the expected waiting cost of a new entrant.

[16] But, by far, not necessary: the uniform quality distribution satisfies regularity by a wide margin, as $\overline{q}(\theta) - q(\theta) = \frac{1-\theta}{2}$.

[17] To be rigorous, if we do not want to use the additional information that $q(\theta)+c\leq 1$ before solving the equation, we need to extend the support of $E(.)$ up to 2: $E(x)\equiv 1$ for $x\in[1, 2]$.

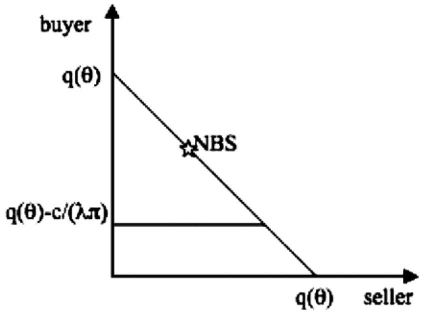


Fig. 2. The Nash Bargaining Solution.

generated by trade is shared. Putting Eqs. (3) and (5) together, we obtain the price distribution:

$$p(\theta) = (1-\lambda)\left(q(\theta) - \overline{q}(\theta^{*}) + \frac{c}{\lambda\pi_{\mathrm{b}}}\right). \tag{10}$$

That is, expected prices are linearly increasing in quality, with zero — by Eq. (6) — being paid for the good of the marginal producer (unless the marginal quality is zero, in which case the lowest price is $p(0)=(1-\lambda)\left(\frac{c}{\lambda}-\overline{q}(0)\right)>0$).[18] Consequently, the sellers do obtain the share corresponding to their raw bargaining power, but not of the full surplus, rather of the incremental surplus over the one provided by the marginal seller. This is consistent with the fact that our bargaining equilibrium is equivalent to the (asymmetric) Nash Bargaining Solution, with the vector of outside options for the buyer and seller, $\left(\overline{q}(\theta^{*}) - \frac{c}{\lambda\pi_{\mathrm{b}}}, 0\right)$, as the threat point and $q(\theta)$ the size of the bargaining surplus.[19] See Fig. 2.

Note that the price at all supramarginal sellers is decreasing in the *marginal* quality, $q(\theta^{*})$. In other words, more competition — additional (though lower quality) sellers active in the market — corresponds to higher prices earned by the incumbents. This is a direct consequence of the fact that the sellers' entry is endogenous, implying that the marginal seller must expect zero in equilibrium, whoever she is. Of course, this does not imply that high quality sellers should prefer to maximize entry, since to achieve that they may have to decrease their matching probability. We will examine this question in more detail in Section 4, where we analyze the effects of the introduction of a quality threshold — as a result of, say, occupational licensing — on the market equilibrium.

Condition (10) also allows us to make our first empirically relevant/testable observation.

**Proposition 2.** *If the price charged by the marginal seller is zero (positive), the "price-quality ratio",* $\frac{p(\theta)}{q(\theta)}$, *is strictly increasing (decreasing) in the quality of the good traded.*

**Proof.** From Eq. (10) it is immediate that

$$\frac{p(\theta)}{q(\theta)} = (1-\lambda)\left[1 - \frac{\overline{q}\left(\theta^{*}\right) - \frac{c}{\lambda\pi_{\mathrm{b}}}}{q(\theta)}\right], \tag{11}$$

which — for $\lambda\in(0, 1)$ — is strictly increasing (decreasing) in $q(\theta)$, as long as $\overline{q}(\theta^{*}) - \frac{c}{\lambda\pi_{\mathrm{b}}} > (<) 0$. From (6) and (10), this corresponds to whether, the price charged by the marginal seller is zero or positive. □

That is, the proportion of the gains from trade appropriated by the sellers is increasing with the quality of their good as long as the buyers can always expect a positive surplus. Since the latter condition is usually satisfied, this implies that the mark-up is increasing with quality, as is commonly observed. On the other hand, when the buyers expect a loss in the market (even if a smaller loss than what they would suffer outside), the result is the opposite, the mark-up is decreasing with quality, as it is driven by a fixed (independent of quality) effect. If customers are in dire straits, even low quality sellers can extract significant surplus, so prices are less

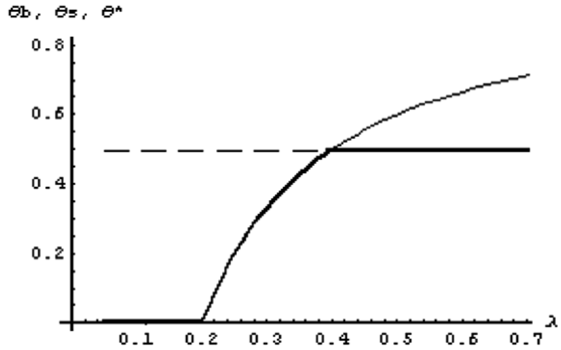


Fig. 3. Marginal seller $\theta^{*}$ as a function of $\lambda$; $\theta_{\mathrm{s}}$ is the flat line and $\theta_{\mathrm{b}}$ is the increasing curve.

[18] For $c/\lambda$ high enough to result in a negative buyer's option $\left(\overline{q}(\theta^{*}) - \frac{c}{\lambda\pi_{\mathrm{b}}} < 0\right)$, the price charged may exceed the quality provided. This is a direct consequence of the fact that the benefit to a buyer from trade is not just $q(\theta)$ but also the avoidance of the waiting cost. Of course, this can only happen in equilibrium because, by assumption, there are buyers with negative enough opportunity costs who are still willing to enter the market. If this were not the case, the market would fail for $c/\lambda$ high enough.

[19] See Binmore et al. (1986) for a discussion of this equivalence.

dispersed (c.f. the small(er) price dispersion observed, say, for emergency plumbing services).

At this point an example is useful to illustrate the mechanics of the market equilibrium and to preview the main points that we will be making in the sequel.

### 3.2. An example

Consider a market where potential qualities are distributed uniformly in [0, 1], i.e. $q(\theta)=\theta$, and assume that entry is given by $E(x)=\frac{1}{(2-x)^2}$. The efficient quality provision is given by the solution to $\frac{1}{(2-x)^2}=1-x$, that yields $\theta^c=.534$. The first best level of welfare is thus $W^{\text{fb}}=(1-\theta^c)\overline{q}(\theta^c)-\int_0^{1-\theta^c}E^{-1}(x)\mathrm{d}x=\frac{(1-.534)(1+.534)}{2}-\int_0^{.466}\left(2-\frac{1}{\sqrt{x}}\right)\mathrm{d}x=.357+.434=.791.$

Let us fix the delay cost at, say, $c=.1$, and examine the market equilibrium for different values of the bargaining parameter $\lambda$. Eq. (8) yields $\theta_\text{b}(c,\lambda)=1-\frac{.2}{\lambda}$, which is positive for $\lambda\geq.2$. Eq. (9) yields $\theta_\text{s}(c)=.494$, and $1-\frac{.2}{\lambda}\leq.494$ if and only if $\lambda\leq.395$. Fig. 3 displays $\theta_\text{s}$ and $\theta_\text{b}$ as a function of $\lambda$, and the determination of the measure of active sellers, $\theta^*$. Thus, for $c=.1$, the market equilibrium is as follows:

1. For $\lambda\leq.2$ all sellers are active, $\theta^*=0$ and the buyers are the short side; $b^*=\frac{100\lambda^2}{(14\lambda+1)^2}$, and $\pi_\text{s}=\frac{10(1-\lambda)}{15-14\lambda}<1$.
2. For $.2<\lambda\leq.395$ not all potential sellers are active, $\theta^*=1-\frac{.2}{\lambda}>0$, but buyers continue to be the short side; $b^*=\frac{10(1-\lambda)\lambda}{1+9(1-\lambda)\lambda}$, and $\pi_\text{s}=\frac{50(1-\lambda)\lambda^2}{1+9(1-\lambda)\lambda}<1$.
3. For $\lambda>.395$ the marginal seller is $\theta^*=.494$, and sellers become the short side; $b^*=1.28\lambda$ and $\pi_\text{b}=\frac{.395}{\lambda}<1$.

The market equilibrium measures of active agents for each side of the market are displayed in Fig. 4. With these measures in hand, welfare, profits and consumer surplus as a function of $\lambda$ are readily evaluated. The results of this exercise are displayed in Fig. 5, where values can also be compared to the first best welfare (top horizontal line). It is apparent that the efficient allocation of bargaining power is one where both sides make the offer with positive probability. Furthermore, the maximum welfare (attained at the value of $\lambda$ which balances the market) still falls short of the first best, as the marginal quality is too low. Let us now fix the bargaining power at, say, $\lambda=1/2$, and examine the market equilibrium for varying levels of the delay cost $c$. Eq. (8) yields $\theta_\text{b}(c,\lambda)=1-4c$, which is positive for $c\leq1/4$. Eq. (9) is now the cubic equation $\frac{1}{(2-x-c)^2}=1-x$. Its unique real solution, $x(c)$, is displayed in Fig. 6, along with $1-4c$; observe that $x(0)=\theta^\text{c}$ and $x(c)$ crosses $1-4c$ from below at $c=.129$. Thus, the market equilibrium for $\lambda=1/2$ is as follows:

1. For $c\leq.129$ sellers are the short side and the marginal seller is $\theta^*=x(c)$; $b^*=\frac{(1-x(c))^2}{2c}$ and $\pi_\text{b}=\frac{2c}{1-x(c)}<1$.
2. For $.129<c\leq.25$ buyers are the short side and the marginal seller is $\theta^*=1-4c$; $b^*=\frac{1}{(1+3c)^2}$ and $\pi_\text{s}=\frac{1}{(1+3c)^2 4c}<1\cdot$

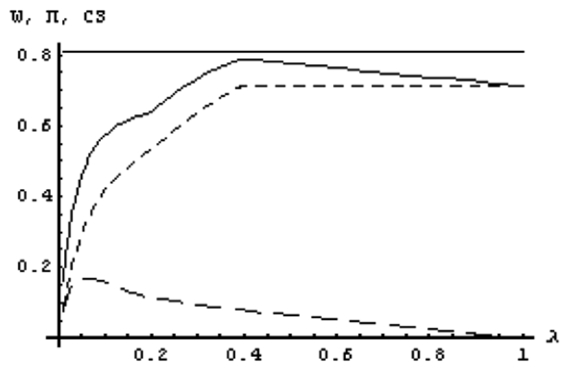


Fig. 5. Welfare, consumer surplus and profits (top, middle and bottom curves); the flat line displays first best welfare.

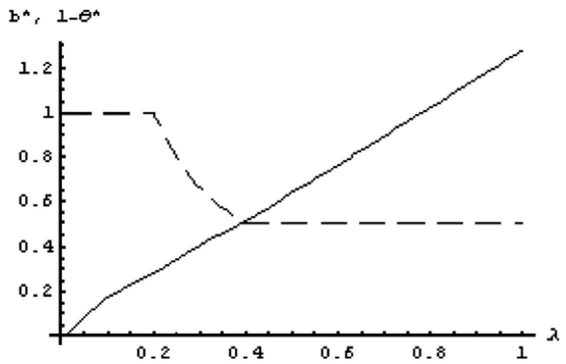


Fig. 4. Buyers ($b^*$ continuous line) and sellers ($1-\theta^*$ dashed line) as a function of $\lambda$.

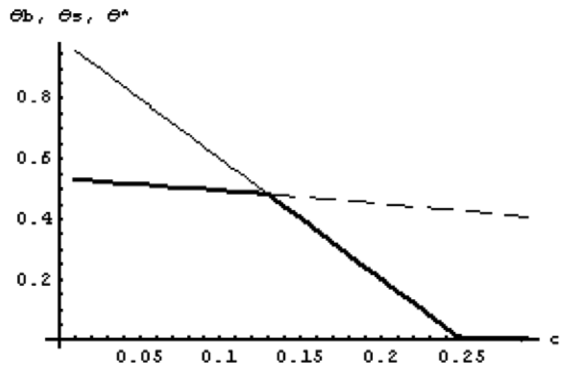


Fig. 6. Marginal seller $\theta^*$ as a function of $c$.

3. For $c > .25$ buyers are the short side with all potential sellers active; $b^* = \frac{1}{(1.5+c)^2}$ and $\pi_s = b^* < 1$.

The measures of agents active on each side of the market as a function of $c$ are displayed in Fig. 7. Note that, as $c \to 0$, $\theta^* \to \theta^c$, $b^* \to \infty$ and $\pi_b \to 0$. Hence, as delay costs vanish, the provision of quality approaches the efficient level. However, the measure of buyers that keeps the market at a steady state grows without bound. In fact, this increase in the queue is faster than the decrease in the search cost. Fig. 8, displays welfare, consumer surplus and profits as a function of $c$, and relates them to the first best welfare; it is clear that efficiency is not attained in the limit. Indeed, at $c=0$, welfare, profits and consumer surplus are all increasing in $c$! Again, the maximum welfare — attained at the value of $c$ which balances the market — still falls short of the first best.

Before turning to the analysis of regulatory tools that might constrain and perhaps correct the market performance, in the next subsections we show that the equilibrium features and comparative statics displayed in this example hold in general, and we shed light on the reasons behind them.

### 3.3. Delay costs

Let us first investigate the behavior of the market equilibrium in general as the per period cost of delay, $c$, varies.

Let us first look at the imperfection caused indirectly by this friction: the deviation of the quality provided by the market from the efficient provision (given by the competitive allocation derived in Section 2.2). As expected, since frictions "lock buyers in", the market equilibrium always has too low marginal — and therefore average — quality. However, as (per period) frictions disappear, we obtain the competitive allocation:

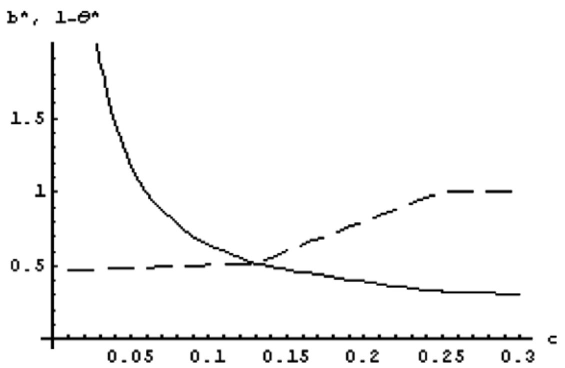


Fig. 7. $1-\theta^*$, and $b^*$, as a function of $c$.

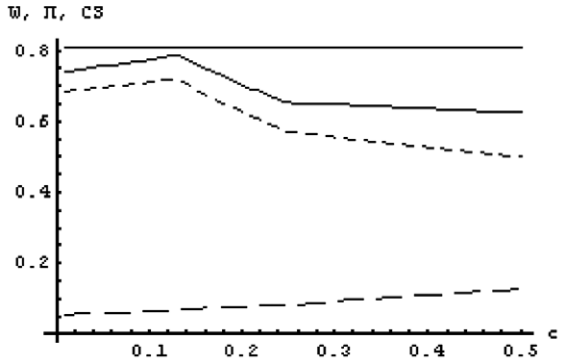


Fig. 8. Welfare, consumer surplus and profits (top, middle and bottom curves) as a function of $c$; the flat line displays the first best welfare.

**Proposition 3.** *For any $\lambda \in (0, 1)$, the marginal quality in the market equilibrium is decreasing in $c$,*[20] *and in the limit as (per period) frictions disappear, $c \to 0$, it coincides with the competitive one.*

**Proof.** Implicitly differentiating their defining equalities, it is easy to see that both $\theta_b(\lambda, c)$ and $\theta_s(c)$ are decreasing in $c$. Since $\lim_{c \to 0} \theta_b(\lambda, c) = 1$ for all $\lambda > 0$ (even if $q(.)$ is not regular!),[21] as $c \to 0$ we always have a sellers' market, and the marginal seller is $\theta_s(0)$, which coincides with $\theta^c$. □

It is important to observe that the convergence to the Walrasian outcome is not complete, though. Notably, the price distribution in the "frictionless" market does not match that of the competitive equilibrium. As we show in Appendix A, the latter leads to a constant buyer's surplus: $q(\theta) - p^c(\theta) \equiv q(\theta^c)$, while the former is given by (7): $p(\theta) = (1-\lambda)\ (q(\theta) - q(\theta^c))$. Thus, the competitive price distribution corresponds to the frictionless market price distribution where almost all the bargaining power is given to the producers ($\lambda \simeq 0$).

For $\lambda > 0$, the buyer's surplus is increasing in quality in the "frictionless" market equilibrium: $q(\theta) - p(\theta) = \lambda q(\theta) + (1-\lambda)\ q(\theta^c)$. This is at first blush surprising: if there are indeed no frictions, why a consumer matched with the marginal seller does not refuse to trade and wait for a better match? The explanation can be extracted from Theorem 1: the measure of buyers in the market, $b^*$, is increasing without bound as the buyers become more and

[20] Until it reaches zero. From then on it is, obviously, constant.

[21] Note that this result does not require $q(.)$ to be regular, since as $c \to 0$ all the roots of $\bar{q}(\theta) - q(\theta) = c$ converge to 1 anyway.

more patient. As a result, the probability of getting rematched in the next period tends to zero as $c$ tends to zero.

Consequently, we do not achieve an efficient outcome even in the limit as (per period) delay cost disappear: while the quality distribution and the measure of new buyers entering the market are the same as in the competitive equilibrium, there exists an infinite queue and the buyers incur positive delay costs on average.

Actually, the limiting outcome is not only inefficient, but it is not even second best! To see this, start by observing that once $c$ is low enough to put us in a sellers' market, the expected payoff of a buyer upon entry, $x_b$, is increasing(!) in $c$.

**Corollary 2.** *In a sellers' market, the expected payoff of a buyer upon entry, $x_b$, is increasing, while the expected payoff upon opting out from a match is decreasing, in the delay cost:* $\frac{dx_b}{dc}>0, \frac{d(x_b-c)}{dc}<0$.

**Proof.** See Appendix B. □

That is, as $c$ decreases, the effect on buyers who are in a match and on those who are not are divergent: while a buyer considering leaving the seller he is matched with will find it easier to switch (as expected), a buyer just entering the matching process will actually face a more costly wait! This latter phenomenon opens the possibility of *welfare* being increasing in the delay cost. That is, contrary to intuition, the more patient traders are in a decentralized market the lower welfare they might enjoy! This is exactly what is displayed in Fig. 8.

The explanation of the "paradox" has a taste of the tragedy of the commons: when buyers decide to enter the queue they do not take into account the negative externality imposed on the rest. As the cost of waiting decreases,[22] during a transitory phase even more buyers enter the market. However, as we have seen above, this is accompanied by a *decrease* in the number of sellers, since now it is easier to switch away from a low quality seller. As a result the queue grows larger, eventually dissuading entry sufficiently so that it stabilizes at the level of the sellers present in the market. This means that we end up with fewer buyers entering per period than before, implying that the expected profit upon entry has decreased. Since prices have also decreased, the sellers are worse off as well. Thus — since by Proposition 3 and the fact that $b$ (and thus the amount of trade) is decreasing in $c$, welfare is always decreasing in $c$ in a buyers' market as well — we have proved that

[22] In the classic commons example this would correspond to a lower (private) cost of maintaining a goat.

**Proposition 4.** *The delay cost that maximizes welfare is not 0. Rather, it is $c_w=E^{-1}(1-\theta_b)-q(\theta_b)$, the delay cost that equates the number of buyers and sellers in the market.*

At first glance, one might think that the above result is driven by the endogeneity of the buyers' entry process. However, this is not the case, as we show next.

Consider a balanced market. When entry is exogenous, a decrease in the delay cost will, obviously, have no effect on entry. Instead, initially it will decrease the number of sellers, since now it is cheaper to switch away from them. Since entry stays constant, this will result in more unmatched buyers. Eventually, the queue will be sufficiently long, so that the expected payoff upon switching is the same as before and the marginal seller is the $\theta=1-e$, once more. What has happened to welfare? Well, if $x_b-c$ stayed constant while $c$ decreased, it must be the case that $x_b$ has decreased. Since production and prices are unchanged, this implies that welfare has also decreased. Consequently, Proposition 4 holds for exogenous entry as well:

**Corollary 3.** *For exogenous entry the welfare optimal delay cost is $c=\lambda\,(\bar{q}(1-e)-q(1-e))>0$, which balances the market.*

Pulling all these results together, we can see that if the free market is a buyers' market then a regulator would like to decrease waiting costs, since that would have two beneficial effects: it would increase the marginal quality and decrease the measure of idle sellers. On the other hand, if the free market is one where the consumers are queuing, then the trade-off between improving the quality distribution and reducing waiting time is clearly decided in favor of the latter. The importance of the cost of waiting in the evaluation of social welfare is a recurrent and key observation in our analysis.

### 3.4. Bargaining power

When analyzing the effects of the distribution of bargaining power, the key observation to make is that for any interior values of the parameters, changes in $\lambda$ can move the equilibrium between a sellers' and a buyers' market. In a buyers' market the marginal seller is increasing in $\lambda$ — since the higher the buyers' bargaining power the better is their outside option —, while no queues are formed. Consequently, since the original level of marginal quality was too low, an increase in $\lambda$ will always lead to a welfare improvement. However, at some level of $\lambda$ the market will turn into a

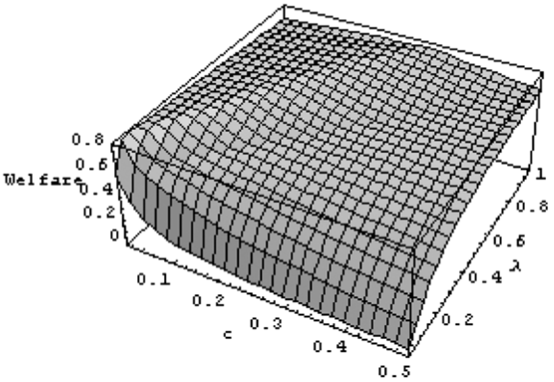


Fig. 9. Welfare as a function of $c$ and $\lambda$.

sellers' market. In a sellers' market we have the opposite situation. The marginal seller is independent of $\lambda$, while the queue length is increasing in it (since the better deal a buyer can get the more waiting he is willing to put up with). Consequently, welfare is decreasing in $\lambda$. As a result, we have the following:

**Proposition 5.** *The welfare maximizing (second best) bargaining power, $\lambda_w$, is the one that leads to an equal number of buyers and sellers entering the market, if feasible:* $\lambda_w = \min\left\{1, \frac{c}{\overline{q}(\theta_s) - q(\theta_s)}\right\}$ $\left(\text{or } \min\left\{1, \frac{c}{\overline{q}(1-e) - q(1-e)}\right\}\right.$, *for exogeneous entry*$\left.\right)$.

Propositions 4 and 5 are two sides of the same coin. Together they demonstrate that the welfare function always has a type of contract curve (that is, a locus of maxima) in $c-\lambda$ space that forms a "ridge:" at any point along it welfare is maximized simultaneously[23] in both $c$ and $\lambda$. This feature would indicate that the two "tools" are interchangeable. However, the "ridge" also has the particularity that it can only "spend time" at one edge, the $\lambda=1$ one. In other words, for any $\lambda$, we can find a low enough $c$ that puts us in a sellers' market, but not *vice versa*. This is illustrated in Fig. 9 that displays the 3-dimensional picture of welfare as a function of $c$ and $\lambda$ for the parameters of the example of Section 3.2.

Having fully characterized the free market, we turn to the possibility of (self-)regulation next.

## 4. Licensing

In this section we analyze the distributive and welfare effects of a minimal level of quality, $q^*$, directly imposed by a (self-)regulator. Note that this quality threshold acts as a constraint on the market and therefore it necessitates a re-examination of the equilibrium behavior of the traders rather than "just" being a comparative statics exercise.

In Section 3 we have seen that not even the "frictionless" free market leads to the first best welfare. Since it does lead to the competitive level of quality supplied, this implies that, as a second best, the most efficient outcome in a "frictionless" market should involve a somewhat different marginal quality from the competitive one. Whether higher or lower, depends on how the expected cost of waiting varies with the marginal consumer. This is not obvious, since there are two effects at work — in opposite directions. On the one hand, increasing the marginal quality lowers the number of available sellers, so it increases the wait for a match. On the other hand, such an increase also decreases the surplus, which makes it possible to maintain the equilibrium entry rate with a shorter queue, which decreases the wait for a match. How this trade-off is decided is the first question we investigate.

Note that the expression for the buyers' expected payoff upon entry, Eq. (5), is not affected by the quality threshold. On the other hand the marginal quality is now given by

$$q(\theta^*) = \max\{q^*, x_b - c, 0\}.$$

Assume that $q^* > \max\{x_b - c, 0\}$, so that the quality threshold is binding (otherwise, it has no effect on the market equilibrium), and let $\theta^* = q^{-1}(q^*)$. With a binding quality threshold the identity of the sellers active in the market is directly driven by the constraint. Consequently, the measure of buyers in the market is sufficient to characterize the equilibrium allocation:

**Proposition 6.** *When the quality threshold is binding, the measure of buyers in the market is given by*

$$b^* = \begin{cases} E\left(\overline{q}(\theta^*) - \frac{c(1-\lambda)}{\lambda}\right), & \text{if } E\left(\overline{q}(\theta^*) - \frac{c(1-\lambda)}{\lambda}\right) \le 1 - \theta^* \\ \lambda(1-\theta^*)\left(\frac{\overline{q}(\theta^*) - E^{-1}(1-\theta^*)}{c} 1\right), & \text{otherwise.} \end{cases}$$

**Proof.** See Appendix B. □

Turning to the price distribution, notice that it is still determined by Eq. (10). Using the queue length given by Proposition 6 we obtain:

**Corollary 4.** *In a sellers' (buyers') market, the prices at supramarginal sellers are increasing (decreasing) in a binding quality threshold.*

In a buyers' market, a rise in the marginal — and therefore the average — quality improves the buyers' outside option, resulting in lower prices. In a sellers' market there is a countervailing effect, which turns out

[23] Note that this does not imply that the level of welfare achieved along this curve is constant.

to be the stronger one: the waiting time increases as a result of the restriction in supply. Consequently, prices increase.

When entry is exogenous we cannot have a binding quality constraint in a sellers' market, since there would be too much entry for the market to stabilize. In a buyers' market the trades are completely determined by the two constraints:

**Corollary 5.** *When the quality threshold is binding and entry is exogenous, the measure of buyers in the market is* $b^*=e$*, while the measure of sellers is* $1-\theta^*\geq e$*.*

Turning to the socially optimal quality constraint, we have that:

**Proposition 7.** *The welfare maximizing (binding) quality threshold is the one that equates the measure of buyers and sellers, determined by:*

$$E^{-1}(1-\theta^*) = \overline{q}(\theta^*) - \frac{c(1-\lambda)}{\lambda},$$

*for endogenous entry and* $\theta^*=\theta^c(e)=1-e$*, for exogenous entry.*

**Proof.** See Appendix B. □

In other words, when the unregulated market is a sellers' market an imposed minimum quality cannot improve welfare.[24] This has the important implication that licensing cannot alleviate the inefficiency caused by congestion (in a sellers' market). On the other hand, if the unregulated market is a buyers' market then there is room for improvement.

When entry is endogenous, the optimal threshold will be strictly below $\theta^c$. The distortion from the competitive level of marginal quality is increasing in the friction ($c$) and decreasing in the buyers' bargaining power ($\lambda$). When entry is exogenous, we can obtain (full) efficiency by imposing the optimal quality constraint.

In either scenario, if the regulator were able to set both the bargaining power and a quality threshold, the first best outcome could be guaranteed:[25]

**Corollary 6.** *For any* $c\in(0, 1]$*, there exists a* $\lambda\in(0, 1)$*, such that setting the quality threshold at the efficient level the market equilibrium leads to the competitive outcome.*

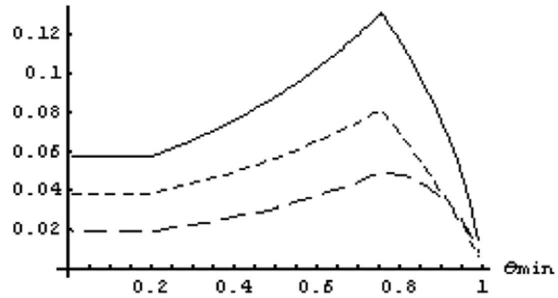


Fig. 10. The profit-maximizing minimal quality can lead to an improvement for both sides.

**Proof.** See Appendix B. □

When the (aggregate) profit-maximizing seller's cartel can set the quality standard, we have that the quality threshold is determined independently of the bargaining powers:

**Proposition 8.** *When entry is endogenous, the profit-maximizing quality threshold always results in a sellers' market. If interior, this marginal quality satisfies*

$$q(\theta^*) + c = E^{-1}(1-\theta^*) - (1-\theta^*)\frac{dE^{-1}(1-\theta^*)}{d\theta^*}.$$

**Proof.** See Appendix B. □

Note that the second term on the right-hand side is positive, and that if it were zero, the solution of the first-order condition would coincide with the unregulated outcome. That is, when the solution is interior, we have that self-regulation will be welfare improving as long as — the unregulated outcome is a buyers' market and — entry is not too inelastic.[26] In fact, for some environments the minimal quality set for the benefit of active sellers can be beneficial for the buyers as well. Fig. 10 displays welfare, consumer surplus and profits (top, middle and bottom curves) as a function of the minimum quality standard for an environment[27] where licensing by the college increases Consumer Surplus.

As we have seen above, exogenous entry is not compatible with a binding quality constraint in a sellers' market. Since in a buyers' market prices are decreasing, while the amount of trade stays constant, we have that

[24] Interestingly, a constraint imposing a maximum on the marginal quality would help, but it would be rather difficult to implement (possibly with a lower bound imposed on number of licensees?).

[25] Because of our restriction that the per period cost cannot exceed the maximum quality, setting the delay cost (and the marginal quality) may not be sufficient to achieve the first best.

[26] In this latter case, the marginal quality may be set so high that welfare is below the one that would result in the unregulated market.

[27] The corresponding parameters are: $c=.2$, $\lambda=.5$, $E(x)=1/(2-x)^5$, $q(\theta)=\theta$.

**Corollary 7.** *When entry is exogenous, a sellers' cartel prefers not to impose a quality threshold.*

**Proof.** See Appendix B. □

## 5. Conclusions

We have shown that in a market with vertical differentiation, if sellers can sell only one unit per period and buyers do not observe quality prior meeting a seller, there is a need for public intervention even in the absence of ex post asymmetric information about the quality at the point of service. The equilibrium outcome involves excess supply (too low marginal quality) and possibly excess demand (queuing customers) relative to the first best. As excess demand is inversely proportional to the unit cost of delay, the inefficiency is robust to the elimination of frictions (delay costs): while excess supply disappears, we have permanent excess demand in the limit.

While it is convenient to think of this inefficiency as the result of an externality, it is important to note that when the cost of delay tends to zero the externality caused by a new buyer entering the queue is also approaching zero. Consequently, the existence of externalities for positive costs does not directly imply that the inefficiency continues to be present at the limit.

From a welfare maximizing point of view, the avoidance of delay costs turns out to be more important than the optimal quality mix being provided. Thus, the optimal quality threshold will never increase the queues even if that would result in a better quality distribution. Since the queue increases in the quality threshold, this also means that if there is a queue in the unregulated market, minimum quality standards will not be able to improve welfare. Under self-regulation by an association of the sellers, since the active sellers are only indirectly affected by the queuing costs, entry is restricted excessively compared to the social optimum (thereby, in general, creating queues), however not necessarily relative to the unregulated outcome. The more elastic is the consumers' entry condition (distribution of opportunity costs) the better is the seller's association optimal decision for welfare.[28]

For brevity's sake, we have not analyzed a number of additional regulatory tools. For example, we have not analyzed the effects of price fixing. It is rather straightforward that if it were feasible to set a price schedule then $p(\theta)=q(\theta)-q(\theta^c)$ would guarantee the first best. However, in practice it is not realistic to assume that such a price schedule can be implemented. A constant price imposed on the entire market should still lead to an improvement, as forcing the price to be the same at all sellers gives a stronger incentive to switch away from a low quality seller, thereby increasing the marginal quality.

Other interesting generalizations include the case where the opportunity cost of the sellers is positive — capturing the effect of a costly investment (training) — or the effects of a license fee either imposed on everyone or on newcomers only. Our model is clearly well-suited as a workhorse for such further analyses.

Finally, an interesting question is to look at how much of the intuition arrived at in this paper can be carried over to a "standard" goods market, where the buyers have heterogeneous valuations and the sellers have heterogeneous costs for one indivisible unit of a good. As shown by Anwar and Sákovics (2007), the limiting inefficiency result persists in the standard model if and only if the good is perishable; i.e. there is a non-negative probability that it perishes overnight.[29]

## Acknowledgements

We are grateful to Santi Sánchez-Pagés, Jonathan Thomas and seminar audiences at the Universities of Bonn, Cambridge, Dundee, Edinburgh, Exeter, Leicester and Stirling, as well as at the Stockholm School of Economics and University College London, for their useful comments. Inés Moreno has provided invaluable help with the illustrations. Financial support from Ministerio de Educación y Ciencia (SEJ2006-02079 and CONSOLIDER-INGENIO 2010-CSD2006-00016) and CREA-Barcelona Economics is gratefully acknowledged.

## Appendix A

In order to provide further insight, we would like to arrive at the competitive allocation from first principles as well. In the current context, this is non-trivial for two reasons.

First, since our set-up is dynamic — where in each period new participants enter the market –, in order to construct supply and demand we cannot just take a snapshot of the stock of traders present in the market. Rather, we have to look at all the traders who could potentially enter the market at any point in time.[30] By

[28] Note that, if the bargaining took place under asymmetric information, the association would have an additional reason to increase licensing standards, so the social value of self-regulation would be further increased.

[29] In the current model, a service can be thought of as certainly perishing: it disappears overnight with probability one.

[30] Strictly speaking, what we need to look at is the (discounted) "sum" of potential entrants.

stationarity, and following Gale (1987), the maximizer of the discounted sum of surplus is the same as that of the (per-period) flow of potential traders. In our model the demand flow is given by the new buyers who could potentially enter in each period, while the supply flow is given by all the sellers who could potentially produce in each period. That is, we can identify the competitive allocation by looking at the demand and supply that is newly generated in each period.

The second difficulty stems from the fact that we are not dealing with a homogeneous good, rather with a (vertically) differentiated one. As a result, the Law of One Price turns into the Law of One Surplus: Since the products are differentiated, consumers do not simply compare prices when deciding which seller to buy from. Rather, they evaluate different products according to the surplus they can obtain from their consumption.[31] Thus, in a competitive market it is the buyer's surplus available, which determines both whether a buyer enters the market, and if yes, from which seller does he purchase. Then in a competitive *equilibrium*, we must have — for the usual reasons — the buyer's surplus, $z(\theta)=q(\theta)-p(\theta)$, constant across sellers. Thus, the quantity demanded (per period) — as a function of the "market buyer's surplus" — is $E(z)$. The supply function — again as a function of $z$ — can be constructed from the constraint that no seller would be willing to sell below cost, that is, we must have $q(\theta)\geq z$ for each active seller. This implies that if they have to offer $z$ to the consumers, only sellers above $\theta=q^{-1}(z)$ will be willing to sell, leading to a supply of $1-q^{-1}(z)$. Inverting supply and demand the Marshallian way, we obtain that the quantity traded (per period) in the competitive equilibrium, $Q$, solves $q(1-Q)=E^{-1}(Q)$, which yields the marginal trader, $\theta^c$, as the solution to $q(\theta^c)=E^{-1}(1-\theta^c)$, just as when we directly maximize welfare (in Proposition 1). Finally, note that the Walrasian buyer's surplus equals the marginal quality, $z=q(\theta^c)$. See Fig. 1.

## Appendix B

**Proof of Proposition 1.** Assume $t^*=b^*$. Then (1) is clearly increasing in $\theta^*$, so at the optimum $b^*=1-\theta^*$. Similarly, if $t^*=1-\theta^*$ (1) decreasing in $b^*$ so at the optimum $b^*=1-\theta^*$. Consequently, at the optimum we must have $t^*=b^*=1-\theta^*$. Substituting in and differentiating (1) with respect to $\theta^*$ we obtain the first-order condition

$$-\overline{q}(\theta^*)+(1-\theta^*)\overline{q}'(\theta^*)+E^{-1}(1-\theta^*)=0.$$

Noting that $\overline{q}'(\theta^*)=\dfrac{\overline{q}(\theta^*)-q(\theta^*)}{1-\theta^*}$ yields

$$E^{-1}(1-\theta^*)=q(\theta^*)$$

as claimed (the second-order condition is trivially satisfied).

When entry is exogenous, we have the feasibility constraints that $e\leq b^*$ and $e\leq 1-\theta^*$. Since $W(e)$ is increasing in $\theta^*$ and decreasing in $b^*$ the result follows. □

**Proof of Theorem 1.** Let us look at the two possible market scenarios — buyers' and sellers' markets — separately.

Sellers are the short side: $b\geq 1-\theta^*=t^*$.

In this scenario Eq. (7) implies that[32] $E(q(\theta_s)+c)=1-\theta_s$, which — as we have seen above — has $\theta_s(c)$ as its unique solution. Using this equation and substituting Eq. (5) into $E(x_b)=t^*=1-\theta^*$ we have

$$q(\theta_s)+c=E^{-1}(1-\theta_s)=\overline{q}(\theta_s)-\frac{c(1-\lambda\pi_b)}{\lambda\pi_b}.$$

In a sellers' market, buyers are not always matched: $\pi_b=\frac{1-\theta_s}{b_s}<1$. Substituting in, we obtain

$$q(\theta_s)+c=\overline{q}(\theta_s)-\frac{c(b_s-\lambda(1-\theta_s))}{\lambda(1-\theta_s)}.$$

[31] Given the indivisibility inherent in a service being provided, we cannot remedy this problem by simply converting to a unit price of quality.

[32] Note that if $q(\theta^*)$ were 0, then Eq. (7) would imply that $E^{-1}(1)-c\leq 0$. Since $E^{-1}(1)=1$ this is equivalent to $c\leq 1$. Therefore, $1\geq c$ implies that we need not take the positive parts in Eq. (7).

Solving for $b_s$, we obtain

$$b_s = (\overline{q}(\theta_s) - q(\theta_s))\frac{\lambda(1-\theta_s)}{c}. \tag{12}$$

Recall that the sellers' market solution obtains if and only if $b_s \geq 1-\theta_s$. It is easy to see that $\lambda[\overline{q}(\theta_s)-q(\theta_s)]=c$ is the same as Eq. (12) with $b_s$ replaced by $1-\theta_s$. Therefore, since by the regularity of $q(.)$, $\overline{q}(\theta_s)-q(\theta_s)$ is decreasing in $\theta_s$, $b_s \geq 1-\theta_s$ if and only $\theta_s \leq \theta_b$.

Buyers are the short side: $b=t^* \leq 1-\theta^*$.

Now all the buyers get matched in every period, so $\pi_b=1$. Substituting back into Eq. (6), we obtain $q(\theta_b) = \max\{\overline{q}(\theta_b) - \frac{c}{\lambda}, 0\}$, which — as we have seen above — has $\theta_b(\lambda, c)$ as its unique solution. In the buyers' market, substituting Eq. (5) into $E(x_b)=t^*=b_b$ we obtain

$$b_b = E\left(\overline{q}(\theta_b)\frac{c(1-\lambda)}{\lambda}\right). \tag{13}$$

This equilibrium prevails if and only if $b_b \leq 1-\theta_b$. When $\theta_b=0$, the inequality is satisfied, since $E(.)$ is a probability distribution function. Otherwise, by Eq. (6), $\overline{q}(\theta_b) - \frac{c(1-\lambda)}{\lambda} = q(\theta_b) + c$, so $b_b = E(q(\theta_b)+c)$. Recall that $\theta_s(c)$ denotes the unique solution to $1-\theta=E(q(\theta)+c)$. Then, by the increasing nature of $E(.)$ and $q(.)$, we have that the buyers' equilibrium exists if and only if $\theta_b \leq \theta_s(c)$. □

**Proof of Corollary 2.** From Eq. (5) we have that in a sellers' market

$$x_b = \overline{q}(\theta_s) - \frac{c(b^*/(1-\theta_s)-\lambda)}{\lambda}.$$

Substituting in for $b^*$ from Theorem 1 and simplifying, we obtain

$$x_b = q(\theta_s) + c.$$

Differentiating and calculating $\frac{d\theta_s}{dc}$ by implicitly differentiating $E(q(\theta_s)+c)=1-\theta_s$, we obtain

$$\frac{dx_b}{dc} = \frac{-q'(\theta_s)E'(q(\theta_s)+c)}{1+q'(\theta_s)E'(q(\theta_s)+c)} + 1 > 0.$$

Similarly,

$$\frac{d(x_b-c)}{dc} = \frac{-q'(\theta_s)E'(q(\theta_s)+c)}{1+q'(\theta_s)E'(q(\theta_s)+c)} < 0. \quad \square$$

**Proof of Proposition 6.** Sellers are the short side: $b^* > 1-\theta^* = E(x_b)$.

In this scenario, $\pi_b \equiv \frac{1-\theta^*}{b^*}$. Moreover, $x_b = E^{-1}(1-\theta^*)$. Eq. (5) translates into

$$E^{-1}(1-\theta^*) = \overline{q}(\theta^*) - \frac{c(b^* - \lambda(1-\theta^*))}{\lambda(1-\theta^*)}.$$

Solving for $b^*$, we obtain

$$b^* = \lambda(1-\theta^*)\left(\frac{\overline{q}(\theta^*) - E^{-1}(1-\theta^*)}{c} + 1\right). \tag{14}$$

Then we have that the sellers' market solution obtains if and only if

$$1-\theta^* < \lambda(1-\theta^*)\left(\frac{\overline{q}(\theta^*) - E^{-1}(1-\theta^*)}{c} + 1\right). \tag{15}$$

Buyers are the short side: $b^* = E(x_b) \leq 1-\theta^*$.

In this case all the buyers get matched in every period, so $\pi_b \equiv 1$. Substituting back to Eq. (5), we obtain

$$E^{-1}(b^*) = \overline{q}(\theta^*) - \frac{c(1-\lambda)}{\lambda}. \tag{16}$$

For this to be the equilibrium we need

$$E^{-1}(1-\theta^*) \geq \overline{q}(\theta^*) - \frac{c(1-\lambda)}{\lambda}. \tag{17}$$

Notice that this inequality is the complement of Eq. (15). □

**Proof of Proposition 7.** Note that there exists a threshold value of $\theta^*$, such that below it we are in a buyers' market while above it we are in a sellers' market. Assume first, that we are in a buyers' market. From Eq. (1) we have that

$$\frac{\mathrm{d}W}{\mathrm{d}\theta^*} = b^* \overline{q}\,'(\theta^*) + \frac{\mathrm{d}b^*}{\mathrm{d}\theta^*}\left[\overline{q}(\theta^*) - E^{-1}(b^*)\right]. \tag{18}$$

The first term is clearly positive, since the average quality is increasing in the marginal one. Since we are in a buyers' market $E^{-1}(b^*)$ is the expected surplus of a buyer, which is less than the average quality as well (c.f. Eq. (16)). Consequently, all we need in order to complete this part of the proof is to show that in a buyers' market $\frac{\mathrm{d}b^*}{\mathrm{d}\theta^*} \geq 0$. That trivially follows form Proposition 6, since both $E(\theta^*)$ and the average quality are increasing in $\theta^*$. Consequently, welfare is increasing in a (binding) quality constraint. That is, the planner would like $\theta^*$ as large as possible, while it still results in a buyers' market.

Turning to the sellers' market, we have

$$\begin{aligned}
\frac{\mathrm{d}W}{\mathrm{d}\theta^*} &= \frac{\mathrm{d}\left((1-\theta^*)\overline{q}(\theta^*) - \int_0^{1-\theta^*} E^{-1}(x)\mathrm{d}x - \left(\lambda(1-\theta^*)\left(\frac{\overline{q}(\theta^*) - E^{-1}(1-\theta^*)}{c} + 1\right) - (1-\theta^*)\right)c\right)}{\mathrm{d}\theta^*} \\
&= (1-\theta^*)\overline{q}\,'(\theta^*) - \overline{q}(\theta^*) + E^{-1}(1-\theta^*) - c(1-\lambda) + \lambda\left(\overline{q}(\theta^*) - E^{-1}(1-\theta^*)\right) \\
&\quad - \lambda(1-\theta^*)\left(\overline{q}\,'(\theta^*) - \frac{\mathrm{d}E^{-1}(1-\theta^*)}{\mathrm{d}\theta^*}\right) \\
&= \left[(1-\theta^*)\overline{q}\,'(\theta^*) - \overline{q}(\theta^*) - c + E^{-1}(1-\theta^*)\right](1-\lambda) - \frac{\lambda(1-\theta^*)}{E'(E^{-1}(1-\theta^*))} \\
&= \left[E^{-1}(1-\theta^*) - q(\theta^*) - c\right](1-\lambda) - \frac{\lambda(1-\theta^*)}{E'(E^{-1}(1-\theta^*))},
\end{aligned}$$

where we have used that $\overline{q}\,'(\theta^*) = \frac{\overline{q}(\theta^*) - q(\theta^*)}{1-\theta^*}$. This derivative is clearly negative, since the fact that the quality constraint is binding implies that the term in the square brackets is negative. □

**Proof of Corollary 6.** It is straightforward that if we set the quality threshold at the efficient quality and the market happens to clear in such a way that the measure of sellers and buyers in the market is the same — and therefore there is no inefficiency caused either by waiting or by too low demand — we have the first best outcome. Consequently, all we need to prove is that it is possible to set $\lambda$ in order to ensure that the market indeed clears in that way. Recall that the producer of the efficient marginal quality — given by the solution to $q(\theta)=E^{-1}(1-\theta)$, — is independent of both $\lambda$ and $c$. On the other hand, by Proposition 6, we have a symmetric market when $E^{-1}(1-\theta^*) = \overline{q}(\theta^*) - \frac{c(1-\lambda)}{\lambda}$. Using the equality defining the optimal marginal producer, means that we need $q(\theta^c) = \overline{q}(\theta^c) - \frac{c(1-\lambda)}{\lambda}$ or $\overline{q}(\theta^c) - q(\theta^c) = \frac{c(1-\lambda)}{\lambda}$. The left-hand side of this last equation is a number in (0,1). Thus the proof boils down to showing that for any $c \in (0, 1]$, any number in (0,1) can be achieved by a judiciously choosing a $\lambda \in [0, 1]$. This is indeed the case, since $\frac{1-\lambda}{\lambda}$ is decreasing from $\infty$ to 1 in $\lambda$. □

**Proof of Proposition 8.** If we are in a sellers' market, by Eq. (3), the aggregate profits are given by $\pi = (1-\lambda)\int_{\theta^*}^1 [q(\theta) - x_b + c]\mathrm{d}\theta$. Recall that $x_b = E^{-1}(1-\theta^*)$. Substituting into the profit function and differentiating, we obtain the first-order condition.

In a buyers' market not all sellers get matched, so $\pi = \frac{b*}{1-\theta*}(1-\lambda)\int_{\theta*}^{1}[q(\theta) - x_b + c]d\theta = \frac{E\left(\overline{q}(\theta*) - \frac{c(1-\lambda)}{\lambda}\right)}{1-\theta*}(1-\lambda)\int_{\theta*}^{1}\left[q(\theta) - \overline{q}(\theta*) + \frac{c}{\lambda}\right]d\theta$, where we used $x_b = E^{-1}(b*)$ and $b* = E\left(\overline{q}(\theta*) - \frac{c(1-\lambda)}{\lambda}\right)$. Integrating and simplifying, we get $\pi = E\left(\overline{q}(\theta*) - \frac{c(1-\lambda)}{\lambda}\right)(1-\lambda)c/\lambda$. Since the average quality is increasing in the marginal producer, the seller's cartel will never want to be in a buyers' market. □

**Proof of Corollary 7.** Aggregate profits are given by

$$\frac{e}{1-\theta^*}(1-\lambda)\int_0^1\left[q(\theta) - \overline{q}(\theta^*) + \frac{c}{\lambda}\right]d\theta = e(1-\lambda)\frac{c}{\lambda},$$

which is clearly constant in the marginal quality. □

## References


Anwar, Ahmed, Sákovics, József, 2007. A decentralized market for a perishable good. B.E. Journal of Theoretical Economics 7 (1) (Contributions), Article 7.

Arrow, Kenneth J., 1963. Uncertainty and the welfare economics of medical care. American Economic Review 53 (5), 941–973.

Bester, Helmut, 1988. Bargaining, search costs and equilibrium price distributions. Review of Economic Studies 55, 201–214.

Bester, Helmut, 1993. Bargaining versus price competition in markets with quality uncertainty. American Economic Review 83 (1), 278–288.

Binmore, Kenneth, Rubinstein, Ariel, Wolinsky, Asher, 1986. The Nash Bargaining Solution in economic modelling. RAND Journal of Economics 17, 176–188.

Blouin, Max R., 2003. Equilibrium in a decentralized market with adverse selection. Economic Theory 22, 245–262.

COM, 2004. Report on Competition in Professional Services. Communication from the EU Commission 83, http://europa.eu.int/comm/competition/liberal_professions.

Davis, Steven J., 2001. The quality distribution of jobs and the structure of wages in search equilibrium. NBER Working Paper, no. 8434.

Gale, Douglas, 1987. Limit theorems for markets with sequential bargaining. Journal of Economic Theory 43, 20–54.

Fershtman, Chaim, Fishman, Arthur, 1994. The 'perverse' effects of wage and price controls in search markets. European Economic Review 38, 1099–1112.

Kranton, Rachel, 2003. Competition and the incentive to produce high quality. Economica 70, 385–404.

Leland, Hayne, 1979. Quacks, lemons, and licensing: a theory of minimum quality standards. Journal of Political Economy 87, 1328–1346.

OECD, Directorate for Financial, Fiscal and Enterprise Affairs, Committee on Competition Law and Policy, 23 February 2000. "Competition in Professional Services" http://www.oecd.org/dataoecd/35/4/1920231.pdf.

Paterson, Iain, Fink, Marcel, Ogus, Anthony, 2003. The economic impact of regulation in the field of liberal professions in different Member States. Research report for the European Commission. http://europa.eu.int/comm/competition/publications/publications/index.html#liberal.

Ponsatí, Clara, Sákovics, József, 1998. Rubinstein bargaining with two-sided outside options. Economic Theory 11, 667–672.

Shapiro, Carl, 1986. Investment, moral hazard, and occupational licensing. Review of Economic Studies 53 (5), 843–862.

Stigler, George, 1971. The theory of economic regulation. Bell Journal of Economics 2, 3–21.

Stiglitz, Joseph, E., 1979. Equilibrium in product markets with imperfect information. American Economic Review 69 (2), 339–345.

Wilson, Charles, 1979. Equilibrium and adverse selection. American Economic Review 69 (2), 313–317.

Wilson, Charles, 1980. The Nature of equilibrium in markets with adverse selection. Bell Journal of Economics 11 (1), 108–130.