

Selection, Heterogeneity and Entry in Professional Markets*

Marco Alderighi^a

Università della Valle d'Aosta, Italy.

Università Bocconi, Milano, Italy.

Claudio A. Piga^b

University of Keele, UK.

RCEA, Rimini, Italy.

May 13, 2013

Abstract

We analyze two different cases of entry regulation in professional markets: first, when licensing is a requirement for becoming a professional (lawyers); second, when entry and price restrictions are applied on a geographical basis (pharmacists). Both cases are investigated within a circular model of localized competition and heterogeneous players. The analysis reveals that licensing introduces a selection mechanism which is effective in preventing entry of inefficient players in markets with large ex-ante heterogeneity. Furthermore, because in the second case excessive entry is reduced as the degree of heterogeneity increases, our analysis lends support to a policy that simultaneously relaxes entry and price restrictions.

JEL classification: L11, D61.

Keywords: Liberal professions; Cost/quality heterogeneity; Entry regulation; Localized competition; licensing.

*We would like to thank the Editor, the Coeditor and two anonymous referees of this Journal for their valuable suggestions and comments. Paul Dobson, André de Palma, Francesco Passarelli, Joanna Poyago-Theotoky, Jacques Thisse, Chris Wilson and the participants of 2010 IIO Conference in Vancouver, of the 2010 EARIE Conference in Istanbul, of the CCRP Workshop organized by The City University, London, and all seminar participants at CORE (Belgium) and York (UK) also kindly provided useful feedback.

^aE-mail: m.alderighi@univda.it

^bCorresponding Author. Keele Management School, University of Keele, Staffordshire, ST5 5BG. E-mail: claudio.piga@gmail.com

1 Introduction

The degree of efficiency in markets for such liberal professions as lawyers, notaries, architects, engineers and pharmacists has been the subject of extensive investigation in Europe (European Commission, 2004).¹ The high level of regulation characterizing the European market, in the form of either State regulation or self-regulation by professional bodies, has been deemed by some as unnecessary and harmful, while by others as compelling and vital. For the latter, regulation is usually sustained by arguments associated with the so-called “public interest view” (i.e., regulation addresses market failures due to asymmetric information, and/or externalities and/or public good provision). In contrast, those supporting the “private interest approach” claim that, due to regulatory capture, many regulatory mechanisms may serve the private interests of professional bodies’ members more than those of the general public (Becker, 1983; Peltzman, 1976; Philipsen, 2009; Posner, 1974; Stigler, 1971). For instance, in the case of entry restrictions, a requirement to have demonstrated basic competence is clearly justified where consumers are not well placed to assess the quality of service, as is usually the case in markets for professional services. This has to be weighed against the risk that necessary qualification thresholds may be set too high to constrain entry and thus benefit incumbent producers, who largely control the professional bodies.

Based on the premise that in the real world some degree of firm heterogeneity is the rule rather than the exception, in this study we deviate from the usual market failure versus regulator capture arguments to provide a different and, in some sense,

¹According to the EU Directive on Recognition of Professional Qualifications (2005/36/EC), liberal professions are “those practised on the basis of relevant professional qualifications in a personal, responsible and professionally independent capacity by those providing intellectual and conceptual services in the interest of the client and the public”.

complementary model for liberal professions, which allows an evaluation of the welfare properties associated with 1) the entry requirement of licensing, 2) the entry restrictions capping the maximum number of firms and 3) fixed or minimum prices.

Even within the European Union, professional regulation presents significant differences among countries and professions, although there are also some important similarities worldwide. As far as entry restrictions are concerned, for all professions a university degree in the relevant field is required; for a subset of them (lawyers) the exercise of the profession is conditional on the further acquisition of a licence which is obtained by passing an examination and on spending a period of apprenticeship under the supervision of a qualified professional (Kleiner, 2000). For other professions, e.g. pharmacists, no licensing is needed; in this case, however, rules on entry based on demographic and geographic criteria often make it impossible to open a new outlet in markets where such criteria are binding (Schaumans and Verboven, 2008).²

We propose a set-up that accommodates two main cases based on whether or not a licence is required for the entry in the liberal profession. When licensing is not required, entry is modelled as a two-stage game, where prospective entrants incur a set-up sunk cost before starting production and the related competitive stage. To account for licensing, we introduce a preliminary stage, where prospective professionals have to incur the licensing sunk cost before deciding whether to start their activity or not. The opportunity cost of the time spent to complete the apprenticeship period largely determines the magnitude of the licensing cost. Thus, entry in professions where licensing is required entails a three-stage process and two different

²Licensing and quantitative entry restrictions are not always mutually exclusive. For instance in Italy they are both used to regulate entry into the notary profession, which is generally found to be highly regulated across Europe (Philipsen, 2009).

sunk costs (i.e., the licensing cost as well as the production set-up cost as in the no-licensing case)

Both types of entry games (i.e., with licensing and without) are analyzed in this study by modeling the relevant market as a circular city model of localized competition with heterogeneous costs' firms (Salop, 1979; Syverson, 2004; Vogel, 2008). To remain in keeping with the terminology used in the existing literature, we will use the terms "professional" and "firm" interchangeably.

The two games have different informational structures, which impact on the characteristics of the ensuing equilibrium. In the two-stage game, before moving to the final stage, players have no information about their own production costs, while in the licensing game, a professional learns her own cost after the licensing phase. Thus, this second model entails a selection mechanism: only those professionals with sufficiently low production costs will decide to enter the production stage by paying the second set-up sunk cost.

We study the properties of the selection mechanism by focusing on how it responds to differences across markets in their supply-side characteristics. In particular, we investigate whether the upper threshold of the ex-post cost distribution determined in the selection stage reacts to the degree of ex-ante cost heterogeneity characterizing the potential entrants in a market. That is, we ask whether the selection mechanism is capable to weed out the least efficient professionals precisely in those markets where there is a larger risk that relatively inefficient ones may enter.

Since the selection mechanism occurs only in the licensing game, we thus investigate a potentially beneficial effect of maintaining licensing as an entry requirement in liberal professions. Our findings indeed point out that following the licensing stage, the selection mechanism induces a truncation from above of the ex-ante cost distribu-

tion and that such truncation intensifies as the degree of ex-ante cost heterogeneity increases. That is, the maximum level of cost draw guaranteeing non-negative profits for an operative professional is inversely related with the ex-ante cost variance. Our model thus predicts that only the professional with lower costs will decide to become a professional and those with higher costs will move to other jobs as, e.g., in Leland (1979): “Doctors (or potential doctors) [...] may not be willing to remain in (or enter) the market.. (p. 1329)”. Thus, our analysis suggests that licensing may play a crucial role in determining high levels of efficiency across markets, and especially in those markets where heterogeneity of prospective professionals is large. Furthermore, this result provides an additional rationale for licensing, which, unlike previous contributions, is obtained in the absence of informational asymmetry between the service providers and the buyers (Kleiner, 2000; Leland, 1979; Shapiro, 1986; Philipsen, 2009).

This model can be applied more in general to describe the entry decisions in markets with localized competition. In this respect, our three-stage entry game is related with the work of Syverson (2004). Apart from an alternative modelling of the pricing competition stage, the focus of the two papers is rather different. Syverson (2004) investigates how such demand-side characteristics as market density can determine productivity differences across markets, while our main focus is on the supply-side characteristics (i.e., cost heterogeneity) and the role of a regulatory mechanism (i.e., licensing).

In the two-stage game without licensing we study the inter-play between a type of regulatory restriction on entry and pricing, by focussing in particular on the insights that can be gained by extending the analysis to the case of cost heterogenous firms. Deneckere and Rothschild (1992) highlight the general property of excessive

entry in circular city model of localized competition amongst homogeneous cost firms. By introducing heterogeneous firms who decide whether to enter without knowing their rivals' costs, we show that excessive entry reduces as the degree of cost heterogeneity increases. This is because the first-best number of firms increases with cost heterogeneity faster than the free-entry one. Therefore, the gap between the number of firms in the free-entry market equilibrium and the socially optimal number of firms shrinks as cost heterogeneity expands. Our analysis also indicates that uniform pricing, which may arise from recommended and fixed prices by professional bodies, is likely to be a particularly harmful form of regulatory restriction.

The next Section describes real-world examples of entry restrictions in professional markets; it is followed by the model's set-up and the characterization of the equilibrium prices as well as the condition for a unique Nash equilibrium in pure strategies. The three-stage entry game with a selection mechanism is developed in Section 4, which is followed by the two-stage entry game without licensing. Section 6 concludes. The Appendix contains the proofs of all the Propositions, Lemmata and Corollaries.

2 Entry qualifications in professional markets

In many countries, before an individual is allowed to practice a profession, a certain set of regulatory conditions must be fulfilled (Philipsen, 2009). Examples of measures include minimum periods of education and/or professional experience, mandatory registration, establishment requirements, licensing.

As a starting example, consider the legal profession in England and Wales, where only those who are qualified as a lawyer (solicitor or barrister) may represent parties

before a Court; note, however, that anyone, whether a qualified lawyer or not, is able to give legal advice: extra-judicial legal advice is often given by non-lawyers in, for instance, the fields of tax, business and planning. To enter the solicitor's profession, candidates must undertake three phases of training, which are described in the Solicitors' Act 1974: an academic phase, a vocational one and a training contract (Vickers, 2001, p. 52). The academic phase consists of a three-year undergraduate degree in law; this is not binding, since individuals with a different undergraduate degree may undertake a one-year extra full-time programme and be admitted to the next stage if they pass the Common Professional Examination. The vocational step (phase two) consists of the Legal Practice Course (LPC), a one-year postgraduate qualification which is accredited by the Law Society, the professional association and governing body of solicitors whose main duties include dealing with complaints against solicitors and disciplinary matters, and issues practising certificates. Finally, there is a two-year training contract (salaried) with a firm of solicitors, during which the trainees gain experience of the practice of law in three areas of their choosing under the supervision of a qualified solicitor. The training also covers topics which, according to the Law Society, are best studied once students have some work experience (advocacy and oral communication, financial awareness and business accounts, and ethics and client responsibilities). After fulfilling the training requirements above, all solicitors in private practice must obtain a practising certificate (issued annually) from the Law Society. To become a barrister, a similar process involving three steps with a final pupillage period has to be followed. As far as the legal profession is concerned, the above licensing process entailing, in addition to a relevant university degree, the fulfilment of two mandatory requirements (passing of an examination testing the necessary entry level competencies, plus an apprentice-

ship period) appears to be largely prevalent in the EU, as the cases of France, Italy and Germany in Paterson *et al.* (2007) indicate.³ However, in the USA prospective lawyers only have to sit the state bar examination as well as the Multistate Professional Responsibility Examination to satisfy the moral character and fitness standards in that state; in 43 jurisdictions they are required to take continuing education courses following bar admission, i.e., there is no compulsory traineeship period (see <http://international.lawsociety.org.uk/ip/americas/610/practise>).

The accountant profession offers another important distinction. In the UK, France, Netherlands, Germany and Italy, the licensing model above applies only to specific higher-level tasks, such as statutory auditing, which are exclusively reserved to individuals possessing the necessary requirements and who are registered with the relevant Recognised Supervisory Board (e.g., the Institute of Chartered Accountants in England and Wales - ICAEW, or “La Compagnie Nationale des Commissaires aux Comptes - CNCC” in France). Other services (e.g., internal audit, advice on financial controls, preparation of management accounts, bookkeeping) are not regulated and can be performed by anyone (Vickers, 2001; Paterson *et al.*, 2007). For these non-statutory services, the licensing model is often replaced by a certification system (Shapiro, 1986), where members to a professional accountancy body (one which does not act as a Recognised Statutory Boards for statutory audit) are individuals who have voluntarily enrolled themselves in a study programme managed by the body itself; the body also acts as the guarantor for the programme’s quality standards.⁴

³Similarly, in Upper Canada the licensing process consists of two components, which are taken after the completion of a university degree: (a) licensing examinations; and (b) an articling program, which is ten months in length. An online “professional responsibility and practice course” must be completed during the articling term and requires completion of an assessment with the licensee’s articling principal - see www.lsuc.on.ca/LicensingProcessSiteMap

⁴In England, two of these bodies have chartered status, i.e., they have received a formal State

A third system of entry into a profession is one where a university degree in a relevant programme is the only required qualification. This was the case for Spanish lawyers until November 2011: law graduates were automatically eligible for admission to the local bar association, without any further requirements for training or examinations. Under the new regime, specified in the Royal Decree of 16 June 2011, law graduates will also need to complete a 2-year work experience placement and pass a state exam (see <http://international.lawsociety.org.uk/ip/europe/580/practise>). To become a pharmacist in Portugal, Sweden, Italy and Belgium, no license is required; this is not the case in Germany and Ireland, where the three-phases licensing process applies (Paterson *et al.*, 2007; Schaumans and Verboven, 2008).

3 The model

Consider a circular city of unitary length with uniform density D . There are $N \geq 2$ equidistant firms located in $n \in \mathbb{L} \equiv \{1, 2, \dots, N\}$. Firms have different unit variable costs $c_n \geq 0$ with $c_n \in [c_L, c_H]$ and offer an identical quality level $\theta > 0$.⁵

The utility a consumer, who is located at distance $d \in [0, 1/N]$ from firm n , obtains from buying a service is $U_{d,n} = v - td - p_n$, where t is the unit transport cost, p_n is the uniform price charged by firm n and v is the reservation value (which absorbs the quality level). We assume v sufficiently high so that each consumer buys a unit of the good and the market is fully covered.

Let p_n , p_{n-1} and p_{n+1} denote the prices charged by firm n and its two immediate recognition. They are the Chartered Institute of Management Accountants (CIMA) and the Chartered Institute of Public Finance and Accountancy (CIPFA).

⁵ Under quality observability, this set-up is isomorphic to one where firms' heterogeneity is modeled by allowing firms to have identical unit cost c but different quality levels $\theta_n \in [\theta_L, \theta_H]$, when $c_n = c - \theta_n \in [c_L = c - \theta_H, c_H = c - \theta_L]$ for any n .

neighbors.⁶

The marginal consumer between n and $n + 1$ ($n - 1$) lies at a distance S_n^R (S_n^L) from firm n , where:

$$S_n^R = \frac{p_{n+1} - p_n}{2t} + \frac{1}{2N}; \quad S_n^L = \frac{p_{n-1} - p_n}{2t} + \frac{1}{2N} \quad (1)$$

Standard computations yield the set of first order conditions (Tirole, 1988, p. 283):

$$p_n = \frac{2c_n + p_{n-1} + p_{n+1}}{4} + \frac{t}{2N}, \quad n \in \{1, 2, \dots, N\}. \quad (2)$$

3.1 Price Equilibrium

Following Eaton and Lipsey (1978), we keep the requirement that competition is localized, that is, each firm competes with its two adjacent firms on the captive market represented by the closer consumers located between them.

Condition 1 (*No mill-price undercutting*) $S_n^R \geq 0$; $S_n^L \geq 0$, for $n \in \mathbb{L}$.

Firms do not reduce prices in such a way to grab all the market of one or both 1-step neighbors. From a technical viewpoint, Condition 1 rules out the possibility that an inefficient firm is driven out of the market when it competes against highly efficient neighboring rivals; it therefore imposes a limit on the maximum allowable amount of cost heterogeneity in the model (Alderighi and Piga, 2010).

The price equilibrium with heterogenous firms presents a number of characteristics which do not feature in the standard, homogenous firms case. Such characteristics are now briefly illustrated, since they describe the equilibrium of the last stage

⁶ With a little abuse of notation, we use the convention that if $n \pm i \notin \mathbb{L}$, with $i = 1, \dots, N - 1$, the index refers to a firm $n \pm i \mp N \in \mathbb{L}$.

of the two entry games on which we focus the attention in this study.⁷

Define $\underline{l}(N) = \lceil \frac{1-N}{2} \rceil$, and $\bar{l}(N) = \lceil \frac{N-1}{2} \rceil$, where $\lceil x \rceil$ is the approximation of x to its larger integer.⁸

Lemma 1 *If Condition 1 is satisfied, then system (2) has a unique solution. The market equilibrium prices are given by:*

$$p_n^* = \sum_{i=\underline{l}}^{\bar{l}} w_{\bar{l}-|i|} c_{n+i} + k, \quad n \in \mathbb{L}, \quad (3)$$

where w_i -s and k are constants.

Despite the localized competition assumption, the market equilibrium arises from the interaction of a sequence of chain-linked, inter-locked sub-markets (Rothchild, 1982), where each firm's cost affects the pricing of both direct and non-direct competitors through a transmission mechanism whose properties are defined by the weights in the following Lemma.

Lemma 2 *For any $N \geq 2$ the weights w_i and k in Lemma 1 are:*

$$w_i = 4w_{i-1} - w_{i-2} \text{ for } i = 1, \dots, \bar{l} - 1 \quad (4)$$

$$w_1 = (3 - \underline{l} - \bar{l}) w_0 \quad (5)$$

$$4w_{\bar{l}} = 2 + 2w_{\bar{l}-1} \quad (6)$$

$$\sum_{i=\underline{l}}^{\bar{l}} w_{\bar{l}-|i|} = 1, \quad (7)$$

$$k = t/N. \quad (8)$$

First, weights decrease with distance, so that a shock in a firm's costs propagates throughout the market but has a larger impact on the price of its closer competitors.

⁷For more on the characteristics of this pricing equilibrium, see Alderighi and Piga (2012), which focuses on its geographical properties and on how these impact on the competitive process among vertically related firms.

⁸Through this notation, we can effectively manage the cases of N being odd or even, plus take advantage of the symmetry of the locations around the two halves of a circle.

Second, because weights sum to one, (3) is a generalization of the price equilibrium in the standard case with identical cost firms. Third, weights are independent of the cost differential between any pairs of firms, implying, for instance, that a high-cost firm will assign the same weight to its own cost regardless of whether it faces a high- or low-cost direct competitor.⁹

Lemma 3 *Equations (3)-(8) represent a unique Nash price equilibrium in pure strategies when Condition 1 is satisfied, i.e., if:*

$$c_H - c_L < \rho_c(N) k, \quad (9)$$

where $k = t/N$ and $\rho_c(N) = (w_{\bar{1}} - w_0)^{-1}$.

Eq. (9) indicates the maximal cost heterogeneity such that there always exists an indifferent consumer located between two neighboring producers, i.e. a high cost firm facing tough competition from two low-cost neighboring rivals always has a non-negative market share. Unlike the homogenous firms' case, the vector of market prices (3) induces a distortive allocative outcome where consumers who should patronize a low-cost firm end up buying from a higher cost firm. Indeed, for any N :

Lemma 4 *The first-best full information allocative solution can be obtained by setting:*

$$p_n^F = c_n + k_F, \quad (10)$$

where k_F can be set freely with the only caution that the delivered price does not exceed the reservation price of consumers.

⁹ To calculate the numerical values of the weights in (4)-(7), we use the following approximation which holds for any i when N is large: $w_{\bar{1}} = y = 1/\sqrt{3}$; $w_{\bar{1}-i}/w_{\bar{1}-i-1} = x = 2 + \sqrt{3}$; $w_{\bar{1}-i} = y \cdot x^{-i}$.

From Lemmata 1, 2 and 4, after setting $k_F = k = t/N$ we can compare equilibrium prices under the free market and first-best scenarios (for any N). Because in the former case firms base their prices not only on their own costs (as in the latter situation) but also on those of the opponents, it follows that in the market equilibrium efficient low-cost firms charge prices above the first-best solution and inefficient high-cost ones charge prices below it. A productive inefficiency arises because high-cost (low-cost) firms enjoy a larger (smaller) than optimal market share.

4 Entry with licensing

We model entry in liberal professions as either a three-stage or a two-stage game, depending on whether the licensing requirement is present or not. The former case is developed in this Section, while the latter in the next one. The last stage of both types of game corresponds to the price competition of the previous Section.

There is an important difference between the two games. Licensing entails the acquisition of information on a firm's own level of efficiency which may induce some firms to abandon the market before the actual entry stage, after they compare their marginal cost (or intrinsic ability) with the average marginal cost that is expected to prevail in the market. In other words, licensing triggers a selection mechanism that restricts access to those firms whose cost levels fall in the lower part of the ex-ante industry distribution of costs. While this truncation of the cost distribution from above has been already discussed within a different informational set-up in Syverson (2004), a novel result in this Section highlights the effectiveness of the selection mechanism in relation to the degree of firm heterogeneity in the market. More precisely, we show that the cost threshold below which actual entry takes place decreases as the ex-ante cost heterogeneity increases.

The timing of the licensing game is as follows:¹⁰

Stage 1 - Licensing; Out of a pool of Λ potential candidates, $M \leq \Lambda$ prospective entrants decide to fulfill the legal requirement of obtaining a licence to exercise a given profession; by investing F_L they acquire private information on their own cost but not on that of their rivals;

Stage 2 - Selection and actual entry; Firms decide whether to exit (i.e., they pursue an outside option) or to enter the professional market by incurring a set-up cost F_p to start production;

Stage 3 - Full Information Pricing; Prices are set under a full information scenario; i.e., according to the Lemmata 1 and 2.

The first stage reflects the fact that not all the individuals with a relevant degree seek to become licensed professionals. Stage 2 entails that some firms abandon the market even if they have fulfilled all the necessary formal requirements. For example, in both the legal and the medical profession it is not uncommon to observe individuals opting not to enter the profession even after obtaining the license (Leland, 1979; Pashigian, 1977).

The value of F_p includes such expenses as, for instance, the present discounted value of fees paid for the mandatory membership to Professional Bodies, the time spent to set up the practice, to hire collaborators, etc.

The licensing cost F_L arises as a consequence of a number of aspects. While the cost to acquire a university degree is not sunk, as the degree can be used to

¹⁰There is an important difference between the structure of our game and that presented in Syverson (2004) where firms set their prices based only on their cost type but not on their rivals' actual cost realizations. While this may be a realistic assumption in the short-run, our approach closely mimics a complete information, long-run equilibrium where firms have learnt to set prices from which they would not unilaterally want to deviate (which, given Lemmata 1 and 2, implies a full knowledge of all the firms' costs and locations as in Vogel, 2008).

pursue a number of alternative career options, during the apprenticeship period a prospective professional, e.g., someone who has recently gained a degree in Law, has to combine the general know-how obtained from the university studies with the acquisition of the knowledge of both job-specific skills and practical aspects of the profession. This process, which determines a professional's intrinsic productivity, is highly idiosyncratic and therefore is unlikely to yield a purely deterministic outcome, in the sense that, for instance, individuals with good academic credentials may not necessarily be very apt to tackling the more practical or relational aspects of the profession. Indeed, as noted by Holmstrom (1999), productive abilities are revealed over time through the observation of performance, and therefore by the end of the apprenticeship period, prospective professionals become aware of their efficiency (i.e., production cost). The magnitude of the licensing cost is therefore represented by the opportunity cost of foregoing remunerated job opportunities during the apprenticeship period. This is assumed to be the same for all the prospective professionals, since they all have an identical outside option, as it would be the case for graduates with little or no previous work experience.

To derive the properties of the entry game equilibrium, we assume that costs are identically and independently distributed:

Assumption 1 *Let $\tilde{c}_n \in [c_L, c_H]$ be a random variable with cumulative distribution G , density g , mean value $E[c_n] = \bar{c}$, variance $E[(c_n - \bar{c})^2] = \sigma^2$ and covariance $E[(c_n - \bar{c})c_m] = 0, \forall n, m \in \mathbb{L}$ and $n \neq m$.*

We solve the model by backward induction. The firms' profits in stage 3 are computed using Lemmata 1 and 2 and the restriction (9) on maximal cost heterogeneity is maintained.

4.1 Selection

In the second stage, a generic firm n knows c_n , M and the prior distribution of costs G . Because costs are randomly and independently distributed, learning its own cost does not change a firm's beliefs on its rivals' costs. Nonetheless, in the selection phase firms can form their beliefs as to the cost distribution that will emerge in the price competition stage, where only a subset $N \leq M$ of firms may be involved. Based on such beliefs, each firm assumes a posterior distribution of the opponents' costs G_e , which reflects each firm's beliefs on its rivals' decision on whether to stay or exit the market. That is, the prior and posterior distribution may differ because with a large cost heterogeneity, the producers drawing a sufficiently high cost expect to gain profits that are below the fixed set-up cost F_p and therefore abandon the market before entering production.

Under the assumption that there is at least one opponent that is going to enter the market, the expected profit (gross of pre-entry fee F_L) of firm n , if it enters the market (together with $N - 1 \geq 1$ other competitors) is:

$$\begin{aligned} E_{G_e} \left[\tilde{\Pi}_n | c_n, N, M \right] &= \frac{D}{2t} \cdot E_{G_e} [(\tilde{p}_n - c_n) \cdot (\tilde{p}_{n-1} + \tilde{p}_{n+1} - 2\tilde{p}_n + 2k)] - F_p \\ &= \frac{D}{t} \cdot E_{G_e} [(\tilde{p}_n - c_n) \cdot (\tilde{p}_{n-1} - \tilde{p}_n + k)] - F_p, \end{aligned} \quad (11)$$

where E_{G_e} means that the expectations are taken using firm n 's posterior cost distribution on the actual entrants, G_e and the superscript \sim denotes stochastic variables. The second line in (11) derives from the fact that in stage 2 firm n treats \tilde{p}_{n-1} and \tilde{p}_{n+1} in an identical manner. Further, even if firm n is a monopolist, its profit must be bounded from above, since it is limited by the consumers' willingness to pay v . Therefore:¹¹

¹¹The case of monopoly can generate problems of existence of equilibria when the market is too

$$E_{G_e} \left[\tilde{\Pi}_n | c_n, 1, M \right] = D (v - c_n - t/2) - F_p < \infty. \quad (12)$$

Define $\bar{c}_e = E_{G_e} (c_{n+i})$, $\sigma_e^2 = E_{G_e} (c_{n+i} (c_{n+i} - \bar{c}_e))$, $\forall i \neq 0$.

Lemma 5 *Under Assumption 1,*

a) *for $N = 1$, the expected profit of firm n is given by (12); for $N \geq 2$:*

$$E_{G_e} \left[\tilde{\Pi}_n | c_n, N, M \right] = \frac{D}{t} \left((w_d (\bar{c}_e - c_n) + k)^2 + W_L(N) \sigma_e^2 \right) - F_p, \quad (13)$$

where $w_d = (1 - w_l) = (w_l - w_{l-1})$, $W_L(N) = \sum_{i=\mathbb{L}_0} w_{l-|i|} (w_{l-|i+1|} - w_{l-|i|}) > 0$, $\mathbb{L}_0 \equiv \mathbb{L} \setminus \{0\}$, $\lim_{N \rightarrow \infty} W_L(N) = \bar{W}_L = \frac{2}{9}\sqrt{3} - \frac{1}{3} \simeq 0.051567$ and $\lim_{N \rightarrow \infty} w_d = (1 - 1/\sqrt{3}) \simeq 0.42265$.

b) $E_{G_e} \left[\tilde{\Pi}_n | c_n, N, M \right]$ *is decreasing in c_n and F_p , and increasing in \bar{c}_e , σ_e^2 and D .*

Values of $W_L(N)$ are given in Table 1. Provided that cost differentials are sufficiently small, Lemma 5 holds for each firm and for each possible N . Thus, sequential rationality leads to the conclusion that each firm expects that the opponents deciding to enter the market are sufficiently efficient, i.e., their cost is not greater than a threshold level α . Therefore, the posterior distribution of such firms' costs corresponds to the prior distribution $G(x)$ truncated at α . That is, $G_e(x) = G_\alpha(x) = \min \{G(x)/G(\alpha), 1\}$.

Conjectures on $\alpha \in [c_L, c_H]$ also affect the number of firms that participate to the production stage: $\tilde{N} = \tilde{N}(\alpha, M)$. Since cost distributions are independent, we compute the probability that there are $\eta \leq M$ (including n) firms whose costs fall

small for two firms but large enough for one firm. See for example Levin and Peck (2003) for a detailed analysis on entry with at least two potential entrants.

Table 1: The sequence of equilibrium parameters in the Entry Games (percentage values).

N	2	3	4	5	8	12	20	∞
W_I	22.222	24.000	23.611	23.269	23.030	23.020	23.020	23.020
W_L	11.111	8.000	6.250	5.540	5.169	5.157	5.157	5.157

under the entry threshold α :

$$\omega_\eta = \Pr(\tilde{N} = \eta) = \binom{M}{\eta} G(\alpha)^\eta (1 - G(\alpha))^{M-\eta}.$$

It follows that if firm n decides to enter the market, provided that in the first stage there are M firms, its expected profit is:¹²

$$\sum_{\eta=1}^M \omega_\eta E_{G_\alpha} \left[\tilde{\Pi}_n | c_n, \eta, M \right], \quad (14)$$

which represents the expected profit in the selection phase for given c_n , M and α . Equivalently, for M large, the central theorem guarantees that $\tilde{N}(\alpha, M) \rightarrow N(\alpha, M) = G(\alpha)M$ since $\omega_N \rightarrow 1$; (14) becomes:

$$E_{G_\alpha} \left[\tilde{\Pi}_n | c_n, M \right] = E_{G_\alpha} \left[\tilde{\Pi}_n | c_n, G(\alpha)M, M \right] \quad (15)$$

In the subsequent analysis, we assume M large so that the expected profit can be approximated by equation (15). Note that $E_{G_\alpha} \left[\tilde{\Pi}_n | c_n, M \right]$ retains most of the properties of $E_{G_e} \left[\tilde{\Pi}_n | c_n, N, M \right]$ in Lemma 5, i.e.

Lemma 6 $E_{G_\alpha} \left[\tilde{\Pi}_n | c_n, M \right]$ is increasing in D and decreasing in c_n , F_p and M .

On the one hand, we are mainly interested in the comparative static analysis

¹²When firm n decides to enter the market it has costs lower than or equal to α and expects that the other firms entering the market have costs lower than or equal to α .

concerning the impact of a change in cost heterogeneity on the expected profit; on the other, this is not a straightforward task to investigate because the expected profit in (15) is not only affected by the variance of the prior cost distribution σ , but it depends on \bar{c}_α , σ_α and $G(\alpha)$, i.e., the mean, the variance and the shape of the posterior distribution. In order to avoid unnecessary complications, we assume that G is uniformly distributed.¹³

Assumption 2 Let $\mathcal{G} = \{G^\sigma, \sigma \in (0, \hat{\sigma})\}$ be a family of uniform distributions with average \bar{c} and variance σ^2 .

Define $\Phi(\alpha, M, \sigma) = E_{G_\alpha^\sigma} [\tilde{\Pi}_n | \alpha, M]$ as the expected profit of the least efficient firm ($c_n = \alpha$) remaining in the market of size $N = G^\sigma(\alpha)M$ after selection.

Lemma 7 $E_{G_\alpha} [\tilde{\Pi}_n | c_n, M]$ is decreasing in α .

Since $E_{G_\alpha} [\tilde{\Pi}_n | c_n, M]$ is decreasing in α the expected profit of a generic firm n is negatively affected by an increase in α and the subsequent higher number of competitors, even if these are on average slightly less efficient. Thus, if $\Phi(c_H, M, \sigma) < 0$, then for every M , there exists an α^* , such that: $E_{G_{\alpha^*}} [\tilde{\Pi}_n | c_n, M] \leq 0$ when $c_n \geq \alpha^*$. Therefore, the optimal entry threshold α^* is implicitly given by:

$$E_{G_{\alpha^*}} [\tilde{\Pi}_n | \alpha^*, M] = 0 \quad (16)$$

Lemma 8 Given (16), under Assumption 2 and $\Phi(c_H, M, \sigma) < 0$:

$$\frac{d\alpha^*}{dM} < 0; \quad \frac{\partial \alpha^*}{\partial D} \Big|_M > 0. \quad (17)$$

¹³The following analysis holds for a broader class of distributions, see: Alderighi and Piga (2012b).

Moreover, when $\Phi(c_H, M, \sigma) > 0$, $\alpha^* = c_H$ so that: $\frac{d\alpha^*}{dM} = 0$; $\frac{\partial\alpha^*}{\partial D}|_M = 0$. Note how the second inequality of (17) seems to suggest that an increase in the market density D allows more inefficient firms to profitably stay in the market. However, this is not an equilibrium result for the full game because the partial derivatives in Lemma 8 are calculated for fixed M ; hence they do not capture how a change in D affects the equilibrium number of firms in the pre-entry stage.

4.2 Licensing stage

The number of firms entering the licensing stage is obtained by assuming that firms correctly anticipate the outcomes of both the selection and the pricing stages. Prior to paying the licensing fee F_L , firm n is unaware of its own type c_n , so it can base its entry decision only on the expected profit it would gain, which is a function of the endogenously determined number of entrants M into the licensing stage. All firms acquiring a licence enter the selection stage, but only if $c_n \leq \alpha^*$, firm n starts production after paying F_p , gaining $E_{G_{\alpha^*}}[\tilde{\Pi}_n|c_n, M]$; otherwise it stays out and gains 0.

From (15) and Assumption 2, the expected profit of a firm deciding to enter the licensing stage, gross of the entry fee F_L , depends, amongst other things, on the prior and posterior distributions of costs G and G_{α^*} and on the value of α^* :

$$E[\tilde{\Pi}_n|M] = \int_{c_L}^{\alpha^*} E_{G_{\alpha^*}}[\tilde{\Pi}_n|c_n, M] dG(c_n) \quad (18)$$

Lemma 9 *Under Assumption 2, $E[\tilde{\Pi}_n|M]$ is monotonically decreasing in M .*

Therefore, we can use the zero-profit condition to determine the equilibrium number of entrants in the first stage, M_S :

$$E[\tilde{\Pi}_n|M_S] = F_L, \quad (19)$$

and $N_S = M_S G(\alpha^*(M_S))$. Under Lemma 9 the solution is unique.

4.3 Main Results

The next two Propositions report comparative static results: the first constitutes the main thrust of this Section and shows new insights into the functioning of the selection mechanism by deriving predictions on the relationship between the truncation point of the prior distribution α^* and the variance of the same distribution, the latter being a measure of firm heterogeneity; the second analyzes the link between selection and market density.

Proposition 1 *Under Assumption 2:*

$$\frac{dM_S}{d\sigma} > 0; \quad \frac{d\alpha^*}{d\sigma} \leq 0 \text{ if } \Phi(c_H, M_S, \sigma) \leq 0. \quad (20)$$

The first inequality confirms the result presented in Lemma 5.b), that cost variance increases expected profits and hence the number of potential entrants. Indeed, expected profits are convex in c_i so that, for the Jensen inequality, $E(\Pi(\tilde{c})) > \Pi(E(\tilde{c}))$ (Spulber, 1995). Thus, cost heterogeneity has a positive effect on M_S : as more potential professionals are attracted into the market, competition intensifies. This implies an increase in proximity among firms, and greater possibility for consumers to substitute one firm's services with another. Selection, and the associated exit of high-cost firms, takes place because these firms perceive they cannot compete adequately with low-cost ones, if the cost gap is sufficiently high and, therefore, they prefer to exit since, in the competition stage, it is unlikely that they will recover the second set-up sunk cost due to their low efficiency.

The selection mechanism induced by cost heterogeneity is formally illustrated in the second inequality. The cut-off point α^* , i.e., the level of cost at which the prior

distribution is truncated, is not effective when the expected profit $\Phi(c_H, M_S, \sigma)$ of a firm with cost c_H is non-negative. That is, $\alpha^* = c_H$ if $\Phi(c_H, M_S, \sigma) > 0$.

However, when selection is active, i.e. $\alpha^* < c_H$ because expected profits are negative at c_H , the cut-off point is negatively related to cost heterogeneity. Therefore, there exists a value of the cost gap ($c_H - c_L$) beyond which a progressive truncation at $\alpha^* < c_H$ of the prior cost distribution occurs: the thrust of Proposition 1 is to show that the cut-off point α^* reduces as the cost gap goes beyond a certain level. The analysis thus bears important policy consequences: the selection mechanism becomes more severe precisely in those situations where it is more needed, that is, when there is a risk that highly inefficient firms may remain and operate in the market.

Proposition 2 *Under Assumption 2,*

$$\frac{d\alpha^*}{dD} < 0; \quad \frac{dM_S}{dD} > 0; \quad \frac{dM_S}{dF_L} < 0; \quad (21)$$

The first inequality mirrors the theoretical predictions in Syverson (2004): the upper bound of the firms' posterior cost distribution decreases in demand density. Part of the explanation of this result is associated with the second inequality: more firms enter the licensing stage in denser markets, attracted by a higher level of expected profit. Therefore, an increase in density plays a qualitatively similar role as cost variance. Because the average distance separating any two firms reduces, their products become closer substitutes and the ensuing intensification of competition makes it less likely for a high-cost firm to retain any positive market share. The posterior cost distribution obtains therefore from a truncation of the prior distribution from above, whose magnitude is larger in denser markets.

It is not easy, however, to derive a clear-cut prediction with regards to the impact of market density on the equilibrium number of firms, N_S . In this case, a direct and an indirect effect are at play:

$$\frac{dN_S}{dD} = \overbrace{\frac{dM_S}{dD}}^{>0} G(\alpha^*) + g(\alpha^*) \overbrace{\frac{d\alpha^*(M_S)}{dM}}^{<0} \overbrace{\frac{dM_S}{dD}}^{>0} M_S \gtrless 0 \quad (22)$$

On the one hand, in a denser market larger profits intensify entry in the first stage, thereby also increasing the number of firms in the production stage (direct effect); on the other, the competitive pressure from a larger number of firms push towards the exclusion of less competitive firms (indirect effect). The overall effect depends on the net balance between these two forces. Simulations obtained assuming a uniform distribution suggest that the first effect dominates: N_S is increasing in D but at a lower rate than M_S .

Finally, the third inequality constitutes a standard result, which has however important implications in professional markets. On the one hand, a higher licensing fee discourage potential entrants, thereby creating a potential restriction to entry and therefore to competition; on the other, the licensing sunk cost plays a crucial role and its elimination would be likely associated with efficiency losses. Indeed, Proposition 1 shows that the licensing stage induces self-selection of most efficient professionals in situations where heterogeneity is large.

Figure 1 further illustrates how the selection mechanism operates as a function of the cost gap (which is a proxy for variance) $c_H - c_L$. For low levels of the cost gap, the optimal number of firms in the licensing stage and in the production stage coincide ($M_S = N_S$ and $\alpha^* = c_H$). That is, there is no selection when heterogeneity is small in magnitude. However, as heterogeneity increases, the selection mechanism

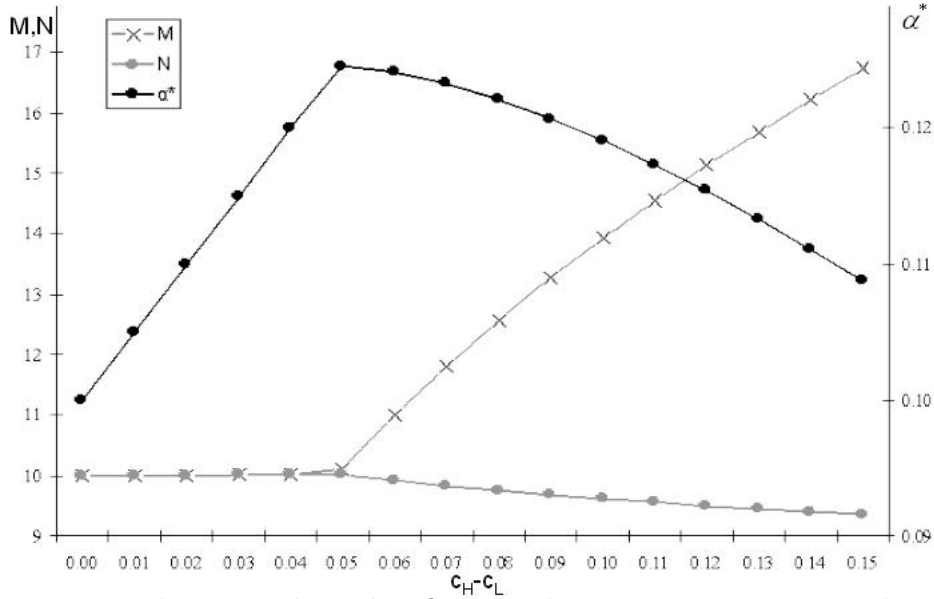


Figure 1: The optimal number firms in the pre-entry stage M and in the production stage N (left scale). Optimal entry threshold (right scale). $F_L = 0.2$, $F_p = 0.8$, $\bar{c} = 0.1$, $c_H - c_L \in [0, 0.14]$, $D = 100$ and $t = 1$.

determines a cut-off point after which $\alpha^* < c_H$ and M_S increases at a fast rate. When the decision to enter is made under uncertainty, an increase in the costs' variance leads to an increase in the expected profit in the licensing stage, even when firms can anticipate that in the subsequent selection stage a number of firms will not continue. In Figure 1, the selection mechanism is responsible for: 1) the decreasing trend in N_S due to α^* being decreasing in cost variance when selection is in place; 2) the exit of the $M_S - N_S$ least efficient firms ($M_S - N_S$ also increases with variance); 3) the increase in the average level of market efficiency, $\bar{c}_e = \frac{\alpha^* + c_L}{2} < \bar{c}$.

To sum up, the foregoing analysis, in addition to being able to replicate some recent results explaining the relatively higher efficiency of firms operating in denser markets, has illustrated how the licensing stage may have beneficial implications for allocative efficiency even in the absence of the traditional rationales generally advo-

cated to justify licensing as a regulatory instrument in liberal professions (Philipsen, 2009). Indeed, our set-up abstracts from any form of market failure due to informational asymmetries, negative externalities or the provision of a public good (Kleiner, 2000). Nonetheless, we obtain that the licensing stage is instrumental in allowing a selection mechanism to operate by driving inefficient firms out of the market. Most importantly, our main prediction is that such a mechanism is more restrictive in markets characterized by higher levels of ex-ante cost heterogeneity.

4.4 Compulsory vs. voluntary apprenticeship

The foregoing analysis suggests that licensing, modeled as a period of compulsory apprenticeship, can be welfare enhancing since it makes professionals learn about their type: they decide whether to incur the entry production cost only if they find out their type is low enough. Therefore, there appears to be a private benefit for professionals from undertaking a period of apprenticeship. In this section, we analyze the circumstances under which licensure can be replaced by a voluntary apprenticeship scheme into which the potential entrants find it in their own interest to enrol. The analysis will also investigate the rationale for not having either forms of apprenticeship.

To derive Figure 2, we consider two different scenarios, one in which the firms pay an apprenticeship cost F_L and learn their type, the other where they do not incur the cost and therefore are unaware of their type when they enter the production phase. In the first scenario, the parameter values are such that selection is always active ($D = 100$, $F_p = 0.8$, $t = 1$, $c_L = 0.05$, $c_H = 0.15$).

The social gain function in Figure 2 is obtained by calculating the difference between the social welfare in the second scenario, i.e., when compulsory licensure or

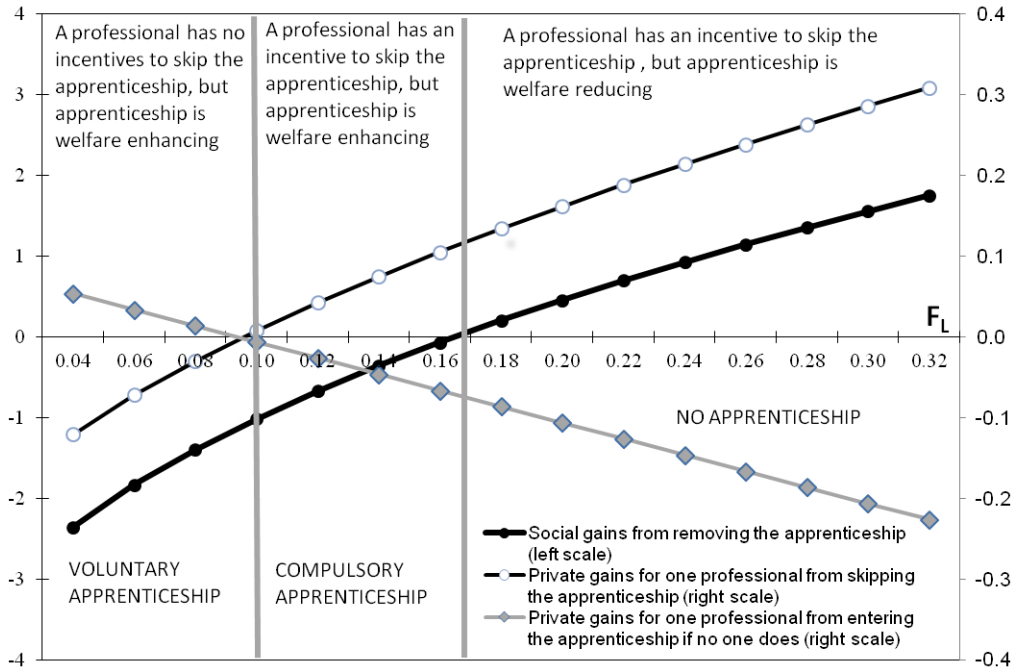


Figure 2: Voluntary vs. compulsory apprenticeship.

voluntary apprenticeship are ruled out, and the social welfare arising from having the potential entrants undertaking a stage where they learn their own type (i.e., the first scenario). For $F_L < 0.17$, not having any form of apprenticeship is welfare reducing. Intuitively, for a given ex-ante cost heterogeneity, when apprenticeship costs are high, the benefits from excluding inefficient firms are outweighed by the excessive burden due to the apprenticeship scheme.

From an individual viewpoint, the private incentive to undertake the apprenticeship varies with its cost. When F_L is below 0.10, the expected profit from skipping the apprenticeship scheme (when everyone else does it) is lower than the one when the apprenticeship is undertaken. This is because the apprenticeship has an option value: buying it allows a high-cost firm the possibility to exit to avoid negative profits, whereas, without it, the same type of firm would not exit and would face tough competition (i.e., would make negative profits). Similarly, when everyone is

skipping the apprenticeship, there is a private benefit from doing it. Therefore, when F_L is low enough, every potential entrant finds it privately profitable to enter an apprenticeship scheme and there is no need for a compulsory licensure. However, for intermediate values of F_L , every individual would find it beneficial to skip the apprenticeship stage, although it is socially desirable. In such a case, there is room for compulsory entry regulation.

From a policy viewpoint, the entry mechanism depends crucially on the value of F_L . In practice, this largely depends on the requirements set in the licensing process. While, on the one hand, it guarantees that the professional has achieved a minimum level of competence, on the other it is necessary to consider the incentives that professional bodies face to inflate the period of apprenticeship and the difficulty of the examination, thereby reducing the number of firms and thus raising the rents for the incumbents (Leland, 1980; Shaked and Sutton, 1981; Shapiro, 1986). In some cases, larger entry costs can reduce the number of potential entrants and can lower the quality of the service (Kleiner, 2000). Using data on US labor market after controlling for gender, age, race and education level, Kleiner and Krueger (2011) and Weeden (2002) find that licensing occupations receive a premium of about 9% and 18%, respectively. In the US dentistry market, Kleiner and Kudrle (2000) find that tougher licensing thresholds do not improve the quality of the service, but raise earnings of practitioners. These situations are consistent with our results if we suppose that entry costs F_L are proportional to the duration of licensure τ_L , and that prospective professionals become fully aware of their type after a period τ_0 . A socially inefficient outcome arises whenever $\tau_L > \tau_0$. This may provide a rationale for the decision taken by the Italian Government led by PM Monti, which decreed to shorten the apprenticeship period of many professions (see D.L. N. 1, 24/01/12).

4.5 Licensing and quality

The results obtained thus far rule out information asymmetries between buyers and sellers (i.e. quality is observable). As suggested in fn. 5, reinterpreting the results in Propositions 1 and 2 in terms of quality θ_n implies that market selection operates to exclude firms with lower quality and firms entering in the production stage are those with higher quality. Because of Proposition 1, the selection mechanism is more effective when there are high ex-ante quality differences among prospective professionals, and so some forms of voluntary or compulsory apprenticeship may be welfare improving. Moreover, a higher degree of ex-ante heterogeneity provides larger private and social incentives to enter the selection phase, therefore increasing the number of participants and the quality of selected ones. Hence, the selection mechanism works to reduce the ex-post heterogeneity especially when quality differences are high.

In the extant literature, licensure is seen as a training period which may enhance professional competencies (Shapiro, 1986). In our set-up, if we further assume that quality increases with the length of the training period, having a licensing phase turns out to be more beneficial than before: in addition to the selection mechanism (which shakes out lower quality workers), licensure increases the average quality of the professionals. The two effects reinforce each other: the selected professionals have higher quality, and consumers receive larger utility. In practice, if the apprenticeship period increases by δ the quality of all trainees, in equilibrium, there is the same number of prospective practitioners M and professionals N ; the quality adjusted prices reduces by δ ; social benefits increase by δD and consequently, there is shift to the right of the critical thresholds depicted in Figure 2.

Different quality perception among consumers may lead to different forms of market organization. For instance, when quality is appreciated by some customers

more than others, there is room for some professionals to take a voluntary decision to invest in additional quality (e.g. a certificate). Iossa and Jullien (2010) analyze the market for lawyers and show that being certified increases the rate of success in litigations. This is because less experienced lower-court judges, who are more concerned about their careers' prospects, tend to support more favorably the theses of certified lawyers than those of uncertified ones. The authors point out that there can be an excessive number of certified lawyers. Shapiro (1986) contrasts the case of certification and licensing when there is imperfect observability of quality. He also finds that under certification, there is excessive investment (signalling) by professionals. In his set-up neither licensing or certification dominates the other.

Our set-up can be extended to discuss the case of imperfect observability of quality when reputation matters. This can be done by assuming, as in Shapiro (1986), that the pricing stage is divided into two different periods, one lasting W and the other $1 - W$. In the first period, lawyers are young and they have not already revealed their quality to consumers (they have no reputation), and consequently their remuneration is uniform and independent from their types. In the second period, thanks to direct experience, word-of-mouth or availability of public information on lawyer performance (Iossa and Jullien, 2010), consumers can distinguish good lawyers from bad ones, and the high quality professionals receive a higher pay. Because, before acquiring a good reputation, the better professionals cannot gain larger revenues as their quality is indistinguishable from that of bad types, the selection mechanism is less effective. Imperfect observability therefore plays against the possibility to encourage a larger participation in the first stage, since in the second stage the difference in returns that accrue to good or bad lawyers is less pronounced than in the case where quality is immediately observable. In turn, this

corresponds to a shift to the left of the critical thresholds in Figure 2.

5 Entry without licensing

For some professions, acquiring a relevant university degree is the only formal condition to start production. For instance, in Belgium and Portugal all individuals gaining a degree in pharmaceutical sciences do not need to satisfy any further educational requirement and are consequently deemed qualified to run an establishment in any part of the country (Paterson *et al.*, 2007; Schaumans and Verboven, 2008). There is, therefore, no licensing stage involved. However, in Belgium and Portugal, as well as in most EU countries, the number of pharmacies allowed to operate within a given geographic area is fixed and depends on the area population.

Quite interestingly, in such countries as Ireland and Germany, where pharmacists have to undertake a full licensing process that involves an apprenticeship period and an exam, geographical entry restrictions are not applied. It appears therefore that licensing and geographical entry restrictions constitute mutually exclusive regulatory mechanisms, the presence of the latter being generally explained by the public interest motive according to which unregulated markets would generate excessive entry and therefore a lower level of social welfare (Philipsen, 2009).

We consider a two-stage entry game with the entry stage followed by the production stage. This is directly obtained from the licensing entry game of the previous Section by setting $F_L = 0$, i.e., no licensing. Before entering, each potential entrant only knows the costs distribution but has no information on the actual realization of its own and its potential opponents' costs, as in Assumption 1. In the first stage, those potential entrants choosing to enter have to pay a set-up cost F , the others remain out of the market and gain zero profit. After investing F and occupying a

random equidistant location, in the production stage price competition takes place among the N entrants as in Section 3, so that locations and costs are common knowledge. We solve the model by backward induction. The price equilibrium in the production stage is therefore given by Lemma 2 under the existence conditions of Lemma 3.

5.1 Free market entry equilibrium

Firms are risk-neutral and have perfect foresight, so that they enter the market if their expected profit is non-negative. Since costs are unknown in the first stage, firm n 's random profit is:

$$\tilde{\Pi}_n = \frac{D}{t} (\tilde{p}_n - \tilde{c}_n) \cdot \left(\frac{1}{2} (\tilde{p}_{n-1} + \tilde{p}_{n+1}) - \tilde{p}_n + k \right) - F. \quad (23)$$

Lemma 10 *The expected profit of firm n is:*

$$E \left[\tilde{\Pi}_n | N \right] = \frac{D}{t} (\sigma^2 W_I(N) + k^2) - F \quad (24)$$

where $W_I(N) = \sum_{i=\bar{l}}^{\bar{l}} w_{\bar{l}-|i|} \left[(w_{\bar{l}} - w_{\bar{l}-1}) - (w_{\bar{l}-|i|} - w_{\bar{l}-|i+1|}) \right] > 0$, with $w_{-1} := w_{\bar{l}+1}$; $\lim_{N \rightarrow \infty} W_I(N) = \bar{W}_I = 1 - \frac{4}{9}\sqrt{3} \simeq 0.23020$.

The values of $W_I(N)$ are reported in Table 1: they monotonically decrease for $N \geq 3$, and rapidly converge toward \bar{W}_I . From Lemma 10, the expected profit $E \left[\tilde{\Pi}_n | N \right]$ decreases in N and F and increases in D , σ and t .¹⁴ As in the licensing game, the expected profits increase with cost heterogeneity. Because the profit function is convex in c_i , in the extreme case where any two adjacent firms' costs are

¹⁴Taking the derivative of $E \left[\tilde{\Pi}_n | N \right]$ with respect to t , we obtain that the expected profit is increasing in t when $\sigma < k/\sqrt{\bar{W}_I} \simeq 2.09k$. In the support $[c_L, c_H]$ the maximal variance arises from a distribution where all the mass is equally concentrated on the extremes: $\sigma = (c_H - c_L)/2$; hence, under (9) the expected profit is always increasing in t for any distribution.

perfectly negatively correlated, i.e., $E[(c_n - \bar{c})c_{n+1}] = -\sigma^2$ and N is even, it can be shown that the expected profit (24) reduces to $E[\Pi] = \frac{D}{t} (k^2 + \frac{4}{9}\sigma^2) - F$.¹⁵ That is, firms prefer to “gamble” even if there is a fifty-fifty chance of drawing a high cost. Furthermore, when there is perfect positive pairwise correlation among costs, $\sigma^2 = 0$ and the profits are the same as the case of no uncertainty and no heterogeneity.

Proposition 3 *From Lemma 10, the equilibrium number of firms in the market, N_M , is implicitly given by the following equation: $t/N_M^2 + \sigma^2 W_I/t(N_M) = F/D$. For N large, the number of firms is approximated by:*

$$N_M \simeq \sqrt{\frac{t}{F/D - \bar{W}_I \sigma^2/t}} \quad (25)$$

It is noteworthy that N_M converges to the value in Salop (1979) when firms’ costs are identical, i.e., $N_M = \sqrt{tD/F}$ when $\sigma^2 = 0$. More importantly, the game with cost heterogeneity yields a higher number of firms in equilibrium, relative to that in the traditional case. However, the two cases are not perfectly comparable given the different informational structure they assume. Therefore, to gauge the extent by which the free-entry outcome in the last Proposition is sub-optimal, we need to derive the Pareto optimal outcome under cost heterogeneity.

5.2 Pareto Efficient Entry

In the spirit of Salop (1979), we now compare the free-entry market equilibrium in Proposition 3 with the first-best optimum that would be chosen by a social planner. It is worth stressing that relative to the welfare analysis in Salop (1979), introducing cost heterogeneity brings about an extra distortion associated with the total production cost incurred in the economy. This is captured by the difference between

¹⁵Perfect negative correlation is obtained when the 1-step neighbors of a c_L -type firm are c_H -types, and vice versa. Calculations are available on request.

the market equilibrium prices in Lemma 1 and the first-best prices in Lemma 4. Indeed, the market equilibrium pricing rule induces an allocative inefficiency such that consumers who should patronize a low-cost firm end up buying from a higher cost firm, thereby increasing the total and average costs incurred in the industry. It follows therefore that the first-best optimal number of firms chosen by a benevolent social planner is such that it minimizes the expected sum of transport, production and fixed costs (see Appendix, proof of Lemma 4).

Proposition 4 *a) The first-best optimal number of firms is:*

$$N_F = \frac{1}{2} \sqrt{\frac{t}{F/D - 0.5\sigma^2/t}}. \quad (26)$$

b) There is less excessive entry under cost heterogeneity, because $\bar{W}_I < 0.5$, so that $N_M/N_F < 2$ when $\sigma > 0$.

Part b) of the Proposition refers to the fact that the ratio $N_M/N_F = 2$ when firms are identical. Its corollary is that the greater the degree of heterogeneity, the less excessive market entry is, with the proviso that condition (9) determines a lower bound for N_M/N_F . The welfare-enhancing effect of cost heterogeneity could be seen as unexpected, given the extra allocative distortion that the social planner has to manage in this case. From Proposition 3 and eq. (39) in the Appendix, it appears clear that an increase in variance raises a firm's expected profit less that it reduces the industry's expected total cost (in particular, the average variable cost). Therefore, the regulator is willing to allow more firms in the market, so as to shift production towards the more efficient plants, since the gain from doing so offsets the increase in expected transport cost. In other words, excessive entry is reduced because the first-best number of firms, N_F , increases with cost heterogeneity faster than the free-entry one, N_M .

The overall implications of the foregoing analysis on the regulation of liberal professions point in opposite directions. On the one hand, when we consider the more realistic case of heterogeneous firms, the free market outcome tends to be more closely aligned to the first-best equilibrium, thereby suggesting there is less need to restrict the number of firms in professional markets with more heterogeneity. On the other, in the model the distance between the two outcomes remains sufficiently large. Indeed, numerical simulations show that N_M/N_F is indeed monotonically decreasing in σ and reaches a minimum value around 1.78, when $(c_H - c_L)$ in Lemma 3 takes its maximum possible value (with $F = t = 1$, $D = 100$, $\bar{c} = 0.1$). Therefore, the motivation for entry restrictions on the maximum number of firms advocated in the public interest view appears to continue to be theoretically relevant also in our setting.

5.3 Discussion

To further highlight the welfare-enhancing role of cost heterogeneity, consider the following three different pricing schemes:

- *uniform* (e.g., fixed or minimum prices set by professional bodies correspond to this case; equivalently, the regulator may not differentiate between different types of suppliers, as in the case of the Belgian pharmacies described in Schaumans and Verboven (2008), where urban and rural pharmacies are compensated uniformly);
- *market* (prices are set according to Lemma 2);
- *first-best* (the marginal cost pricing in eq. (10)).

Each of these is characterized by a different expected average social cost, $E[C(N)]$, given by the sum of the expected average transport, production and fixed costs. Analytically, and without loss of generality, we have

$$E[C(N)] = \frac{4\bar{c} + k}{4} + \frac{NF}{D} - \frac{\sigma^2}{2k}\chi, \quad (27)$$

where $\chi = 0$ under uniform pricing, $\chi \simeq 0.65$ under market pricing, and $\chi = 1$ under first-best pricing - see (39) in the Appendix for the first-best case. Note that the cost associated to uniform pricing is not affected by σ , and that the three costs coincide when $\sigma = 0$.¹⁶

Figure 3 plots $E[C(N)]$ under the three pricing scenarios as a function of the number of firms. It depicts the relative position of a number of possible outcomes: F is the first-best solution; S is the second best solution (the social planner's choice of entrants when they set market prices); M is the free-market outcome; and U is the social planner's choice of entrants under uniform pricing. Assuming that the first-best solution F cannot be implemented (e.g., due to lack of information about cost heterogeneity), policy intervention should be determined on the basis of the optimal mix between the possible price setting alternatives and the related entry restrictions.

Figure 3 shows that uniform price regulation is dominated by the free-market pricing outcome, and that it performs very poorly when heterogeneity is taken into account.¹⁷ With regards to entry, the free market equilibrium in M , defined in Proposition 3, tends to be less inefficient than U , suggesting that some form of entry regulation may be needed. Indeed, the combination of free market prices with entry restrictions on the number of firms (point S) constitutes a form of second-best, since it entails a level of $E[C(N)]$ only slightly above that of F , but is relatively

¹⁶They differ due to the impact that each pricing mechanism has on market shares, which in turn affects the expected transport and production costs, as the proof of Proposition 4 in the Appendix highlights.

¹⁷In a study of the U.S. real estate market, where commissions rates are fixed at the 6% level and barriers to entry are very low, Hsieh and Moretti (2003) report evidence indicating that fixed rates lead to highly wasteful entry.

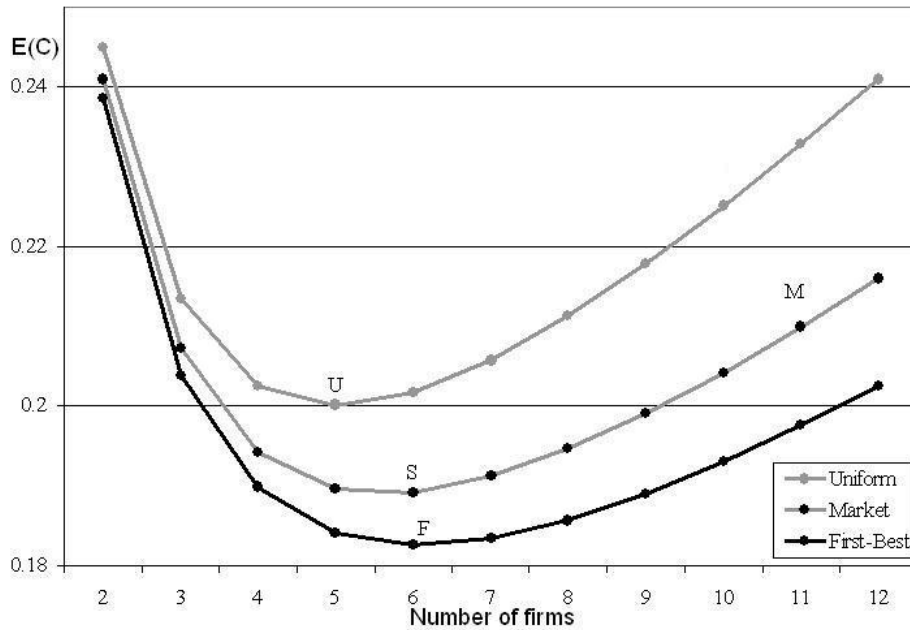


Figure 3: Expected average total cost under different pricing scenarios. F is the social optimum; M is the free-market outcome; S is the second best solution (the social planner's choice of entrants when they set market prices). Under uniform pricing, the planner's choice is in U .

easier and less costly to implement than the first-best outcome since it does not need information on firms' costs.

To conclude, in the presence of sufficiently large cost heterogeneity, welfare improvements can be achieved by replacing uniform pricing with a policy combining both entry restrictions and a price mechanism that more closely reflects firms' underlying costs. Our analysis therefore lends theoretical support and is largely consistent with some policy recommendations found in the empirical literature. In their counterfactual analysis on the Belgian pharmacies, Schaumans and Verboven (2008) consider two policy measures and find that if the Belgian authorities: 1) relaxed the restriction on entry by allowing twice as many outlets per given area population

size, and at the same time 2) they cut the regulated prices non-uniformly by favoring the pharmacies in smaller markets (which are likely to have higher costs), then the number of pharmacies and the related geographical coverage would remain very similar, if not slightly higher, than the pre-intervention period. Therefore, the combined measures implicitly describe a regulatory intervention that would move the Belgian pharmacies market from point U (i.e., fewer firms and uniform prices) towards points F and S (slight increase in firms' number and prices reflecting underlying cost conditions) in Figure 3.¹⁸ Based on our analysis, such a move would be welfare-enhancing.¹⁹

6 Conclusions

Professional services are a key sector in modern economies. Such a sector is characterized, however, by high levels of regulation, resulting from a mix of State regulation, self-regulation and custom and practice. In European Commission (2004), all EU member States and competition authority were invited to assess to what extent existing professional regulations and rules truly serve the public interest and can be objectively justified. To this purpose, the Commission suggested the application of a “proportionality test” according to which each regulatory measure should be explained by an explicitly stated objective detailing how the chosen measure represents the least restrictive mechanism of competition to effectively attain the stated objective. The Commission’s recommendation implicitly calls for further analysis

¹⁸Because market density is fixed at D in both cases, the number of firms per capita is therefore larger under S .

¹⁹This conclusion is based on efficiency grounds - see eq. (27). However, if the regulator is more concerned about equity (e.g., when public service obligation is mandated), then an analysis based on the difference between the highest and lowest price ($\Delta p = (c_H - c_L)\chi$) paid by consumers would indicate that uniform pricing ($\chi = 0$) should be clearly the most preferred solution. The market pricing ($\chi = 0.65$) would continue to be the intermediate one, and the first-best pricing ($\chi = 1$) scheme would be the one with the largest price dispersion.

on the economic properties associated with each regulatory practice.

In this paper we have investigated the economic implications of regulations pertaining both to entry requirements in professional markets and to price setting. Licensing appears to be particularly important in a market with heterogeneous firms, since it is capable to induce a mechanism such that only the most efficient firms in the market become operative. Such a selection mechanism intensifies its effect as the degree of heterogeneity increases. Our findings therefore point at a generally beneficial effect of licensing, and may explain why a formal licensing law capturing many features present in many Western World economies, has been recently introduced in China for the legal professions (Philipsen, 2009).

For relatively simpler professional tasks, it is often argued that licensing may constitute an excessively restrictive rule, which, if relaxed, might benefit consumers. For example, prices dropped significantly in the United Kingdom after the lawyers' reserved rights to provide conveyancing services were removed (European Commission, 2004). We investigate the extreme case of no licensing, to find that when the first-best outcome cannot be implemented, the second-best policies should be aimed at combining the greatest possible flexibility in price setting, so that prices reflect the underlying efficiency of each firm, with some limits on the number of practitioners in the market. From a practical viewpoint, our analysis suggests, in line with the evidence available in the literature, that a reform removing uniform pricing in favor of a liberalized price setting system may also likely necessitate a re-adjustment of the restrictive rules on entry based on a geographic market's demographics.

Appendix

Proof of Lemma 1. See Alderighi and Piga (2012).

Proof of Lemma 2. From Lemma 1, firm n equilibrium strategy is:

$$p_n = w_{\bar{l}}c_n + w_{\bar{l}-1}(c_{n-1} + c_{n+1}) + w_{\bar{l}-2}(c_{n-2} + c_{n+2}) + \dots + w_0c_{n+\bar{l}} + k. \quad (28)$$

By the same token, the price charged by the adjacent firms located at $n - 1$ and $n + 1$, respectively, are:

$$\begin{aligned} p_{n-1} &= w_{\bar{l}}c_{n-1} + w_{\bar{l}-1}(c_{n-2} + c_n) + w_{\bar{l}-2}(c_{n-3} + c_{n+1}) + \dots + w_0c_{n+\bar{l}-1} + k \\ p_{n+1} &= w_{\bar{l}}c_{n+1} + w_{\bar{l}-1}(c_n + c_{n+2}) + w_{\bar{l}-2}(c_{n-1} + c_{n+3}) + \dots + w_0c_{n+\bar{l}} + k. \end{aligned} \quad (29)$$

Now substituting (28) and both equations in (29) in (2), and collecting the similar terms, we obtain:

$$\begin{aligned} &c_n (4w_{\bar{l}} - 2 - 2w_{\bar{l}-1}) + (c_{n-1} + c_{n+1}) (4w_{\bar{l}-1} - w_{\bar{l}-2} - w_{\bar{l}}) + \\ &\quad + (c_{n-2} + c_{n+2}) (4w_{\bar{l}-1} - w_{\bar{l}} - w_{\bar{l}-2}) + \dots + \\ &+ (c_{n+\bar{l}} + c_{n+\bar{l}-1}) (4w_1 - w_0 - w_2) + c_{n+\bar{l}} (4w_0 - 2w_1) + (2k - 2t/N) = 0. \end{aligned} \quad (30)$$

In order for this equality to hold for every possible configuration of c_n , $n \in \mathbb{L}$, all the expressions in round brackets must be equal to zero, thereby proving (4)-(6) and (8). Finally, (7) emerges by replacing $c_n = c$ for $n \in \mathbb{L}$ in (30).

Proof of Lemma 3. See Alderighi and Piga (2012).

Proof of Lemma 4. We focus on the generic arc of length $\frac{1}{N}$ defined by the locations of firms $n - 1$ and n . Because demand is inelastic, the first-best allocative solution is reached when the sum of production and transport costs is minimized. Let d_n and $d_{n-1} = (\frac{1}{N} - d_n)$ be the distance of the pivotal consumer from two adjacent

firms. The average distance between a generic consumer on the arc and the seller she patronizes is: $\bar{d} = N \left(\int_0^{d_{n-1}} x dx + \int_0^{d_n} x dx \right) = \frac{N}{2} (d_{n-1}^2 + d_n^2)$. Hence, the average transport costs among all consumers on the arc is:

$$TC = \bar{d}t = \frac{N}{2} (d_{n-1}^2 + d_n^2) t \quad (31)$$

The proportion of consumers patronizing firm $n - 1$ and firm n are, respectively, Nd_{n-1} and Nd_n . The average (variable) cost of producing goods for a consumer is:

$$VC = Nd_{n-1}c_{n-1} + Nd_n c_n. \quad (32)$$

The social planner problem can be expressed as $\min_{d_n \geq 0} (TC + VC)$, yielding: $d_n^* = \frac{1}{2N} + \left(\frac{c_{n-1} - c_n}{2t} \right)$. Its implementation can be obtained by using the prices in the Lemma as policy instruments.

Proof of Lemma 5. Multiply the terms within the square brackets in (11) to obtain: $E[\tilde{p}_n \tilde{p}_{n-1}] - E[\tilde{p}_n^2] + E[\tilde{p}_n]k - c_n E[\tilde{p}_{n-1}] + c_n E[\tilde{p}_n] - c_n k$. Each term is derived separately. (See also: Alderighi and Piga, 2012b)

Part 5.I Under Assumption 1: $E_{G_e}[\tilde{p}_n^2] = ((w_{\bar{l}}c_n + (1 - w_{\bar{l}})\bar{c}_e) + k)^2 + \sigma_e^2 W_{L,0}(N)$, where $W_{L,0}(N) = \sum_{i=\mathbb{L}_0} w_{\bar{l}-|i|}^2$.

Derivation of Part 5.I. Using (3), $E[c_n] = c_n$, $E[\tilde{c}_{n+i}] = \bar{c}_e$, $E[c_n^2] = c_n^2$, $E[c_n \tilde{c}_{n+i}] = c_n \bar{c}_e$, $E[\tilde{c}_{n+i}^2] = (\sigma^2 + \bar{c}_e^2)$, with $i \in \mathbb{L}_0$ and $E[\tilde{c}_{n+i} \tilde{c}_{n+j}] = \bar{c}_e^2$, $i, j \in \mathbb{L}_0$, $i \neq j$: $E[\tilde{p}_n^2] = E \left[\left(w_{\bar{l}} \cdot c_n + \sum_{i=\mathbb{L}_0} w_{\bar{l}-|i|} \cdot \tilde{c}_{n+i} \right)^2 \right] + 2(w_{\bar{l}} \cdot c_n + (1 - w_{\bar{l}})\bar{c}_e)k + k^2$. The expected value in the squared brackets can be written as: $w_{\bar{l}}^2 c_n^2 + 2w_{\bar{l}}(1 - w_{\bar{l}})\bar{c}_e c_n + \sum_{i=\mathbb{L}_0} w_{\bar{l}-|i|}^2 \sigma_e^2 + \sum_{i,j=\mathbb{L}_0} w_{\bar{l}-|j|} w_{\bar{l}-|i|} \bar{c}_e^2$. Setting $W_{L,0}(N) = \sum_{i=\mathbb{L}_0} w_{\bar{l}-|i|}^2$ and noting that $\sum_{i,j=\mathbb{L}_0} w_{\bar{l}-|j|} w_{\bar{l}-|i|} = (1 - w_{\bar{l}})^2$, we obtain $(w_{\bar{l}}c_n + (1 - w_{\bar{l}})\bar{c}_e)^2 + \sigma_e^2 W_{L,0}(N)$. Summing up previous results completes the calculation. \diamond

Part 5.II Under Assumption 1: $E_{G_e} [\tilde{p}_n \tilde{p}_{n-1}] = (c_n w_{\bar{l}-1} + (1 - w_{\bar{l}-1}) c_e + k) (c_n w_{\bar{l}} + (1 - w_{\bar{l}}) c_e + k) + W_{L,1}(N) \sigma^2$, where $W_{L,1}(N) = \sum_{i=\mathbb{L}_0} w_{\bar{l}-|i|} w_{\bar{l}-|i+1|}$.

Derivation of Part 5.II. Similar to that of Part 5.I. \diamond

a) To derive (13), substitute $E[\tilde{p}_n] = w_{\bar{l}} c_n + (1 - w_{\bar{l}}) \bar{c}_e + k$, $E[\tilde{p}_{n-1}] = w_{\bar{l}-1} c_n + (1 - w_{\bar{l}-1}) \bar{c}_e + \hat{k}$ and the results in Parts 5.I and 5.II into (11). After some simplifications we obtain the thesis. To compute \bar{W}_L and w_d , use the numerical algorithm for the weights in fn. 9. **b)** For (13) to be decreasing in c_n , it sufficient that $w_d(\bar{c}_e - c_n) + k \geq 0$, which holds since $\sqrt{3}k > c_H - c_L \geq |\bar{c}_e - c_n|$ from Lemma 3.

Proof of Lemma 6. It comes from the differentiation of (15).

Proof of Lemma 7. Let $\phi(\alpha, c_n, M, \sigma) = E_{G_\alpha} [\tilde{\Pi}_n | c_n, M]$ and ϕ_x be the derivative with respect to x evaluated at $(\alpha^*(M), c_n, M, \sigma)$, i.e. $\phi_x = \frac{\partial}{\partial x} \phi(\alpha^*(M), c_n, M, \sigma)$. Set, henceforth, without loss of generality $\frac{D}{t} = 1$. From (13), $\phi = A^2 + W_L \sigma_e^2 - F_p$, where $A = w_d(\bar{c}_e - c_n) + k$. We must prove that $\phi_\alpha = 2AA_\alpha + W_L d\sigma_e^2/d\alpha < 0$. Under Assumption 2, after taking the derivatives and dividing the result by $2(\alpha - c_L)$, we obtain: $\left(\frac{\bar{c}_e - c_n}{\alpha - c_L} w_d + \frac{k}{\alpha - c_L}\right) \left(\frac{1}{2} w_d - \frac{k}{\alpha - c_L}\right) + \frac{W_L}{12} < 0$. By noting from Lemma 3 that $k > (c_H - c_L)/\sqrt{3}$, after substituting $w_d \simeq 0.42$ and $W_L \simeq 0.051$, it emerges that the first parenthesis is positive and the second one is negative. Simple computations show that the first term is in magnitude greater than the last one.

Proof of Lemma 8. First inequality. Differentiating (16) with respect to α^* and M , we obtain: $\frac{d\alpha^*}{dM} = -\Phi_M/\Phi_\alpha$ which is negative since $\Phi_M = \phi_M < 0$ from Lemma 6, and $\Phi_\alpha = \phi_\alpha + \phi_{c_n} < 0$ from Lemmata 6 and 7. The second inequality can be obtained using similar arguments.

Proof of Lemma 9. Define:

$$\Psi(\alpha, M, \sigma) = E_{G_\alpha} [\tilde{\Pi}_n |, M] = \int_{c_L}^{\alpha} ((w_d(\bar{c}_e - c_n) + k)^2 + W_L \bar{\sigma}_e^2 - F_p) g d c_n - F_L. \quad (33)$$

Moreover, from (33), after computing the integral,

$$\Psi(\alpha, M, \sigma) = \left(k^2 + \frac{1}{12} (w_d^2 + W_L) (\alpha - c_L)^2 - F_p \right) \frac{\alpha - c_L}{c_H - c_L} - F_L. \quad (34)$$

We have to show that $\frac{d\Psi}{dM} < 0$. When $\Phi(c_H, M, \sigma) \geq 0$ then: $\Psi = \left(\frac{t}{M}\right)^2 + (W_L + w_d^2) \sigma^2 - F_p - F_L$, which is clearly decreasing in M . When $\Phi(c_H, M, \sigma) < 0$, totally differentiating previous expression we obtain:

$$\frac{d\Psi}{dM} = \frac{d\alpha^*}{dM} \cdot \phi(\alpha^*, \alpha^*, M, \sigma) g(\alpha^*) + \int_{c_L}^{\alpha^*} \frac{d}{dM} \phi(\alpha^*, c_n, M, \sigma) g(c_n) dc_n, \quad (35)$$

where the first term is null because $\phi(\alpha^*, \alpha^*, M, \sigma) = \Phi(\alpha^*, M, \sigma) = 0$, and the last term is negative from Lemma 6.

Proof of Proposition 1. Remember from (13) and Assumption 2 that:

$$\Phi(\alpha, M, \sigma) = \left[k - \frac{1}{2} w_d (\alpha - c_L) \right]^2 + \frac{1}{12} W_L (\alpha - c_L)^2 - F_p. \quad (36)$$

In order to determine the sign of $\frac{dM}{d\sigma}$ and $\frac{d\alpha}{d\sigma}$ when $\Phi(c_H, M, \sigma) < 0$, consider the two equilibrium conditions: $\Phi(\alpha, M, \sigma) = 0$ and $\Psi(\alpha, M, \sigma) = 0$. The former determines the equilibrium of the production stage and the latter of the entry stage.

By totally differentiating these expressions, we obtain: $\Phi_\alpha d\alpha + \Phi_M dM + \Phi_\sigma d\sigma = 0$ and $\Psi_\alpha d\alpha + \Psi_M dM + \Psi_\sigma d\sigma = 0$. Thus:

$$\frac{dM}{d\sigma} = -\frac{\Psi_\alpha \Phi_\sigma - \Phi_\alpha \Psi_\sigma}{\Phi_M \Psi_\alpha - \Psi_M \Phi_\alpha}, \quad \frac{d\alpha}{d\sigma} = -\frac{\Phi_M \Psi_\sigma - \Psi_M \Phi_\sigma}{\Phi_M \Psi_\alpha - \Psi_M \Phi_\alpha}. \quad (37)$$

Choosing: $X = (\alpha - c_L)$ and $Y = (c_H - c_L)$, noting that $\frac{d}{d\sigma}(-c_L) = \frac{d}{d\sigma}(c_H) = \sqrt{3}$, we have the following derivatives:

$$\begin{aligned} \Phi_M &= -2\frac{1}{M} \left(k - \frac{w_d}{2} X \right) k \\ \Phi_\alpha &= -2\frac{1}{X} \left(k - \frac{w_d}{2} X \right) \left(k + \frac{w_d}{2} X \right) + \frac{1}{6} W_L X \end{aligned}$$

$$\Phi_\sigma = \sqrt{3} \left(2\frac{1}{X} \left(k - \frac{w_d}{2} X \right) \left(\frac{2X-Y}{Y} k - \frac{w_d}{2} X \right) + \frac{1}{6} W_L X \right)$$

$$\Psi_M = -2\frac{1}{M} \frac{X}{Y} k^2$$

$$\Psi_\alpha = - \left(k^2 - \frac{1}{4} (w_d^2 + W_L) X^2 + F_p \right) \frac{1}{Y}$$

$$\Psi_\sigma = \sqrt{3} \left(\frac{1}{12} (w_d^2 + W_L) (3Y - 2X) X^2 + (2X - Y) k^2 + (2X - Y) F_p \right) \frac{1}{Y^2}$$

Using $\Phi = 0$ to replace $F_p = \left(\left(k - \frac{1}{2} w_d X \right)^2 + \frac{1}{12} W_L X^2 \right)$ in previous expressions, after rearranging the terms and required simplifications, we have: $\Phi_M \Psi_\alpha - \Psi_M \Phi_\alpha = -\frac{w_d}{6} (24k^2 - 12w_d X k - W_L X^2) X k / (MY)$. Substituting $w_d \simeq 0.42$ and $W_L \simeq 0.051$, and noting that $\sqrt{3}k > Y \geq X$ from Lemma 3, it emerges that the previous expression is negative. Thus, the sign of $\frac{dM}{d\sigma}$ and $\frac{d\alpha}{d\sigma}$ is given by the sign of the numerators of the expressions in (37).

Analogously: $\Psi_\alpha \Phi_\sigma - \Phi_\alpha \Psi_\sigma = \frac{w_d}{18} (72k^3 - 60w_d X k^2 + 18w_d^2 X^2 k - 3X^3 w_d^3 - W_L w_d X^3) X / Y^2$, that is positive from similar reasonings, meaning $\frac{dM}{d\sigma} > 0$. Finally: $\Phi_M \Psi_\sigma - \Psi_M \Phi_\sigma = -\frac{\sqrt{3}w_d}{6} (24(Y - X) k^2 + 4w_d (4X - 3Y) X k - 2w_d^2 X^3 - W_L X^2 Y) X k / (MY^2)$, which is negative from previous arguments, meaning $\frac{d\alpha}{d\sigma} < 0$.

If $\Phi(c_H, M, \sigma) > 0$, then $\alpha^* = c_H$. Therefore: $\Psi = k^2 + (W_L + w_d^2) \sigma^2 - F_p - F_L$ which is increasing in M and $\frac{d\alpha^*}{d\sigma} = \frac{c_H}{d\sigma} > 0$.

Proof of Proposition 2. The second and third inequalities are obtained by totally differentiating (19). The first inequality comes from the observation that dividing both sides of (21) the left-hand side is independent of D .

Proof of Lemma 10. Profit (23) can be decomposed as:

$$\begin{aligned} E \left[\tilde{\Pi}_n \right] &= \frac{D}{2t} \left\{ E \left[\tilde{p}_n \tilde{p}_{n-1} \right] - 2E \left[\tilde{p}_n^2 \right] + 2E \left[\tilde{c}_n \tilde{p}_n \right] - E \left[\tilde{c}_n \tilde{p}_{n-1} \right] - E \left[\tilde{c}_n \tilde{p}_{n+1} \right] \right\} + \\ &+ \frac{D}{2t} \left\{ E \left[c_n \tilde{p}_{n+1} \right] \right\} + \frac{Dt}{N^2} - F. \end{aligned} \quad (38)$$

Each term is derived separately. See also: Alderighi and Piga (2012b).

Part 10.I Under Assumption 1: $E[(\tilde{p}_n)^2] = \sigma^2 W_{I,0}(N) + \bar{c}^2 + 2\bar{c}k + k^2$ where $W_{I,0}(N) = \sum_{i=\underline{l}}^{\bar{l}} w_{\bar{l}-|i|}^2 > 0$.

Derivation of Part 10.I. From (3): $E[(\tilde{p}_n)^2] = E\left[\left(\sum_{i=\underline{l}}^{\bar{l}} w_{\bar{l}-|i|} \cdot \tilde{c}_{n+i}\right)^2\right] + 2E\left[\left(\sum_{i=\underline{l}}^{\bar{l}} w_{\bar{l}-|i|} \cdot \tilde{c}_{n+i}\right)\right]k + k^2$. From Assumption 1, we know that $E[\tilde{c}_n^2] = (\sigma^2 + \bar{c}^2)$, and, since \tilde{c}_n and \tilde{c}_{n+i} are uncorrelated: $E[\tilde{c}_n \tilde{c}_{n+i}] = \bar{c}^2$, with $i \in \mathbb{L}_0$. After some manipulations we obtain the thesis. \diamond

Part 10.II Under Assumption 1: $E[\tilde{p}_n \tilde{p}_{n-1}] = \sigma^2 W_{I,1}(N) + \bar{c}^2 + 2\bar{c}k + k^2$ where $W_{I,1}(N) = \sum_{i=\underline{l}}^{\bar{l}} w_{\bar{l}-|i|} w_{\bar{l}-|i+1|} > 0$.

Derivation of Part 10.II. Similar to that of Part 10.I. \diamond

Part 10.III Under Assumption 1: **a)** $E[\tilde{c}_n \tilde{p}_n] = \sigma^2 w_{\bar{l}} + \bar{c}^2 + \bar{c}k$; and **b)** $E[\tilde{c}_n \tilde{p}_{n-1}] = \sigma^2 w_{\bar{l}-1} + \bar{c}^2 + \bar{c}k$.

Derivation of Part 10.III. **a)** $E[\tilde{c}_n \tilde{p}_n] = E\left[\tilde{c}_n \left(\sum_{i=\underline{l}}^{\bar{l}} w_{\bar{l}-|i|} \cdot \tilde{c}_{n+i} + k\right)\right]$. The thesis follows from the same reasonings of Parts 10.I and 10.II by noting that the weight associated to $E[\tilde{c}_n^2]$ is $w_{\bar{l}}$. **b).** Similar to a). The weight of $E[\tilde{c}_n^2]$ is $w_{\bar{l}-1}$. \diamond

Note that $E[\tilde{p}_n \tilde{p}_{n+1}] = E[\tilde{p}_n \tilde{p}_{n-1}]$ and $E[c_n \tilde{p}_{n+1}] = E[c_n \tilde{p}_{n-1}]$. After substituting the results from Parts 10.I, 10.II, and 10.III into (38), most of the terms cancel out, finally yielding (23). To compute \bar{W}_I , use Lemma 1 and the numerical algorithm described in fn 9.

Proof of Proposition 3. It derives directly from Lemma 10.

Proof of Proposition 4. Using (31), (32), Assumption 1 and noting that fixed costs per consumer are $FC = NF/D$, the expected average social cost is:

$$E[C] = E[TC + VC + FC] = \left(\frac{t}{4N} + \frac{1}{2} \frac{N}{t} \sigma^2\right) + \left(\bar{c} - \frac{N}{t} \sigma^2\right) + \frac{NF}{D} \quad (39)$$

References

- [1] Alderighi, M. and Piga, C. A. 2010. “On Cost Restrictions in Spatial Competition Models with Heterogeneous Firms.” *Economics Letters*, 108(1): 40–42.
- [2] Alderighi, M. and Piga, C. A. 2012. “Localized Competition, Heterogeneous Firms and Vertical Relations.” *Journal of Industrial Economics*, 60(1): 46–74.
- [3] Alderighi, Marco and Piga, Claudio A. 2012b. “Selection, Heterogeneity and Entry in Professional Markets ” (December, 28 2012). Available at SSRN: <http://ssrn.com/abstract=2194391>
- [4] Becker, G. S. 1983. “A Theory of Competition among Pressure Groups for Political Influence ”. *Quarterly Journal of Economics*, 98: 371–400.
- [5] Deneckere, R. and Rothschild, M. 1992. “Monopolistic Competition and Preference Diversity ”. *Review of Economic Studies*, 59: 361–373.
- [6] Eaton, C. B. and Lipsey, R. G. 1978. “Freedom of Entry and the Existence of Pure Profit”. *Economic Journal*, 88(351): 455–69.
- [7] European Commission. 2004. “Report on Competition in Professional Services”. *Communication from the Commission*, February, COM. 83.
- [8] Holmstrom, B. 1999. “Managerial Incentive Problems: A Dynamic Perspective”. *Review of Economic Studies*, 66: 169–82.
- [9] Hsieh, C.-T. and Moretti, E. 2003. “Can Free Entry be Inefficient? Fixed Commissions and Social Waste in the Real Estate Industry”. *Journal of Political Economy*, 111: 1076–1122.
- [10] Iossa, E. and Jullien, B.. 2010. “The Market for Lawyers and the Quality of Legal Services”. *CEIS Working Paper* No. 170 (July). Available at <http://dx.doi.org/10.2139/ssrn.1653919>
- [11] Kleiner, M. M. 2000. “Occupational Licensing”. *The Journal of Economic Perspectives*, 14(4):189–202

- [12] Kleiner, M. M. and Krueger, A. B. 2011. “Analyzing the extent and Influence of Occupational Licensing on the Labour Market”. *IZA Discussion Paper Series*, N. 5505, February.
- [13] Kleiner, M. M., and Kudrle, R., T. 2000. “Does Regulation Affect Economic Outcomes? The Case of Dentistry”. *Journal of Law and Economics*. 43(2):547-582.
- [14] Leland, H. E. 1979. “Quacks, Lemons, and Licensing: A Theory of Minimum Quality Standards”. *Journal of Political Economy*, 87: 1328–1346.
- [15] Levin, D. and Peck, J. 2003. “To Grab for the Market or to Bide One’s Time: A Dynamic Model of Entry”. *Rand Journal of Economics*, 34(3): 536–556.
- [16] Pashigian, P. B. 1977. “The market for lawyers: The determinants of the demand for and supply of lawyers ”. *Journal of Law and Economics*, 20: 53–85.
- [17] Paterson, I., Fink, M., and Ogus, A. 2007. “Economic Impact of regulation in the field of liberal professions in different member states ”. *European Network of Economic Policy Research Institutes*, WP No. 52. February.
- [18] Peltzman, S. 1976. “Toward a more General Theory of Regulation ”. *Journal of Law and Economics*, 19: 211–240.
- [19] Philipsen, N. F. 2010. “Regulation of Liberal Professions and Competition Policy: Developments in the EU and China.” *Journal of Competition Law and Economics*, 6: 203–231.
- [20] Posner R. A. 1974. “Theories of Economic Regulation ”. *Bell Journal of Economics*, 5: 335–358.
- [21] Rothschild, R. 1982, “Competitive Behaviour in Chain-Linked Markets”. *Journal of Industrial Economics*, 33(1-2): 57–67.
- [22] Salop, S. C. 1979. “Monopolistic Competition with Outside Goods.” *Bell Journal of Economics*, 10(1): 141–156.

- [23] Schaumans, C. and Verboven, F. 2008. “Entry and Regulation - Evidence from the Health Care Professions.”, *Rand Journal of Economics*, 39(4): 949–972.
- [24] Spulber, D. F. 1995. “Bertrand Competition when Rivals’ Costs are Unknown”. *Journal of Industrial Economics*, XLIII(1): 1-11.
- [25] Shapiro, C. 1986. “Investment, Moral Hazard and Occupational Licensing.” *Review of Economic Studies*, 53: 843–862.
- [26] Stigler, G. J. 1971. “The Theory of Economic Regulation ”. *Bell Journal of Economics*, 2:3–21.
- [27] Syverson, C. 2004. “Market Structure and Productivity: A Concrete Example”. *Journal of Political Economy*, 112(6): 1181–1222.
- [28] Tirole, J. 1988. “*The Theory of Industrial Organization*”. MIT Press: Cambridge, Mass., USA.
- [29] Vickers, J. 2001. “Competition in Professions. A Report by the Director General”. *Office of Fair Trading*, OFT328, March.
- [30] Vogel, J. 2008. “Spatial Competition with Heterogeneous Firms”. *Journal of Political Economy*, 116(3): 423–466.
- [31] Weeden, K. A. 2002. “Why Do Some Occupations Pay More Than Others?”. *American Journal of Sociology*. 108(1): 55-101.