Front. Econ. China 2017, 12(3): 371–399
DOI 10.3868/s060-006-017-0016-1

RESEARCH ARTICLE

Nicolas Boccard, Patrick Legros

Audit Competition in Insurance Oligopolies

Abstract We provide a simple framework for analyzing how competition affects the choice of audit structures in an oligopolistic insurance industry. When the degree of competition increases, fraud increases but the response of the industry in terms of investment in audit quality follows a U-shaped pattern. Following increases in competition, the investment in audit quality will decrease if the industry is initially in a low competition regime while it will increase when the industry is in a high competition regime. We show that firms will benefit from forming a joint audit agency only when the degree of competition is intermediate; in this case, cooperation might improve total welfare and we analyze the effects of contract innovation on the performance of the industry.

Keywords insurance fraud, audit quality, oligopolistic competition
JEL Classification D43, G22, L14, L22

1 Introduction

Insurance fraud is a universal and costly phenomenon. The cost of fraud in the US is estimated by the Insurance Information Institute to be between 10% to 20% of either claims, reimbursements or premiums. Insurance products like wage loss or medical

Received October 14, 2016

Nicolas Boccard
Economics Department, Universitat de Girona, Girona 17071, Spain

Patrick Legros (✉)
Economics Department, Northeastern University, Boston, MA 02115, USA; Universite Libre de Bruxelles (ECARES)
E-mail: plegros@ulb.ac.be

coverage both for physical and psychological traumas have inflated the total bill for insurers. Beyond this natural inflation, Berwick and Hackbarth (2012) estimate fraud to cost 6.7% of US national health expenditures ($\approx$177 bn\$); the burden even rises to 10% when considering the Medicare and Medicaid systems (cf. The Economist (2014)'s synthesis). The new insurance products have also opened the door to "soft" fraud practices i.e., the claim buildup of existing damages to one's car, body or mind. The Insurance Research Council (2015) estimates that claim fraud and buildup added some 15% to the premiums paid for "private passenger auto injury" coverage. According to the same source, about a fifth of claims for bodily injury and personal injury protection had the appearance of fraud, with high prevalence in Florida, New York, Massachusetts and Minnesota. Counter measures include independent medical exams, peer medical reviews and special investigative units.

The insurance sector has in the last twenty years increased efforts to fight fraud: by sharing information,[1] by investing in the training of special investigation units,[2] by using anti-fraud technologies and leveraging the availability of big-data,[3] or by advertising "toughness" with respect to fraud.[4] The sector also successfully lobbied for more stringent laws with the US Insurance Fraud Act of 1994 turning several types of fraud into federal crimes.[5] Furthermore, the National Insurance Crime Bureau (NICB) successfully lobbied in 1999 against a federal bill protecting personal privacy and limiting the use of nationwide databases. State Insurance Fraud Bureaus have grown from 8 in 1990 to 41 in 2015 (for a total of 51 states). However, these efforts vary significantly across countries.

At one extreme, like in the US or in South Africa, there are coordinated efforts in the industry to prevent and fight fraud, as well as active advertising by individual firms. At another extreme, like in most European countries, there are little attempts to co-

---

[1] Nearly all insurers use public databases like the "all claims database," 70% use the National Insurance Crime Bureau automobile database, 60% a database on claims (CLUE).

[2] According to a 1996 report of the Insurance Research Council (IRC), insurers had tripled their fraud control spending in 4 years and virtually all companies had a special investigation unit to investigate fraud.

[3] The Coalition Against Fraud conducts regular surveys among insurer companies. In their most recent survey (2014), 95% of the insurers used anti-fraud technologies (two-thirds of them used a software developed by a vendor), a 88% increase since the 2012 survey and 50% of them stated that the suspicious fraud activity increased significantly since 2011.

[4] The 1996 IRC report suggested that insurers places public awareness as the number one deterrent of fraud.

[5] See also the review by *Insurance Fraud* of state bills at their website.

ordinate the fight against fraudulent activities and often individual firms are reluctant to formally acknowledge the fraud problem;[6] nevertheless, even in these countries, individual firms develop contractual and organizational responses to the problem of fraud.

Differences in regulatory and legal systems could explain such differences across countries: for instance, strong punishments in courts of law make the deterrent effect of investments in fraud detection more effective; privacy laws prevent the sharing of data in the industry and limit industry coordination.[7] In this paper we leave aside differences in legal or regulatory regimes, and focus on a factor that has received little attention in the literature: the degree of competition among insurance companies and its relationship to the investment made by these firms for fighting against fraud.

We analyze an oligopoly model of insurance provision in which firms compete for horizontally differentiated consumers with two instruments: a "price" or pecuniary dimension—basically the premium and the reimbursement—and a "quality" or non-pecuniary dimension—additional services and the thoroughness and speed of the audit of claims. Thoroughness of the audit is important for the insurance company in order to credibly deter fraud; speed is important for the customers since it reduces the cost they will have to bear in case of a loss, hence increases their ex-ante utility from the contract. We analyze how the equilibrium contracts will be located on price and quality and how the relationship between these two dimensions will be modified when the degree of competition changes in the market. The degree of competition is an increasing function of the number of firms in the industry and a decreasing function of the index of differentiation of the consumers.

The combination of price and quality in contracts affects both the incentives of consumers to fraud and the financial returns to the firms. Market equilibrium dictates that customers obtain a certain level of expected utility—typically a non-trivial function of the degree of competition on the market—and firms will readjust their contracts in order to provide this level of utility at lowest cost to them. Hence, competition determines the substitution between the price and the quality instruments, and also the return on organizational or contractual innovations.

We show that the effect of competition on the level of private or cooperative audit

[6] cf. UK Insurance Fraud Taskforce (2016).

[7] For instance, European firms find it difficult to establish a common data base for insurance violations partly because of the privacy laws. In the US, insurance fraud is now considered a criminal offense, which makes private investments against fraud activity more likely to have a deterrent effect.

quality is non-monotonic because there is a shift from price competition to quality competition below a certain degree of competition.[8] This shift implies a U-shaped response of the equilibrium to the degree of competition; as competition increases, quality first decreases and then increases again. This theory provides therefore a rationale for the different attitudes of the US and European assurance industries towards fraud. It is also consistent with the observation that deregulation of the insurance supervision has resulted in appreciable price decreases in Western Europe since 1994 and in Japan since 1998, while contracts include now more non-pecuniary clauses than before.

In our model, most of the interesting effects arise when the degree of competition is intermediate. This is why comparing only perfect competition and monopoly—the two market structures usually considered in the literature—would not be very useful. Beyond this theoretical reason there are obvious empirical reasons for making the assumption of an oligopolistic market. First, there are significant barriers to entry in the industry (reputation, a large risk of bankruptcy, regulatory barriers). Second, in most countries, while there are few insurance companies having market shares above 1%, the concentration is moderate but certainly not low. For instance, in the US, the highest market share in 2015 was 10% (State Farm), and the 10^{th} highest market share was 3%. Table 1 summarizes this information for a variety of large markets.[9]

Table 1 Concentration in the Insurance Market

Period: (2010–2015)		Market share first 10	
Country	# Firms	LIFE	NON LIFE
Germany	568	64%	64%
UK	562	61%	62%
France	324	74%	75%
Spain	272	60%	64%
Italy	138	77%	85%
Japan	99	65%	80%
US	4,335	46%	55%

The rest of the paper is organized as follows. We present the basic model in the next section. We describe the properties of the market equilibrium and the comparative stat-

[8] As such our model belongs to a small theoretical literature, following Hart (1983), on the effect of increasing competition on market performance; see Legros and Newman (2014) for a literature survey.

[9] Data from the National Association of Insurance Commissioners, Insurance Europe and the Japan page at the Insurance Information Institute (all accessed in October 2016).

ics in Section 3. In section 4 we extend the basic model in two directions. First we consider the possibility for the industry to use an external audit agency. The U-shaped pattern that we obtain in the basic model implies the somewhat surprising result that an external audit agency is used in equilibrium only for *low* levels of competition; furthermore the use of an external audit can be welfare improving. In a second extension we consider contractual innovations, like rental car replacement, that increase the“quality” of the contract. We show that these innovations are complementary to the quality of audit. We conclude in Section 5.

2 The Insurance Market

An asset like a car, one's health or housing can be damaged with a small probability β.[10] Owners are risk-neutral with respect to money, value the use of the item at V and have an initial wealth w. If a damage occurs, the cost of repairing the item is L. We assume that

$$V > L > w. \tag{1}$$

Limited liability prevents the agents to borrow to repair the damaged item or such borrowing would have a large shadow price (for instance, the purchase of a new car after an accident will significantly reduce the consumption of other goods). However, they can transfer the risk of the loss to an insurance company in exchange for the payment of a premium. This specification makes the model of oligopoly competition tractable and is a reasonable alternative to the usual formalization of a concave utility function for money. The reservation utility of a consumer is

$$\underline{v} \equiv (1-\beta)\,V + w. \tag{2}$$

A contract specifies a premium P and a reimbursement R. To focus on the organizational problem, we ignore adverse selection problems and assume that wealth w is observable. A contract is *feasible* if the agent is able to pay the premium and if the

[10] We will make clear later on in the paper how small β needs to be, see page 393. To fix ideas, the odds of a car occupant in the US dying in a transportation accident was 1 in 47,718 in 2013 while the lifetime odds were 1 in 606 for a person born in 2013. The frequency of car theft in France was 0.3% in 2015 (109 thousands for 32 millions of vehicles), that of car accident was 0.2% (56 thousands) and that of a burglary was 1.1% (382 thousands for 34 millions of housings).

reimbursement net of the premium covers the cost of damage, i.e., if

$$w \geqslant P, \tag{3}$$

$$R \geqslant L - (w - P). \tag{4}$$

Obviously, (4) combined with (1) requires that the reimbursement is larger than the premium

$$R > P \geqslant 0. \tag{5}$$

If the loss can be verified at no cost by the firm, these conditions characterize feasibility. However, we are interested in situations where losses can be verified only if the insurance firm invests in audit. We assume the following audit technology:

- If the claim is *honest*, audit shows that there is no fraud with probability 1.
- If the claim is *dishonest*, the audit detects the fraud with probability q.

The success probability q is chosen by the firm, e.g., is affected by the experience of the claim adjusters, the ex-ante thoroughness of the description of the item, and the time spent on the audit. The cost of an audit of quality q is $C^f(q)$ for the firm, that is strictly increasing, strictly convex in q and satisfies $C^f(0) > 0$.

While a firm might indeed invest in quality q by hiring experienced claim adjustors and setting strict auditing procedures, there is a decision to be made ex-post of whether or not to audit a claim. We thus assume that the insurance firms can commit *ex-ante* to a quality q of audit but not to its *ex-post* frequency.[11]

The agent is reimbursed unless fraud is established, in which case he pays a fine F to the state.[12] Generally, measures to control fraud impose direct and indirect costs on consumers. For instance, taking pictures of a car when signing a contract reduces the possibility of a consumer to make a claim for a pre-existing default but increases the transaction costs of the contract. Inspection of a damaged car via "authorized experts" limits the possibilities of collusion with the repairer and of false claims but increases the delay for repairs or forces the consumer to free time to go to the

[11] On the issue of commitment to audit, see Khalil (1997) and Picard (1996).

[12] The penalty could include an amount F_0 awarded to the company. We set $F_0 = 0$ to simplify exposition. F is the product of a fine (or disutility of jail term) and of the probability of being found guilty in court. We may interpret the recent criminalization of insurance fraud in the US as an attempt to keep F at a reasonable level despite the increased congestion of the judicial system, that is despite the lower probability of being found guilty in court in "reasonable time."

inspection.[13] An example of indirect cost is the policyholder's perception of "mistrust" by the insurance company. The overall opportunity cost is denoted $C^a(q)$ and is assumed to be increasing and convex in audit quality, with $C^a(0) > 0$.

We consider the following timing of events:

1. Stage 1: Contract choice
   (a) Insurer i chooses an audit quality q^i and offers insurance contracts with premium P^i and reimbursement R^i.
   (b) Agents observe $\left(q^i, P^i, R^i\right)_{i=1}^n$ and decide to purchase or not an insurance contract.
2. Stage 2: Damage and claim
   (a) Agents incur a damage with an exogenous "small" probability β.
   (b) An agent who purchased a contract (P^i, R^i) can claim a loss and ask for a payment R^i.
   (c) The insurance company decides to audit or not the claim and pays the agent according to the result of the audit.

Our two-stage game is solved by backward induction using the concept of Perfect Bayesian Equilibrium (PBE). In the next section we analyze the fraud and control game between a policyholder and its insurer. Then, we compute the expected utility of agents and firms conditional on choosing (q^i, P^i, R^i) and we derive the symmetric equilibrium of the quality and contract competition between insurance companies.

To analyze how the intensity of competition affects the equilibrium design of contracts and the choice of audit quality, we assume that consumers are horizontally differentiated. The *residual demand* facing an individual firm that offers consumers a level of expected utility v when the other firms in the industry offer the level v^* is $D\left(v - v^*\right) = t(v - v^*) + \frac{1}{n}$.[14] Note that $D(0) = \frac{1}{n}$ and $D'(0) = t$. The index

[13] Some contractual innovations tend to reduce these direct costs; for instance as soon as 1989, Allstate, a major US player, introduced a Priority Repair Option to accelerate repairing and reduce the loss adjustment expenses.

[14] This indirect form can be obtained, for instance, in the circular city setting introduced by Salop (1979). Consumers are uniformly distributed over the unit circle and bear a mobility cost $\frac{2}{t}$ per unit of distance while the n firms are located at an equal distance one from another (i.e., $\frac{1}{n}$). Agents buy insurance contracts from the company that offers them the highest expected utility.

$\theta \equiv nt$ proxies for the degree of competition on the market. By varying the mobility cost t or the number of firms we will be able to vary exogenously the intensity of competition. The ongoing deregulation in many countries or the greater visibility of entrants permitted by the information technology "revolution" are examples of exogenous factors affecting t and n.

3 Equilibrium and Audit Response to Competition

In this section we first solve for the game of fraud and audit played by an insurer and one of its policyholders (Propositions 1 and 2). We then show that the equilibrium analysis and the comparative statics can be simply captured in an Edgeworth box in which the "commodities" are the reimbursement and the quality of audit. It is then possible to show the existence of two regimes: one in which the reimbursement is compatible with a premium level that does not make the liability constraint of the consumers binding, the other where this constraint is binding; these two regimes correspond also to equilibrium regimes when the degree of competition is high or low.

3.1 The Fraud and Audit Game

Since a damage is private information to the policyholder, if an insurer were to pay all claims without audit then all policyholders would fraud.[15] However systematic audit is not credible. We therefore have a double moral hazard problem: that of inducing insurance companies to audit with a high probability and that of inducing policyholders to fraud with a low probability. Clearly, the equilibrium is in mixed strategies with σ being the audit probability and τ the fraud probability. The following proposition characterizes the equilibrium levels of fraud and audit (All proofs missing from the text appear in the Appendix).

Proposition 1. In a PBE, the game of fraud and control following a contract (P, R) with a quality q (second stage) has a unique Nash equilibrium; it is in mixed strategies with

$$\sigma^* \equiv \frac{\beta C^f(q)}{(1-\beta)\,(qR - C^f(q)\,)} \; and \; \tau^* \equiv \frac{R}{q\,(R+F)}. \tag{6}$$

[15] In this model all agents will fraud if they can get away with it. A recent study of the Coalition Against Insurance Fraud [4] reveals that people fraud to save money or reduce costs, to get expensive work done they would not otherwise be able to afford and to "get back" at insurance companies.

Observe that while the audit quality q generates future costs for the agent and the firm, it brings benefits too since it influences the desire of the agent to fraud and of the firm to audit. Its effect on the utilities of insurers and consumers can be summarized by the average cost functions $c^f(q) \equiv C^f(q)/q$ and $c^a(q) \equiv C^a(q)/q$. These two identities imply that the average cost functions c^a and c^f are U-shaped and reach minima q^a and q^f that we call *ideal* audit qualities. Following stylized facts[16] we assume that insurers desire a better audit quality, relative to policyholders i.e., that

$$q^a < q^f. \tag{7}$$

Using the equilibrium levels of fraud σ^* and control τ^*, the expected utility levels of an agent u and the expected profit of the firm π are

$$u(q, P, R) \equiv V + w - P - \beta(L - R) - \beta\frac{Rc^a(q)}{R + F},$$
$$\pi(q, P, R) \equiv P - \beta\frac{R^2}{R - c^f(q)}.$$

The ratios $\frac{Rc^a(q)}{R+F}$ and $\frac{R^2}{R - c^f(q)}$ are the surplus loses to insureds and insurers of the fraud. These two equations motivate two observations. First, since both $\pi(q, P, R)$ and $u(q, P, R)$ are increasing with q at 0 and decreasing at 1, offering a contract (q, R, P) with q outside the interval $[q^a, q^f]$ is strictly dominated by offering either (q^a, R, P) or (q^f, R, P). In a PBE, the quality of audit lies in the interval $[q^a, q^f]$, where the bounds are the average cost minimizing levels. Second, if $R > P + L - w$, the insurer can decrease both the reimbursement R and the premium P in order to preserve his market share but increase his per-consumer profit.[17] Therefore, in equilibrium, constraint (4) binds and we can reduce competition on the "pecuniary dimen-

[16] It is well documented—e.g., the white paper (2000)—that insurance salesmen (the demand side) are calling for low audit quality while claim payers (the supply side) are calling for high audit quality.

[17] If $R > P + L - w$ consider $\Delta R < 0$ and $\Delta P = \frac{\partial u}{\partial R}\Delta R = \beta\Delta R\left(1 - \frac{Fc^a(q)}{(R+F)^2}\right) > \beta\Delta R$. As $u(q, P + \Delta P, R + \Delta R) = u(q, P, R)$ and the firm has the same market share. Since $\pi(q, R, P) = P - \beta R\frac{R}{R - c^f(q)}$ and $R\left(R - 2c^f(q)\right) < \left(R - c^f(q)\right)^2$, we have $\frac{\partial\pi}{\partial R} = -\beta R\frac{R - 2c^f(q)}{\left(R - c^f(q)\right)^2} > -\beta < 0$. Hence, the total per-capita profit variation $\Delta\pi$ is greater than $\Delta P - \beta\Delta R > 0$.

sion" to the reimbursement R which becomes a proxy for the premium. We summarize this discussion in the following proposition.

Proposition 2. In a PBE, an insurer chooses a quality between the ideal of firms q^f and of consumers q^a and offers minimal reimbursement with $R = P + L - w$.

Using the equality in the feasibility constraint (4), the liquidity constraint (3) becomes

$$R \leqslant L, \tag{8}$$

while a positive premium (5) yields

$$L - w \leqslant R. \tag{9}$$

Hence in a PBE, the reimbursement varies in the interval $[L - w, L]$ and the premium varies in the interval $[0, w]$. The indirect utility function of a policyholder and the per-capita profit of an insurer can then be written:

$$u(q, R) \equiv V + (1 - \beta)L - R\left(1 - \beta + \frac{\beta c^a(q)}{R + F}\right), \tag{10}$$

$$\pi(q, R) \equiv R\left(1 - \beta \frac{R}{R - c^f(q)}\right) - L + w. \tag{11}$$

Note that a zero premium ($R = L - w$) generates losses for an insurer, hence the liquidity constraint (8) is the only relevant constraint.

To avoid trivialities, we assume that the sum of payoffs $u(q, R) + \pi(q, R)$ is larger than the agent's reservation utility $\underline{v}$ when contract variables q and R vary in their PBE range. This is equivalent to assuming that V is large enough. Precisely:

$$\frac{V}{L} \geqslant \frac{L}{L - c^f(q)} + \frac{c^a(q)}{L + F}, \quad \forall q \in \left[q^a, q^f\right]. \tag{12}$$

3.2 Competition

A symmetric equilibrium of the first stage competition among insurers is denoted (q^θ, R^θ); it yields a unique equilibrium utility level $v^\theta \equiv u(q^\theta, R^\theta)$ to all consumers. The choices of reimbursement R and audit quality q for insurers are limited by the feasibility constraint $R \leqslant L$ and the participation constraint $u(q, R) \geqslant \underline{v}$. When all other firms choose their equilibrium actions q^θ and R^θ, firm i's profit is

$$\Pi_i(v^\theta, q^i, R^i) \equiv D\left(u(q^i, R^i) - v^\theta\right)\pi(q^i, R^i). \tag{13}$$

In a symmetric oligopoly equilibrium, when firm i chooses $\left(q^i, R^i\right)$ to maximize Π_i, taking v^θ as given, we must have in equilibrium $u\left(q^i, R^i\right) = v^\theta$. Whenever $\underline{v} < v^\theta$ and the liquidity constraint $R \leqslant L$ is not binding, the two FOCs for an interior solution are

$$\frac{\partial\,\Pi_i(q,R)}{\partial\,R} = 0 \Leftrightarrow D' u_R \pi + D\pi_R = 0,$$

$$\frac{\partial\,\Pi_i(q,R)}{\partial\,q} = 0 \Leftrightarrow D' u_q \pi + D\pi_q = 0.$$

As $D' = t$ and $D = D(0) = \frac{1}{n}$, those are equivalent to

$$\frac{u_R}{u_q} = \frac{\pi_R}{\pi_q}, \tag{14}$$

$$-u_q \theta \pi = \pi_q. \tag{15}$$

Equation (14) defines the contract curve of an Edgeworth Box with commodities R and q. Note that if the marginal rates of substitution between quality and reimbursement of policyholders $\frac{u_R}{u_q}$ and insurers $\frac{\pi_R}{\pi_q}$ are not equalized, the insurer can offer a better contract $(\tilde{q}, \tilde{R})$ in the sense that it leaves its clients indifferent with respect to (q, R) and preserves its market share but increases its per-capita profit.

Equation (15), which holds for only $\underline{v} < v^\theta$, illustrates the traditional trade-off for a change in the strategic variable q between the marginal effect $-u_q\theta\pi$ (loss of clients) and the infra-marginal effect π_q (higher margin on remaining clients). It characterizes the position of the oligopoly equilibrium on the contract curve. The analysis of the contract curve is therefore crucial to understand the relationship between changes in the exogenous parameter θ and changes in the equilibrium.

3.3 The Edgeworth Box and Comparative Statics

When the liquidity constraint $R \leqslant L$ is not binding, a Pareto optimum is characterized by the equality of the marginal rates of substitution $MRS^a = \frac{u_R}{u_q}$ and $MRS^f = \frac{\pi_R}{\pi_q}$ as depicted on the Edgeworth box of Figure 1. The iso-profit curve (solid line) is tangent to the iso-utility curve (dashed line) at point x, the absolute value of the slope is equal to the marginal rate of substitution. The arrows indicate the direction of increasing payoffs for each type of agent.

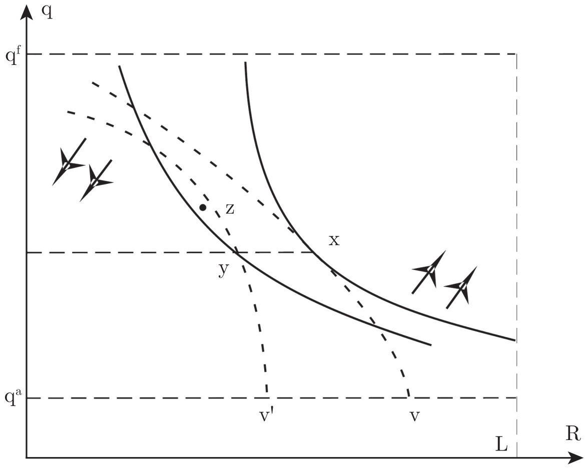

Figure 1 The Edgeworth Box

As the risk β is small, reimbursement and audit quality are substitute "goods" for the insurer and substitute "bads" for the policyholder. Since the risk of damage is small, audit is rare, and audit quality generates only second order effects while the premium generates first order effects. These two properties combine to produce a MRS decreasing with quality.[18] As shown on Figure 1, the MRS of the insurer at y is lower than at x while the opposite holds for the policyholder, thus MRSs are equalized at a point like z, above y.

However, the fact that reimbursement and quality are substitute for the policyholder and for the insurance company does not necessarily imply that these variables are substitute along the contract curve. Lemma 1 below establishes that the contract curve is indeed downward-sloping in the interior of the Edgeworth box. Hence, *in equilibrium,* reimbursements R and audit quality q are substitute.

Lemma 1. The contract curve is downward-sloping in the interior of the Edgeworth box (R, q).

We can now analyze the consequences of changes in the degree of competition. Since increasing utility to the policy holder is costly for the insurance company, it is immediate that consumer utility decreases when competition decreases. Because the

[18] Note that it is the combination of small β and substitution between R and q that yields the decreasing MRS. The property of decreasing MRS is usually a consequence of the complementarity of the goods and the decreasing returns to consumption.

risk of damage is small, most of the competitive pressure applies initially to prices that is to R. Since the contract curve is decreasing in the (R, q) space, it follows that the increase in price is accompanied by a decrease in the quality of audit as competition decreases (see Figure 2). When competition is weak, the insurance premium is equal to the willingness to pay of consumers and the limited liability binds. Weaker competition will then lead to an increase in quality, i.e., move up vertically towards their ideal quality q^f (see Figure 2).

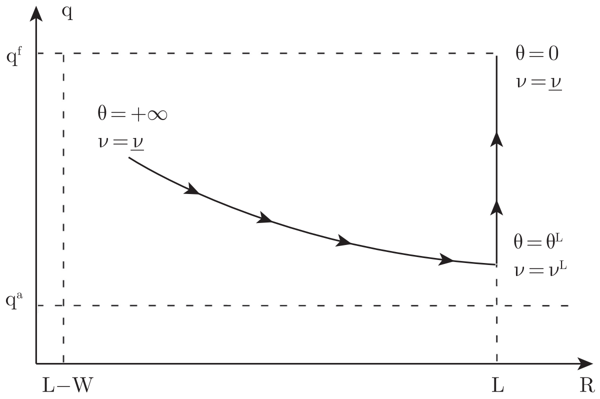

Figure 2 The Equilibrium Path

Proposition 3. There exist two wealth levels $\underline{w}$ and $\bar{w}$ such that

(i) for an intermediate wealth $w \in [\underline{w}; \bar{w}]$, the symmetric equilibrium between insurers has two regimes:

- **weak competition regime** $\theta \in [0; \theta^L]$: equilibrium contracts feature full insurance, maximal premium while quality is decreasing with θ.

- **strong competition regime** $\theta \in [\theta^L; +\infty]$: equilibrium reimbursement and premium decrease with θ while quality increases with θ.

(ii) If $w < \underline{w}$, only the weak regime applies

(iii) If $w > \bar{w}$, only the strong regime applies.

While the quality of audit is not a monotone function of the degree of competition, the level of fraud unambiguously increases when competition increases. That fraud increases is immediate from (6); in the weak competition regime since the premium is fixed while the audit quality tends to decrease, i.e., the *average cost* of audit increases making insurers *less credible* auditors; consumers then fraud more.

The result is less obvious in the strong competition regime because there are two opposite forces at play. First, the increase in quality reduces the cost of audit and firms are *more credible* auditors and fraud should decrease as a consequence. Second, competition reduces the price *and* the benefit of audit: this effect makes firms less credible auditors. Since the price dimension is the main channel of competition, the net effect is a reduction in the incentives of the firm to audit and an increase in the equilibrium level of fraud.

Corollary 1. Fraud increases with the degree of competition.

4 Two Extensions

In this additional section we analyze the consequences of two strategies that the industry might adopt to fight fraud (and alter contracts features). We first consider the cooperative creation of a common audit agency and then the use of replacement clauses in contracts.

4.1 Centralized Audit Agency

In our model but also in practice, a major source of cost for insurers is their inability to commit to an audit frequency. They sometimes launch campaigns to commit to systematic audit of claims but this conduct never lasts very long. One alternative—used frequently in the US or in Europe—is to have a third party offer high quality audit services; the insurance companies remain free to use their own internal audit divisions or the third party's agency.[19]

To simplify we assume that the agency is created in a cooperative fashion by the industry (for instance by selecting the best offer in an auction for auditing services). The costs of creation (if any) are shared between the firms. There is commitment value in the creation of such an agency only if policy holders anticipate that the agency will be effectively used. The timing of events is now:

- The industry creates cooperatively an audit agency with quality $\bar{q}$.
- Insurers choose their control structure q^i – $q_i \leqslant \overline{q}$ – and offer insurance contracts (P^i, R^i).

[19] We ignore commitment on contracts because it is more difficult to enforce and would be probably illegal under current antitrust legislation.

- Consumers observe $\left\{\bar{q}, (q^i, P^i, R^i)_{i=1}^n\right\}$ and purchase an insurance contract or not.
- A policyholder incurs a loss or does not incur a loss; he can then make a claim for reimbursement.
- Insurers decide to audit or not claims and if so to use the agency or not (use $\bar{q}$ or q^i).

We show in Proposition 4 below that the intermediary is a valuable commitment device on quality only if there is "some" competition on quality but none on premiums, that is when $\theta \in \left[\underline{\theta}; \theta^L\right]$. Alternatively, since creating a common agency reduces competition among firms to that on the level of premiums, the value of commitment on audit is large when competition on premiums is not too severe. However, a commitment to a larger audit quality will intensify competition on premiums.

To understand this result, observe first that once an agency of quality $\bar{q}$ has been set-up, an individual firm can only deviate to a higher quality. Indeed if a firm invests in quality $\hat{q} < \bar{q}$, it will choose ex-post to use the common agency since $c^f(\bar{q}) < c^f(\hat{q})$. The new game has therefore a smaller strategy set (with respect to the original model). When $\bar{q} \leqslant q^\theta$, the equilibrium quality in the original model, q^θ, remains a Nash equilibrium of the new game. If there is a positive fixed cost of creating the agency, the insurers will not create it.

Hence the only possibility is to set $\bar{q} > q^\theta$ and two cases must be considered. If at the original equilibrium the liquidity constraint is not binding $(\theta > \theta^L)$ then competition forces firms to decrease the premium R to compensate consumers for the committed high quality of audit; this eventually results in lower equilibrium profits (cf. proof of Proposition 4 in the appendix, 397) and insurers will not create the agency in the first place.

The second case is in the weak regime $(\theta \leqslant \theta^L)$. In this case, it is possible to increase $\bar{q}$ above q^θ without putting too much pressure on premiums i.e., the equilibrium in the new game still features the binding liquidity constraint. We show in the appendix, that the optimal choice by the industry is increasing in θ and that there exists a threshold $\underline{\theta}$ below which the agency sets the monopoly audit quality q^f as shown in Figure 3; hence there is bunching on $[0, \underline{\theta}]$.

Proposition 4 If there is no transaction cost for the creation of an audit agency, it is created only in the weak competition regime $(\theta \leqslant \theta^L)$. The common audit quality $\tilde{q}(\theta)$ is greater than the equilibrium value q^θ and equal to the insurers' ideal level for low levels of competitiveness.

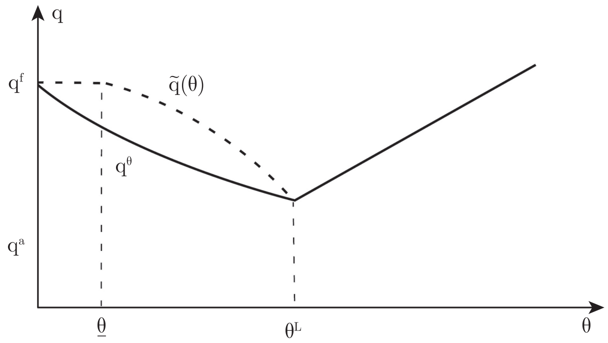

Figure 3 Quality with and without the Agency

Cooperation raises obvious competition policy concerns. To analyze the welfare consequences of cooperation in the industry, assume that the competition authority puts as much weight on consumers as on firms and wants to maximize total surplus. Since the total mass of consumers is one, the government maximizes $\pi(q,R)+u(q,R)$ under the constraints $\pi(q,R) \geqslant 0, u(q,R) \geqslant \underline{v}$ and $R \leqslant L$. Let $\hat{v}$ be the utility achieved at this optimum. By continuity of the equilibrium solution, there exists an index $\hat{\theta}$ such that $v^{\hat{\theta}} = \hat{v}$. It is immediate that the social optimum is the equilibrium outcome corresponding to $\hat{\theta}$.

Whether or not the individual rationality constraint is binding, using the equilibrium condition (15) applied at $\hat{\theta}$, the social optimum also solves the FOC, $\pi_q = -u_q \Leftrightarrow \hat{\theta} = 1/\pi$. This gives us a rather simple way to assess whether the current profit conditions correspond to the social optimum. By a simple algebraic manipulation of the FOC we obtain:

$$\pi_q = -u_q \Leftrightarrow \frac{c_q^f}{-c_q^a} = \left(\frac{R - c^f(q)}{R}\right)^2. \tag{16}$$

Therefore, if $c^f(.)$ has a large curvature (highly convex) then the solution of (16) is close to q^f while if $c^a(.)$ has the largest curvature, the solution of (16) is close to q^a.

Although fears of increased market power for the insurance companies are legitimate, coordination on a common auditing agency can be welfare increasing. Recall indeed that it is a commitment to set a control structure that *reduces* the cost of audit, thus *increases* the desire of firms to audit and ultimately *reduces* fraud in the economy. While consumers are hurt when the quality of audit increases (see equation (10)), total welfare may increase if the cost savings are large enough for firms.

To assess the validity of this claim we note that the external audit agency is created only for $\theta \leqslant \theta^L$ and that welfare depends only on audit quality in this regime.

The change from q^{θ} to $\tilde{q}(\theta)$, resulting from the entry of the agency, will be welfare improving only if $q^{\hat{\theta}}$ is closer to $\tilde{q}(\theta)$ than to q^{θ}. As we can see on Figure 3 above, this amounts to say that $q^{\hat{\theta}}$ is large. Now, the solution of equation (16) is large only if the average cost $c^f(.)$ has a quite larger curvature than $c^a(.)$. Hence the following corollary only enunciates a possibility, not a certainty.

Corollary 2. There exists $\hat{\theta} > 0$ such that the external audit agency can be welfare improving when the degree of competition is intermediate: $\theta \in [\hat{\theta}, \theta_L]$.

4.2 Replacement Clauses

Insurance contracts often include clauses that decrease the opportunity cost of audit for policyholders (e.g., Allstate's priority repair option). A relevant example is when the policyholder of a damaged car is provided free of charge a replacement vehicle while its car is examined by a company expert and then repaired in a "certified" garage.

The effect of a replacement clause can be captured in our model by a reduction $\delta + \epsilon$ of the opportunity cost $C^a(q)$ together with an increase δ of the audit cost $C^f(q)$. It is clear that such a substitution from money to in-kind transfer has to satisfy $\epsilon > 0$ in order to appear in equilibrium, for otherwise a premium reduction would dominate quality improvement for the firm. The sum of equilibrium payoffs $u + \pi$ given by equations (10) and (11) increases with the inclusion of the replacement clause only if the cost terms satisfy

$$\epsilon > \underline{\epsilon} \equiv \delta \left(\frac{R}{R - c^f(q)} \frac{R + F}{R - c^f(q) - \delta} - 1 \right), \tag{17}$$

We can now study how the introduction of the replacement clause affects the equilibrium contracts. We assume that the clause is immediately adopted by all insurers as soon as it is "discovered."

We first consider the strong competition regime. The equilibrium is characterized by the *no-arbitrage* equation (14) linking premium and quality on the one hand and the *market* equation (15) linking a *marginal* effect (LHS) and an *infra-marginal* effect (RHS) on the other hand (recall $\theta = D'/D$). The introduction of a replacement clause reduces the per-capita profit π, the LHS of (15) but at the same time increases the marginal per-capita profit π_q, the RHS of (15).[20] Hence, all insurers have an incentive to put more weight on their revenue improving strategies i.e., to increase either

[20] The same occurs if one considers the equation $-u_R \theta \pi = \pi_R$.

premium or quality. As the two variables are substitute along the equilibrium path, intuition would suggest that one variable increases while the other variable decreases. However the effect of the replacement clause is to move the whole contract curve upward and in fact both reimbursement and quality are increased as a result of the introduction of the replacement clause.

Proposition 5. The introduction of the replacement clause increases the audit quality and the premium (whenever not already binding). Furthermore insurers' profits increase and consumer's utility does not fall.

5 Conclusion

We consider a model of insurance oligopoly in which firms compete both on prices (premium and reimbursement) and on the quality of audit. If a firm increases audit quality, it affects the residual demands of all other firms. Therefore audit quality has the flavor of a public good, leading to standard free riding and possibly under-investment problems. However, while standard free-riding reasoning would suggest that the quality of audit decreases with the degree of competition, we show an inverted U-shaped relationship between competition and audit quality.

We believe that by bridging industrial organization and audit performance our approach will prove useful for empirical work on insurance fraud. Our theoretical findings seem consistent with stylized facts about the US and the French markets. As already noted in the introduction, the US industry is characterized by a high degree of competition and firms advertise aggressively about the quality of their control structure. In France there is a lower degree of competition and firms are viewed as "soft" on the question of fraud (no awareness campaign, no appearance of the word "fraud" on web sites). Consistent with our findings, French firms have cooperated to create a control agency of high quality (ALFA) which is quite active with over 28 thousand identified fraudulent claims in the auto insurance market in 2013.

Acknowledgements This research was started while both authors were at the University of Liege and was partly financed then by a TMR network contract FMRX-CT98-0203. Financial support from Generalitat de Catalunya (AGAUR, XREPP) and Ministerio de Economia y Competitividad (contract ECO2013-45395-R) are gratefully acknowledged. Legros acknowledges the financial support of the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013 / ERC grant agreement No. 339950).

References

Aghion P, Dewatripont M, Rey P (1999). Competition, financial discipline and growth, Review of Economic Studies. 66(4): 825–852

Berwick D M, Hackbarth A D (2012). Eliminating waste in US health care. JAMA, 307(14): 1513–1516

Border K, Sobel J (1987). Samurai accountant: A theory of auditing and plunder. Review of Economic Studies, 54(4): 525–540

Coalition Against Fraud (2014). The state of insurance fraud technology—A study of insurer use, strategies and plans for anti–fraud technology SAS. https://goo.gl/oMI6zX

Graetz M, Reinganum J, Wilde L (1989). The tax compliance game: Toward an interactive theory of law enforcement. Journal of Law Economics and Organization, 2(1): 1–32

Hart O (1983). The market mechanism as an incentive scheme. Bell Journal of Economics, 14: 366–382

Insurance Research Council (2015). Fraud and buildup and auto injury insurance claims. https://goo.gl/CuXJHo

Khalil F (1997). Auditing without commitment. Rand Journal of Economics, 28: 629–640

Legros P, Newman A (2002). Contracts, ownership, and industrial organization: Past and future. Journal of Law, Economics and Organization, 30(1): i82–i117

Melumad N D, Mookherjee D (1989). Delegation as commitment: The case of income tax audits. Rand Journal of Economics, 20(2): 139–163

Mookherjee D, Png I (1989). Optimal auditing, insurance and redistribution. Quarterly Journal of Economics, 104: 399–415

OECD (2015), OECD Insurance Statistics 2007–2014. http://dx.doi.org/10.1787/ins_stats–2015–en

Picard P (1996). Auditing claims in the insurance market with fraud: The credibility issue. Journal of Public Economics, 63: 27–56

Reinganum J, Wilde L (1985). Income tax compliance in a principal–agent framework. Journal of Public Economics, 26(1): 1–18

Salop S (1979). Monopolistic competition with outside goods. Bell Journal of Economics, 10:141–56

Tennyson S (1997). Economic institutions and individual ethics: A study of consumer attitudes toward insurance fraud. Journal of Economic Behavior and Organization, 32(2): 247–65

Townsend R (1979). Optimal contracts and competitive markets with costly state verification. Journal of Economic Theory, 21:265–293

UK Insurance Fraud Taskforce (2016). https://goo.gl/6aoJq3

White paper (2000). National insurance fraud forum organized by the coalition against insurance fraud. The International Association of Special Investigation Units and the National Insurance Crime Bureau, Washington, D.C.

Appendix

Proof of Proposition 1

Characterization of the equilibrium of the fraud and control game in a perfect Bayesian equilibrium.

Step 1. Candidate equilibria

A customer's strategy is to defraud with probability σ while an insurer's strategy is to control claims with probability τ. The rate of claims is the sum of truthful ones and fraudulent ones: $\beta + (1-\beta)\sigma$. Because a verification is successful only with probability q (and cost F) and defrauders don't suffer being audited, the utility functions of the consumer and the insurer are

$$u(\sigma,\tau) = V + w - P - \beta(L-R) + (1-\beta)\sigma\left((1-\tau q)R - \tau q F\right) - \beta\,\tau\, C^a(q),$$
$$\pi(\sigma,\tau) = P - \beta R - C^f(q)\tau\left(\beta + (1-\beta)\sigma\right) - (1-\beta)\sigma(1-\tau q)R.$$

The agent's best reply is given by the frequency τ that makes the coefficients of σ nil in $u(\sigma,\tau)$. We obtain

$$\tau \overset{\textstyle<}{\underset{\textstyle>}{=}} \tau^* \equiv \frac{R}{q(R+F)} \Rightarrow \sigma \in \overset{\{1\}}{\underset{\{0\}}{[0;1]}}. \tag{18}$$

The slope coefficient of τ in π is $S \equiv \sigma(1-\beta)\left(qR - C^f(q)\right) - \beta C^f(q)$. Three cases can occur:

► If $\tau < \tau^*$ then $\sigma = 1$ so that $S = (1-\beta)qR - C^f(q)$. Hence the optimal behavior is to play $\tau = 1$ if $S > 0$. We have a contradiction if $\tau^* \leqslant 1$. Otherwise the equilibrium is $\sigma = 1$, $\tau = 1$. The last possibility is $S < 0$ so that $\tau = 0$ is optimal; in that case the equilibrium is $\sigma = 1$, $\tau = 0$.

► If $\tau > \tau^*$ then $\sigma = 0$ so that $S = -C^f(q)\,\beta < 0$ implying that τ should optimally be set equal to 0, a contradiction.

► If $\tau = \tau^*$ then any $\sigma \in [0;1]$ is optimal and to make τ^* optimal as well we look for the value of σ that equalizes S to zero. Formally

$$\sigma \overset{\textstyle>}{\underset{\textstyle<}{=}} \sigma^* \equiv \frac{\beta C^f(q)}{(1-\beta)\left(qR - C^f(q)\right)} \Rightarrow \tau \in \overset{\{1\}}{\underset{\{0\}}{[0;1]}}. \tag{19}$$

If $\sigma^* < 0 \Leftrightarrow R < C^f(q)/q$ or $\sigma^* > 1 \Leftrightarrow C^f(q)/q > (1-\beta)R$ then $S < 0$ thus $\tau = 0$ is optimal and $\sigma = 1$ is the optimal response. We have a pure strategy equilibrium. When $0 \leqslant \sigma^* \leqslant 1$ and $\tau^* \leqslant 1$ the Nash equilibrium is (σ^*, τ^*). Since the agent is indifferent between fraud and honesty, his equilibrium utility is $u(q,R,p) = V + w - P - \beta\left(L - R + \dfrac{RC^a(q)}{(R+F)q}\right)$ while the equilibrium per-capita profit of the insurer is $\pi(q,R,p) \; = \; P - (\beta + (1-\beta)\sigma^*)\,R = p - \beta R \dfrac{R+F}{R - C^f(q)/q}$.

Step 2. Elimination of the pure strategy equilibria

The pure strategy equilibrium $(\sigma = 1, \tau = 0)$ leads to payoffs $\hat{a} = V + w - \beta L - P + R$ and $\hat{\pi} = P - R$. Thus, the contract is producing no economic surplus. Since the no-losses condition of the insurer yields $P \geqslant R$, the agent is better off not signing this contract. Hence this situation will not appear in a PBE.

The pure strategy equilibrium $(1,1)$ exists only when $\tau^* > 1 \Leftrightarrow R > q(R+F)$. Payoff are $\hat{\pi} = P - R + (1-\beta)qR - \phi_\pi(q)$ and $\hat{a} = V + w - \beta(L + C^a(q)) - (1-\beta)q(R+F) + R - P$. We claim that offering (q,R) is a dominated strategy. Indeed, reducing R and P so as to keep $P - R + (1-\beta)qR$ constant maintains $\hat{a}$ and $\hat{\pi}$ constant. The insurer can therefore bring the equality $R = q(R+F)$. With this new contract, the threshold is $\tau^* = 1$ and $\hat{a} = u(\sigma^*, \tau^*)$ because $\tau^* = 1$ makes the coefficients of σ nil in $u(\sigma,\tau)$. Hence, by switching from the pure strategy equilibrium to the mixed one, the insurer keeps its customers (their utility level remains constant). Still, there is a benefit of inducing the mixed strategy equilibrium with $\sigma^* < 1$ because $\left.\dfrac{\partial \pi}{\partial \sigma}\right|_{\tau=\tau^*} = -C^f(q)\tau^*(1-\beta) < 0$ means that $\pi(\sigma^*,\tau^*) > \hat{\pi}$ when $R = q(R+F)$. ■

Proof of Lemma 1

The contract curve linking all Pareto optima is decreasing in the (R,q) *space.*

i) $MRS^a = \dfrac{u_R}{u_q}$ and $MRS^f = \dfrac{\pi_R}{\pi_q}$ are positive.

For β small, $u_R = -1 + \beta - \beta \dfrac{Fc^a}{(R+F)^2} < 0$ and $\pi_R = 1 - \beta \dfrac{R\left(R - 2c^f\right)}{\left(R - c^f\right)^2} > 0$.

Since c^a is increasing and c^f is decreasing on $\left[q^a, q^f\right]$, $u_q = \dfrac{-\beta R}{R+F} c^a_q < 0$ and $\pi_q = \dfrac{-\beta R^2}{(R - c^f(q))^2} c^f_q > 0$, which yields positive MRSs.

ii) We show that $\left.\dfrac{\partial MRS^f}{\partial R}\right|_{q=cte} > 0$ and $\left.\dfrac{\partial\, MRS^a}{\partial\, R}\right|_{q=cte} < 0$.

As $\pi_{RR} = -2\beta \dfrac{(c^f)^2}{\left(R - c^f\right)^2}$ and $\pi_{qR} = 2\beta R \dfrac{c^f_q c^f}{(R - c^f)^3} < 0$, we have for β small,

$$\left.\frac{\partial\, MRS^f}{\partial\, R}\right|_{q=cte} = \frac{\pi_{RR}\pi_q - \pi_R\pi_{qR}}{(\pi_q)^2} > 0 \text{ since the } \pi_R \text{ effect dominates the } \pi_q \text{ effect.}$$

Likewise $u_{RR} = \beta\frac{Fc^a}{(R+F)^3} > 0$ and $u_{qR} = -\beta\frac{Fc_q^a}{(R+F)^2} < 0$ imply $MRS_R^a = \frac{u_{RR}u_q - u_R u_{qR}}{(u_q)^2} < 0$. ■

Proof of Proposition 3

There exist two wealth levels $\underline{w}$ and $\bar{w}$ such that for an intermediate wealth $w \in [\underline{w}\,;\bar{w}]$, the symmetric equilibrium between insurers has two regimes
(i) weak competition regime $\theta \in [0;\theta^L]$: equilibrium contracts feature full insurance, maximal premium while quality is decreasing with θ.
(ii) strong competition regime $\theta \in [\theta^L;+\infty]$: equilibrium reimbursement and premium decrease with θ while quality increases with θ.
If $w < \underline{w}$, only the weak regime applies while if $w > \bar{w}$, only the strong regime applies.

We use a series of lemmas. In the first, we show that consumer utility decreases as we move down on the contract curve (when $R \nearrow$ and $q \searrow$). Next we use this property to characterize the solution of $\hat{P}(L,v) \equiv \max_{q,R} \Pi(v,q,R)$ under constraints $u(q,R) \geqslant v$ and $R \leqslant L$ with $\Pi(v,q,R) \equiv D\,(u(q,R)-v)\,\pi(q,R)$. Lastly, we relate the equilibrium of the insurer competition for a given θ to $\hat{P}(L,v^\theta)$ where v^θ is the equilibrium utility of consumers. ■

Lemma 2. Along the contract curve the consumer utility increases as R increases and q decreases.

Proof. Observe that $u(q,R)$ is bounded over $[q^a;q^f] \times [L-w;L]$ and spans an interval $[v_1;v_2]$. The assumption that β is very small implies $u_R < 0$ thus the solution $R = \rho(q,v)$ to the equation $v = u(q,R)$ is also bounded over $[q^a;q^f] \times [v_1;v_2]$. Let us now introduce

$$\Phi^q(q,R) \equiv -\frac{\pi_q}{u_q} = \frac{-R\,(R+F)}{(R-c^f(q))^2}\frac{c_q^f(q)}{c_q^a(q)}, \tag{20}$$

$$\Phi^R(q,R) \equiv -\frac{\pi_R}{u_R} = \frac{1-\beta R\dfrac{R-2c^f(q)}{(R-c^f(q))^2}}{1-\beta+\beta\dfrac{Fc^a(q)}{(R+F)^2}}. \tag{21}$$

The solution $q^*(v)$ to $\frac{\pi_R}{\pi_q} = \frac{u_R}{u_q}$ evaluated at $R = \rho(q,v)$ also solves $\Phi^q\,(q,\rho(q,v)) =$

$\Phi^R(q,\rho(q,v))$. We want to show that $q^*(v)$ is increasing. Then it will be clear that $R^*(v) \equiv \rho(q^*(v),v)$ is decreasing since ρ_q and ρ_v are negative. This will indirectly prove that the relationship between R and q on the contract curve is negative and also that the consumer utility v increases when R decreases (as $R^v \searrow$).

Observe that for any v in $[\underline{v};\bar{v}]$, $\Phi^R(q,\rho(q,v))$ is bounded positive over $\left[q^a;q^f\right]$ with a lower bound $\underline{\Phi}^R$ while $\Phi^q(q,\rho(q,v))$ varies from $+\infty$ to 0 over $\left[q^a;q^f\right]$. Hence the equation has a solution; we assume that it is unique. To obtain q_v^* we differentiate the equation $\Phi^q - \Phi^R = 0$ using $\rho_q = -\dfrac{u_q}{u_R} < 0$ and $\rho_v = \dfrac{1}{u_R} < 0$.

$$\begin{aligned}
0 &= \left(\Phi_q^q - \Phi_q^R\right) q_v^* + \left(\Phi_R^q - \Phi_R^R\right)\left(\rho_q q_v^* + \rho_v\right)\\
&= \left(\Phi_q^q - \Phi_q^R\right) q_v^* + \left(\Phi_R^q - \Phi_R^R\right)\frac{1 - q_v^* u_q}{u_R}\\
\Rightarrow q_v^* &\left(\left(\Phi_q^q - \Phi_q^R\right) - \left(\Phi_R^q - \Phi_R^R\right)\frac{u_q}{u_R}\right) = \frac{\Phi_R^R - \Phi_R^q}{u_R}\\
\Rightarrow q_v^* &= \frac{\Phi_R^R - \Phi_R^q}{\left(\Phi_q^q - \Phi_q^R\right)u_R - \left(\Phi_R^q - \Phi_R^R\right)u_q}.
\end{aligned} \tag{22}$$

Defining $X \equiv R\dfrac{R-2c^f}{\left(R-c^f\right)^2}$ and $Y \equiv \dfrac{Fc^a}{(R+F)^2} - 1$, we have $\Phi^R = \dfrac{1-\beta X}{1-\beta Y}$ and derive

$$\Phi_R^R = \beta\frac{X_R(1-\beta Y) - Y_R(1-\beta X)}{(1-\beta Y)^2} \quad \text{and} \quad \Phi_q^R = \beta\frac{X_q(1-\beta Y) - Y_q(1-\beta X)}{(1-\beta Y)^2} < 0.$$

Observe that

$$\Phi_q^q = \Phi^q\left(\frac{c_{qq}^f}{c_q^f} - \frac{c_{qq}^a}{c_q^a} + \frac{2c_q^f}{R-c^f}\right) < 0 \quad \text{and} \quad \Phi_R^q = \frac{2R+F}{(R-c^f)^3}\frac{c^f c_q^f}{c_q^a} < 0.$$

When $\beta \ll 1$, Φ_R^R and Φ_q^R are second order with respect to Φ_R^q and Φ_q^q . As $u_R \gg u_q$ we obtain

$$\begin{aligned}
\lim_{\beta\ll 1} q_v^* = \frac{-\Phi_R^q}{\Phi_q^q u_R} &= \frac{-\Phi_R^q}{u_R\Phi^q\left(\dfrac{c_{qq}^f}{c_q^f} - \dfrac{c_{qq}^a}{c_q^a} + \dfrac{2c_q^f}{R-c^f}\right)}\\
&= \frac{(2R+F)\,c^f c_q^f}{(R-c^f)^3\pi_R c_q^a}\frac{1}{\dfrac{c_{qq}^f}{c_q^f} - \dfrac{c_{qq}^a}{c_q^a} + \dfrac{2c_q^f}{R-c^f}} > 0,
\end{aligned} \tag{23}$$

since $c_q^f > 0$, $c_q^a < 0$ and both cost functions are convex. By continuity, there exits a level $\bar{\beta}$ such that $q_v^* > 0$ for all R and q in their respective range and $\beta < \bar{\beta}$. ■

Lemma 3. There exists $\underline{w}$ and $\bar{w}$ such that when $w \in [\underline{w}\,;\bar{w}]$, $\exists\, v^L \in [\underline{v};\bar{v}]$ and the solution of $\hat{P}(L,v)$ is
- $R^v = L$ and $q^v \searrow$ for $v \in \left[\underline{v}; v^L\right]$.
- $R^v \searrow$ and $q^v \nearrow$ for $v \in \left[v^L;\bar{v}\right]$.

Proof. As $u_q < 0$, the equation $u(q,R) = v$ has a unique solution $q = Q(R,v)$ satisfying $Q_R = -\dfrac{u_R}{u_q} < 0$ and $Q_v = \dfrac{1}{u_q} < 0$. Solving $\hat{P}(L,v)$ is equivalent to maximize $\Pi\left(v, Q(R,v), R\right)$ over $[L-w;L]$. The FOC for an interior solution is

$$\left.\frac{u_q}{u_R}\right|_{q=Q(R,v)} = \left.\frac{\pi_q}{\pi_R}\right|_{q=Q(R,v)}. \tag{24}$$

By lemma 2, the solution $R^*(v)$ of (24) is decreasing and $Q\left(R^*(v),v\right)$ is increasing. Furthermore the properties $\pi_q(q^f) = 0$ and $u_q(q^a) = 0$ imply that $Q\left(R^*(v),v\right)$ is always interior to $\left[q^a;q^f\right]$. Lastly, the inequality $\pi(.,L-w) < 0$ implies that only the constraint $R \leqslant L$ can be binding.

The solution to $\hat{P}(L,v)$ is now very simple to obtain: starting from the largest level $\bar{v}$, one decreases v. As long as $v \geqslant v^L \equiv R^{*-1}(L)$, the contract $(R^v,q^v) \equiv \left(R^*(v), Q\left(R^*(v),v\right)\right)$ is optimal i.e., the price increases while the quality decreases. For $v = v^L$, the optimal pair is simply $(L, Q(L,v))$ and as v decreases further the quality now increases (direct effect only).[21] To sum up, the solution of $\hat{P}(L,v)$ is $(L, \min\{q^\pi, Q(L,v)\})$if $v < v^L$ and (R^v,q^v) otherwise.

We need to assess the conditions under which $v^L \in [\underline{v};\bar{v}]$, i.e., when the liquidity constraint binds for the monopoly but not at the fully competitive outcome. The monopoly maximizes π under the constraint $u \geqslant \underline{v}$. He starts from the utility level $\bar{v}$ and moves down along the contract curve (R^v,q^v). If $u\left(q^L,L\right) > \underline{v}$ then the monopoly chooses L and increases quality up to $\underline{q} = \min\{q^\pi, Q(L,\underline{v})\} > q^L$ and this means $v^L > \underline{v}$. At the other extreme of the competition spectrum, if $\pi(q^L,L) > 0$ then the solution of $\hat{P}(L,\,\bar{v})$ features $R < L$ i.e., $v^L < \bar{v}$. The two conditions we have characterized are $u\left(q^L,L\right) = V - \beta\dfrac{L}{L+F}c^a(q^L) > \underline{v} = (1-\beta)\,V + w \Leftrightarrow w < \bar{w} \equiv \beta\left(V - \dfrac{L}{L+F}c^a(q^L)\right)$ and $\pi(q^L,L) > 0 \Leftrightarrow w >$

[21] The candidate value $Q(L,v)$ hits the upper bound q^f if $(V-\underline{v})\dfrac{L+F}{\beta L} \geqslant c^a(q^f) \Leftrightarrow V \geqslant V^* \equiv c^a(q^f)\dfrac{L}{L+F} + \dfrac{w}{\beta}$.

$\underline{w} \equiv \beta L \dfrac{L}{L - c^f(q^L)}$. Thanks to $H1$, $\underline{w} < \bar{w}$ holds, thus $v^L \in \left]\underline{v};\bar{v}\right[$ whenever $w \in \left]\underline{w}\,;\bar{w}\right[$. Obviously if $w \notin \left]\underline{w}\,;\bar{w}\right[$, only one regime applies. Equivalently we may define $v^L \equiv \min\left\{\bar{v}, \max\left\{R^{*-1}(L), \underline{v}\right\}\right\}$. ■

Lemma 4. $\exists\, \theta^L$ such that the equilibrium strategies are
- $R^\theta = L$ and $q^\theta \searrow$ for $\theta \in \left[0;\theta^L\right]$,
- $R^\theta \searrow$ and $q^\theta \nearrow$ for $\theta \in \left[\theta^L;+\infty\right]$.

As in the previous lemma, we solve the problem without constraints and later introduce them.

Step 1. Solve the program $\tilde{P}(\theta) \equiv \max_{v,q} D(v - v^\theta)\, \pi\left(q, \rho(q,v)\right)$.

The FOC of $\tilde{P}(\theta)$ with respect to q for an interior solution is $\dfrac{\pi_R}{\pi_q} = \dfrac{u_R}{u_q}$, thus the optimal quality is $q^*(v)$. The FOC of $\tilde{P}(\theta)$ with respect to v is $0 = \rho_v \dfrac{\partial \Pi_i}{\partial R} = \dfrac{1}{u_R}\left(t\pi + \dfrac{1}{n}\dfrac{\pi_R}{u_R}\right)$ and replacing q by the optimal value $q^*(v)$ we obtain a unique equation[22]

$$\theta = H(v) \equiv \left.\frac{-\pi_R}{\pi .\, u_R}\right|_{\substack{R=R^*(v)\\ q=Q(R^*(v),v)}} = \frac{\Phi^R\left(q^*(v), \rho\left(q^*(v), v\right)\right)}{\pi\left(q^*(v), \rho\left(q^*(v), v\right)\right)}. \tag{25}$$

To show that v^θ is increasing we differentiate (25) to get

$$H' > 0 \Leftrightarrow (1 - u_q q_v^*)\left(\frac{\pi}{u_R}\Phi_R^R + (\Phi^R)^2\right) > \left(\Phi^R \pi_q - \pi \Phi_q^R\right) q_v^*. \tag{26}$$

As Φ_R^R and u_q are multiple of β while Φ^R is not (cf. Lemma 2), the LHS of (26) tends to $\left(\Phi^R\right)^2$ for $\beta \ll 1$, which is bounded away from zero (see definition (21)) while the RHS of (26) tends to zero because both π_q and Φ_q^R are multiple of β.

Step 2. Solve $P(L, v^\theta) = \max_{q,R} \Pi_i(v^\theta, q, R)$ such that $u(q,R) \geqslant \underline{v}$, $R \leqslant L$.

We define $\theta^L \equiv H(v^L)$ for $v^L = R^{*-1}(L)$, $\theta^L = 0$ if $v^L = \underline{v}$ and $\theta^L = +\infty$ if $v^L = \bar{v}$. When $\theta \geqslant \theta^L$, $v^\theta = H^{-1}(\theta) \geqslant v^L \geqslant \underline{v}$, thus $R^\theta = R^*(v^\theta) \leqslant L$ and $q^\theta = Q(R^\theta, v^\theta)$ solve $P(L, v^\theta)$. Yet when $\theta < \theta^L$, $R^*\left(H^{-1}(\theta)\right) > L$ means that all firms want to bind the liquidity constraint $R \leqslant L$. The competition then takes place

[22] We do not need a fixed point argument to derive the equilibrium because the solution of $\tilde{P}(\theta)$ does not depend on the level v^θ. This in turn is due to the linearity of D.

over a single variable, the audit quality varying in $\left[Q\left(R^*(v^L), v^L\right); q^f\right]$. Since the participation constraint $u(q,L) \geqslant \underline{v}$ is equivalent to $q \leqslant Q(L,\underline{v})$, the correct upper bound is thus $min\left\{Q(L,\underline{v}), q^f\right\}$.

The profit function is $\hat{\Pi}_i(q_i) \equiv D\left(u(q_i,L) - u(\hat{q},L)\right)\pi(q_i,L)$ where $\hat{q}$ is the equilibrium quality. The unique FOC to be satisfied at the symmetric equilibrium is similar but different from (25):

$$\theta = G_L(q) \equiv \left.\frac{-\pi_q}{\pi.\,u_q}\right|_{R=L}. \tag{27}$$

Since $\Phi^q_q < 0$ and $\pi_q > 0$ we have $\Phi^q_q\pi > \Phi^q\pi_q \Leftrightarrow G'_L(v) < 0$. The candidate Nash equilibrium quality is $G_L^{-1}(\theta)$; it varies from $G_L^{-1}(0) = q^f$ to $G_L^{-1}(+\infty) = q^a$. The equilibrium is thus $\min\left\{G_L^{-1}(\theta), Q(L,\underline{v})\right\}$. To summarize, the symmetric equilibrium is $\left(R^\theta, q^\theta\right) = \left(L, \min\left\{G_L^{-1}(\theta), Q(L,\underline{v})\right\}\right)$ if $\theta < \theta^L$ and $\left(R^\theta, q^\theta\right) = \left(R^*(H^{-1}(\theta)), Q\left(R^*(H^{-1}(\theta)), H^{-1}(\theta)\right)\right)$ otherwise. ■

Proof of Corollary 1

The frequency of fraud increases with competition.

The frequency of fraud in the symmetric PBE is $\sigma^\theta = \dfrac{\beta c^f(q^\theta)}{(1-\beta)\left(R^\theta - c^f(q^\theta)\right)}$ thus $\dfrac{\partial\sigma^\theta}{\partial\theta} > 0 \Leftrightarrow \dfrac{c^f_q(q^\theta)}{c^f(q^\theta)}\dfrac{\partial q^\theta}{\partial\theta}R^\theta > \dfrac{\partial R^\theta}{\partial\theta}$.

When $\theta < \theta^L$, the equality $R^\theta = L$ implies that $\dfrac{\partial\sigma^\theta}{\partial\theta} > 0 \Leftrightarrow \dfrac{c^f_q(q^\theta)}{c^f(q^\theta)}\dfrac{\partial q^\theta}{\partial\theta}R^\theta > 0$ which holds true since sign $\left(\dfrac{\partial q^\theta}{\partial\theta}\right)$ is equal to $\text{sign}(G_L^{'}) = \text{sign}(c^f_q) = -1$.

When $\theta \geqslant \theta^L$, $v^\theta = H^{-1}(\theta)$, $R^\theta = R^*(v^\theta)$ and $q^\theta = Q\left(R^\theta, v^\theta\right)$ thus $\dfrac{\partial q^\theta}{\partial\theta} = Q_v\dfrac{\partial v^\theta}{\partial\theta} + Q_R\dfrac{\partial R^\theta}{\partial\theta} = \dfrac{1}{u_q}\left(\dfrac{1}{H'} - u_R R^*_v\right)$ and $\dfrac{\partial\sigma^\theta}{\partial\theta} > 0 \Leftrightarrow \dfrac{c^f_q(q^\theta)}{c^f(q^\theta)}\dfrac{R^\theta}{u_q}\left(\dfrac{1}{H'} - u_R R^*_v\right) > R^*_v$. Using $\lim\limits_{\beta\ll 1} q^*_v$ (equation (23)) we finally obtain

$$\frac{\partial\sigma^\theta}{\partial\theta} > 0 \Leftrightarrow \frac{c^\pi_q}{\dfrac{c^f_{qq}}{c^f_q} - \dfrac{c^a_{qq}}{c^a_q} + \dfrac{2c^f_q}{R^\theta - c^f}} < \frac{1}{\frac{2}{R^\theta - c^f}},$$

which is satisfied since convexity of the cost functions implies $\dfrac{c^f_{qq}}{c^f_q} - \dfrac{c^a_{qq}}{c^a_q} > 0$. ■

Proof of Proposition 4

If there are no transaction cost for the creation of an audit agency, it is created only in the weak competition regime ($\theta \leqslant \theta^L$). The common audit quality is $\tilde{q}(\theta)$ is greater than the equilibrium value q^θ and equal to the insurers ideal level for very low levels of competitiveness.

From the argument in the text, it is enough to consider the case $\bar{q} \in (q^\theta, q^f]$ since $\bar{q} < q^\theta$ is a non-credible choice for the industry. Now we claim that in equilibrium, $q = \bar{q}$ for if firms were to choose $\hat{q} > \bar{q}$, this value would have to be part of an equilibrium without the agency and this would be a contradiction. Hence, firms are constrained on quality and compete on R only.

The FOC characterizing an interior equilibrium ($R < L$) is $\theta = J_{\bar{q}}(R) \equiv \dfrac{\Phi^R(\bar{q}, R)}{\pi(\bar{q}, R)}$. Using the fact that $\beta \ll 1$ and the arguments of Lemma 2, we obtain $J'_{\bar{q}} < 0$ hence the equilibrium is $\bar{R}_\theta(\bar{q}) \equiv \min\left\{L, J_{\bar{q}}^{-1}(\theta)\right\}$. If we increase $\bar{q}$ then the whole curve $J_{\bar{q}}$ falls because $\Phi^R_q < 0$ and $\pi_q > 0$; hence $\bar{R}'_\theta < 0$ so that the $\bar{R}_\theta$ curve is decreasing in the (R, q) plan.

The contract curve where $\Phi^R = \Phi^q$ defines a decreasing function $\tilde{R}(q)$.

We plot $\bar{R}_\theta(\bar{q}))$ and $\tilde{R}(\bar{q})$ on Figure A1; they intersect at $\bar{q} = q^\theta$. To prove that $\bar{R}_\theta$ is steeper than $\tilde{R}$ consider $\bar{q} = q_2 > q^\theta$ and $\tilde{R}(\bar{q})$ (point a on Figure A1). As $H' > 0$ (cf. equation (26)) it must be the case that $H > \theta$ at this point a or, in other words, that $\theta < J_{\bar{q}}(\tilde{R}(\bar{q}))$ so that we must increase R to get an equality.

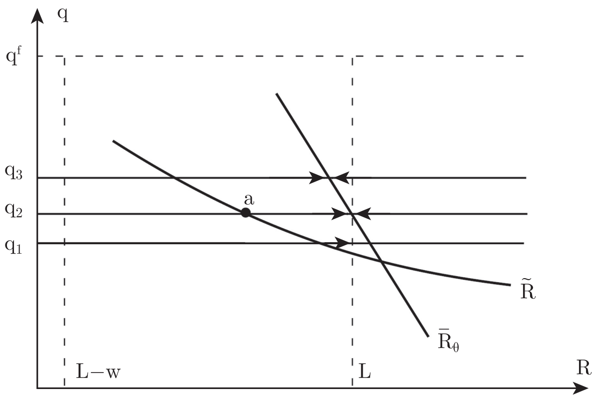

Figure A1 The Equilibrium with an Audit Agency

When $\theta > \theta^L$, any $\bar{q} > q^\theta$ leads to an interior equilibrium $\bar{R}_\theta(\bar{q}) < L$ satisfying $\dfrac{\partial \bar{R}_\theta}{\partial \bar{q}} = \dfrac{\Phi_q^R - \theta\pi_q}{\theta\pi_R - \Phi_R^R}$. Firms' profits are $\bar{\pi} = \dfrac{1}{n}\pi\left(\bar{q}, \bar{R}_\theta(\bar{q})\right)$ thus $\dfrac{d\bar{\pi}}{d\bar{q}} = \dfrac{1}{n}\left(\pi_q + \pi_R \dfrac{\partial \bar{R}}{\partial \bar{q}}\right) = \dfrac{1}{n}\dfrac{\pi_R \Phi_q^R - \pi_q \Phi_R^R}{\theta\pi_R - \Phi_R^R} = \dfrac{\beta}{n}\dfrac{(X - \beta Y)}{\theta\pi_R - \beta Z}$,

where none of X, Y and Z are multiple of β (cf. Lemma 2); hence $\lim_{\beta \ll 1} sign\left(\dfrac{d\bar{\pi}}{d\bar{q}}\right) = sign(\Phi_q^R) < 0$. This means that the industry will not create the common agency when there is a high degree of competitiveness.

When $\theta \leqslant \theta^L$, we see on Figure A1 that if the common agency chooses $\bar{q} = q_3$ then the equilibrium is interior because the quality is so high that firms find it profitable (in the competition) to relax the liquidity constraint. By the argument of the previous paragraph, a better choice for the industry is to reduce the common agency quality $\bar{q}$. At $\bar{q} = q_1$ on the other hand, the equilibrium is constrained at $R = L$ and increasing $\bar{q}$ is desirable for the insurance industry (it increases equilibrium profits). The limit is $\bar{q} = q_2$ such that $\bar{R}_\theta(\bar{q}) = L \Leftrightarrow \theta = J_{\bar{q}}(L)$, it implicitly defines a decreasing function $\tilde{q}(\theta)$ (because the curve $J_{\bar{q}}$ falls when $\bar{q}$ increases). As $\Phi^R > 0$, $\theta = J_{\bar{q}}(L)$ is impossible for very low levels of competitiveness, thus the maximal quality q^f is optimal for the industry when $\theta < \underline{\theta}$ defined by $\underline{\theta} = J_{q^f}(L)$. ■

Proof of Proposition 5

The introduction of a replacement clause increases audit quality and premiums.

We shall treat the difference ϵ between the avoided opportunity cost of consumers and the additional cost of insurers as a constant. We thus only have to show that $dq/d\delta$ and $dR/d\delta$ are positive.

Observe that $\pi_\delta < 0, \pi_{R\delta} > 0, u_{R\delta} > 0, \pi_{q\delta} > 0$ and $u_{q\delta} = 0$ hence $\Phi^R = -\dfrac{\pi_R}{u_R}$ and $\Phi^q = -\dfrac{\pi_q}{u_q}$ increase with δ. Differentiating $\Phi^R = \Phi^q$ with respect to δ keeping R constant we get $q_\delta|_{R=cte} = \dfrac{\Phi_\delta^q - \Phi_\delta^R}{\Phi_q^R - \Phi_q^q}$. It is then easy to obtain $\lim_{\beta \ll 1} q_\delta|_{R=cte} = -\Phi_\delta^q/\Phi_q^q > 0$ as $\Phi_\delta^q > 0$. Likewise $R_\delta|_{q=cte} = \dfrac{\Phi_\delta^q - \Phi_\delta^R}{\Phi_R^R - \Phi_R^q}$ and $\lim_{\beta \ll 1} R_\delta|_{q=cte} = -\Phi_\delta^q/\Phi_R^q > 0$. However these are local variations that do not account for the global change in the equilibrium contract induced by the replacement clause. The system characterizing the equilibrium is $\Phi^R = \Phi^q = \theta\pi$ hence

$$\left\{\begin{array}{c} \Phi_q^R q_\delta + \Phi_R^R R_\delta + \Phi_\delta^R = \Phi_q^q q_\delta + \Phi_R^q R_\delta + \Phi_\delta^q \\ \Phi_q^R q_\delta + \Phi_R^R R_\delta + \Phi_\delta^R = \theta\left(\pi_q q_\delta + \pi_R R_\delta + \pi_\delta\right) \end{array}\right.$$

$$\Rightarrow \left\{\begin{array}{c} \left(\Phi_q^R - \Phi_q^q\right) q_\delta + \left(\Phi_R^R - \Phi_R^q\right) R_\delta = \Phi_\delta^q - \Phi_\delta^R \\ \left(\Phi_q^R - \theta\pi_q\right) q_\delta + \left(\Phi_R^R - \theta\pi_R\right) R_\delta = \theta\pi_\delta - \Phi_\delta^R \end{array}\right.$$

$$\Rightarrow R_\delta = \frac{\left(\Phi_\delta^q - \Phi_\delta^R\right)\left(\Phi_q^q - \theta\pi_q\right) - \left(\theta\pi_\delta - \Phi_\delta^q\right)\left(\Phi_q^R - \Phi_q^q\right)}{\left(\Phi_q^q - \theta\pi_q\right)\left(\Phi_R^R - \Phi_R^q\right) - \left(\Phi_q^R - \Phi_q^q\right)\left(\Phi_R^q - \theta\pi_R\right)},$$

and

$$q_\delta = \frac{\left(\Phi_R^R - \theta\pi_R\right)\left(\Phi_\delta^q - \Phi_\delta^R\right) - \left(\Phi_R^R - \Phi_R^q\right)\left(\theta\pi_\delta - \Phi_\delta^R\right)}{\left(\Phi_R^R - \theta\pi_R\right)\left(\Phi_q^R - \Phi_q^q\right) - \left(\Phi_R^R - \Phi_R^q\right)\left(\Phi_q^R - \theta\pi_q\right)}.$$

We now use the fact that Φ_R^R, $\Phi_q^R, \pi_\delta, \pi_q$ and Φ_δ^R are multiple of β to get $\lim_{\beta \ll 1} q_\delta = -\Phi_\delta^q / \Phi_q^q$ which is positive and $\lim_{\beta \ll 1} R_\delta = \dfrac{\Phi_q^q \pi_\delta - \Phi_\delta^q \pi_q}{-\Phi_q^q \pi_R}$; thus $\lim_{\beta \ll 1} R_\delta > 0 \Leftrightarrow \Phi_q^q \pi_\delta > \Phi_\delta^q \pi_q$

$$\Leftrightarrow \Phi^q \left(\frac{c_{qq}^f}{c_q^f} - \frac{c_{qq}^a}{c_q^a} + \frac{2c_q^f}{R - c^f}\right) \frac{-\beta R^2}{\left(R - c^f\right)^2} > \Phi^q \left(\frac{2}{R - c^f}\right) \frac{-\beta R^2 c_q^f}{\left(R - c^f\right)^2}$$

$$\Leftrightarrow \frac{c_{qq}^a}{c_q^a} - \frac{c_{qq}^f}{c_q^f} > 0,$$

where the last inequality has been established in the proof of lemma 2. ■