Commissions

Roman Inderst* Marco Ottaviani†

October 2008

VERY PRELIMINARY

Abstract

Commissions to third-part agents are pervasive, but have recently received a bad reputation in many industries, including health care, insurance products, or retail financial products such as mortgages. This paper introduces a simple framework to study the various roles that commissions can perform, as well as the implications of different policy alternatives such as mandatory disclosure, increased liability for agents, or caps on commissions or product prices.

*University of Frankfurt, IMFS. Address: Roman Inderst, Johann Wolfgang Goethe-University, Chair of Finance and Economics, Mertonstrasse 17, 60054 Frankfurt am Main, Germany. E-mail: inderst@finance.uni-frankfurt.de.

†Northwestern.

1 Introduction

Firms frequently have to rely on third parties to sell goods to final consumers. In this paper, we are interested in those cases where the chosen agents have also the task of advising customers on whether a particular product is suitable for their needs. Throughout this paper, we will mainly relate our findings to the financial and insurance industry as well as to health care. There, commissions and "kick-backs" that are given to agents more or less openly have recently come under increasing scrutiny.

High-powered incentives paid to supposedly independent insurance agents, most notably by AIG, were targeted by the former New York prosecutor Elliot Spitzer. More recently, commissions earned by brokers of (subprime) mortgages have been identified as one of the main culprits for why business was potentially expanded too aggressively, thereby inducing ("irresponsible" or "predatory") lending to households against their best interest. In addition, households may have been ill-advised in their choice among the skyrocketing number of different mortgage innovations. Also, the UK's experience with deregulation over the last decades has lead a number of cases, including the sale of private pension schemes or, once again, various forms of mortgages, where independent financial advisors or other agents were allegedly incentivized too aggressively, leading to "misselling". While monetary incentives by pharmaceutical companies to GPs or pharmacists are often forbidden, their advice may be influenced by the "soft bribes" that are provided, for instance, through "training courses" in expensive resorts.[1]

This paper develops a simple modelling framework that allows to analyze under different scenarios the role of commissions, as well as the implications of various policies. In our model, commissions can ultimately perform three different roles: i) to (unilaterally) steer an agent's advice and thus sales towards a firm's own product; ii) to provide a countevailing force against a similar attempt by a rival firm whose product is also catered for by the same (common) agent; and iii) to incentivize additional services by the agent, most notably the acquisition of interested customers in the first place. Our main policy options are, first, the imposition of a higher liability standard for "unsuitable selling" and, second, the requirement to make commissions transparent to customers. In addition, we

[1] Another way of influencing the agent, which we are not modeling in this paper, is through providing him with information, which is, for instance, the case in the so-called "detailing process", when GPs are visited by the sales representatives of pharmaceutical companies.

explore the imposition of caps on commissions or product prices. Besides in the health care market, price caps may also apply, though in less immediate ways, with tax-advantaged savings products.[2]

The disclosure of commissions is increasingly required for the sale of financial products to retail customers. In the European Union, this is a requirement since January 2008 (under MiFID). The UK's Financial Service Authority had already earlier imposed a similar obligation. Next, the imposition or strenghtening of a liability requirement for an independent advisor could be administered in several ways. For instance, the agent's sales could be more closely supervised or, likewise, penalties in case of alleged "misselling" could be raised (including the loss of the agent's licence). To be specific, our model presumes a regime of occupational licensing that requires agents to post a "surety bond". By raising this bond, more of the agent's wealth is put at risk in case of alleged "misselling".[3]

In a nutshell, a key result of this paper is that a policy that makes the disclosure of commissions both mandatory and credible may, holding all else constant, not necessarily lead to a more efficient market outcome. While it will typically lead to smaller commissions, this may not necessarily be beneficial. In particular, this could stifle the provision of other services that are performed by the agent and it could shift sales in the "wrong" direction, e.g., by inefficiently expanding the sales of a less efficient firm. A cap on commissions as well as increased liability for the agent may have the same unintended consequences.

In the case of a common agency, the underprovision of non-contractible and not firm-specific services, such as the acquisition of interested customers, may arise from a free-rider problem. Higher commissions under a regime where disclosure is not mandatory (and not credible) may mitigate this inefficiency. But even when we abstract from the (under-)provision of such services, we find that disclosure can still lead to a less efficient outcome. Here, a key starting observation is that the outcome under disclosure is typically not efficient. The intuition for this is that steering the agent's advice through commissions is costly to firms. This holds even more so as a higher commission is paid not only on the "marginal" sale, but also on "inframarginal" sales.[4] In particular, differences in market shares, which depend on the agent's advice, may then not adequately reflect differences

[2] More generally, price caps may be imposed by competition authorities.

[3] For instance, Pahl and Tod (2007) document the widespread use (and cross-state variation) of "surety bonds", as well as other licensing provisions, for the US mortgage industry.

[4] Strictly speaking, in our model a single sale will be made with lower or higher probability. This allows to abstract from the design of nonlinear incentive schemes.

in firms' costs. We derive conditions for when the resulting inefficiency is lower in case commissions are not disclosed, given that the more efficient firm may then have more incentives to further raise its commission than its less efficient rivals. We conduct a similar comparative analysis with respect to the other considered policy variables.

Most prominently, this paper relates to the literature on expert advice. This literature, following Darby and Karni (1973), is surveyed in Dulleck and Kerschbamer (2006). Bolton et al. (2007) provide an application to the case of financial products. The novelty of our setting is that the expert's preferences and thus his advice are influenced by the commissions that he obtains from firms. These firms also determine the pricing of the products.

The expert in our model is an independent agent, who is directly subject to possible penalties (such as the loss of a "bond") in case of providing (allegedly) unsuitable advice. The expert's advice concerns, as in Bolton et al. (2007), the choice between two products. In the single-agency case, only one product is "marketed" through commissions, while in the common-agency case both products are "marketed" through commissions. In Inderst and Ottaviani (2008), instead, a firm sells a single product through an employee, whose compensation package induces him to both locate customers and to provide suitable advice. We also consider such a double-task problem in an extension. In the presently analyzed model, the crucial difference is that with multiple products the benefits from locating a customer are not fully extracted by a single firm.

In our model, the agent always performs for firms the service of advice, which will increase customers' willingness-to-pay and, consequently, the maximum price that firms can extract. In an extension, the agent also performs additional services, such as the acquisition of the customer. This relates the present analysis to various other strands of the literature. In Industrial Organization, there is a literature that analyzes downstream moral hazard, typically in a manufacturer-retailer or franchisor-franchisee relationship, with respect to services such as "promotional effort" (cf. Tirole (2000), chapter 4.2.2.). The case analyzed in this paper where such services are performed under common agency relates to the broader literature studying multi-task settings with common agency (cf. the surveys in Dixit (1996) and Tirole (2001)). Finally, our analysis of optimal commissions relates to the large compensation literature, most notably the literature on the design of compensation schemes for sales representatives (cf. Basu et al. (1985)). The focus there is, however, typically on the classic trade-off between risk-sharing and incentives

The rest of this paper is organized as follows. Section 2 analyzes the basic case where only one product is actively "marketed" by its (monopolistic) provider. Section 3 discusses various policy options in this setting. In Section 4 we extend the analysis to the case of "common agency" where the agent's advice is affected by commissions on different products. Section 5 considers the possibility that the (common) agent can perform additional and potentially non-firm specific services. Section 6 provides a further discussion of some of the modelling assumption. Section 7 offers some concluding remarks.

2 The Basic (Single-Agency) Case

2.1 The Model

In our basic model, there are two products $n = a, b$. A customer purchases at most one product and only one unit of it. His valuation depends on a binary state variable, which for simplicity we also denote by $\theta = a, b$. The customer's utility is equal to v_h if the product "matches" the state, while his utility is $v_l < v_h$ if no such match is obtained. While the customer only has the ("coarse") prior beliefs $\overline{q} = \Pr(\theta = a)$, the agent can observe a more informative signal, which gives rise to the posterior beliefs $q = \Pr(\theta = a)$. These have the (ex-ante) distribution $G(q)$ with density $g(q) > 0$ over $q \in [0, 1]$. The observation of q is the agent's private information.[5]

In our basic model, one product, namely $n = b$, is not actively marketed to the agent, while for product $n = a$ the respective producer, to whom we also refer to as $n = a$, may try to affect the agent's recommendation through paying a commission. This may be a suitable description for many situations where an agent earns nothing or a relatively small fee from recommending a "standard" product or treatment, while he can earn an additional fee or "kick-back" when recommending a "novel" product or treatment. Also, this initial asymmetry provides the most simple case in which to isolate the various effects. Section 4 extends the model to common agency.

[5] This information structure could easily be further endogenized by supposing that the agent observes some signal $s \in [\underline{s}, \overline{s}]$, which in turn is obtained from some state-dependent signal-generating function: $\Psi_\theta(s)$, where $\Psi_a(s)$ dominates $\Psi_b(s)$ in the sense of the Monotone Likelihood Ratio Property. The conditional beliefs $q(s)$ would then be strictly decreasing, with $q(\underline{s}) = 0$ in case $\psi_b(\underline{s}) > \psi_a(\underline{s}) = 0$ and $q(\underline{s}) = 1$ in case $\psi_a(\overline{s}) > \psi_b(\overline{s}) = 0$. A change of probability measure finally allows to calculate expressions directly in terms of the probability distribution over the resulting conditional probability $q \in [0, 1]$.

To influence the agent's advice and, thereby, the buyer's decision, firm $n = a$ may thus specify in the first period of our model, $t = 1$, a commission or fee f. Whether this is observed by customers or not will be one of our key distinctions (and thus policy variables). For the benchmark analysis we suppose that f is not observed by customers. In fact, without policy interference it may simply not be credible for the firm to commit to a given incentive level for the agent.

Firm a also controls the price of its product. (However, below we will also look at the policy option of regulating prices.) The respective price p_a is set in $t = 2$.[6] The cost of product a is given by c_a. For product b, which is not actively "marketed" by its providers, we suppose, for specificity, that it is provided competitively, i.e., at the respective marginal cost: $p_b = c_b$.

In our benchmark case, we suppose that at $t = 3$ a customer arrives with probability one. Hence, the agent needs not exert effort to contact a customer. Likewise, he needs not exert effort to generate the additional information q. Consequently, for now the commission f serves only the role of steering the agent towards product a. The role of inducing the provision of additional services is, instead, analyzed in Section 5. It is also in $t = 3$ that the agent advises the customer on which product to buy. The customer can then follow the agent's advice or choose not to follow it, purchasing none or either one of the two products, a and b.

The provision of advice in $t = 3$ represents a standard game of "cheap talk", in the line of Crawford and Sobel (1982). It is well known that such a game typically has multiple equilibria. More specifically, in our setting there will be an informative equilibrium where the agent's recommendation allows the customer to update his beliefs, as well as an uninformative ("babbling") equilibrium. In what follows, we always focus on the informative equilibrium.

Note also at this point that the agent fully controls the access to customers. Hence, a firm can not bypass the agent through "direct marketing" in order to reach these customers. The agent also does not take ownership of the products. Taken together, these assumptions should apply to the key motivating examples in the Introduction: independent advice on retail financial and insurance products, on the hand, and health care, on the other hand.

[6] The assumption that f is set before p_a captures the notion that f is determined, possibly uniformly for a large number of agents, by some (long-term) negotiated contract.

When recommending the purchase of either product a or b, the agent cares not only about possibly earning a commission. In case his advice turns out to be incorrect, i.e., leading to a "mismatch" and thus to low utility v_l, the agent incurs the costs $w > 0$. As noted in the Introduction, this could represent the minimum net worth (or a surety bond) that is imposed by regulation (cf. also Section 3 for a more detailed discussion).[7] More generally, w may also comprise costs from a loss of reputation or costs that are imposed on the firm by the intervention of regulators or from legal prosecution.[8] For the moment, w may be either a transfer to the regulator or dead-weight loss. Below, we also allow that all or a fraction of w represents compensation for the customer in case v_l is realized. While this complicates expressions, it does not alter results.

Before proceeding to the analysis of the basic model, the restriction to the simple compensation scheme between the agent and firm a deserves some comments. In Section 6, we allow firms to make commissions contingent on whether the agent will subsequently be forced to pay w in case of a resulting "mismatch" or whether this is not the case. In case producers are prevented from "compensating" the agent for all or a fraction of w in case of a "mismatch", e.g., either by an outright prohibition or by the fear to become "vicariously liable" for unsuitable selling, the restriction to such simple commissions, f, is without loss of generality.

We conclude the description of the basic model with the derivation of the efficient benchmark for advice. For this denote $v_a(q) := qv_h + [1-q]\, v_l$ and $v_b(q) := [1-q]\, v_h + qv_l$. Given the information q, the relative net surplus from purchasing product a instead of product b is given by

$$\Delta(q) := [v_a(q) - c_a] - [v_b(q) - c_b]\,,$$

which is strictly increasing in q. Depending on c_n, this could be strictly positive or negative for all q. To make the following problem interesting, we thus suppose that there is an interior cutoff $0 < q_{FB} < 1$ such that $\Delta(q) > 0$ holds for all $q > q_{FB}$ and $\Delta(q) < 0$ holds for all $q < q_{FB}$.

Note finally that we have so far not specified the agent's utility without purchasing any product. For the following analysis, we want to focus on the case where the agent's advice

[7] Our results continue to hold if we added some noise at this stage, such that the agent would lose w only with some probability in case of a "mismatch", but also with positive probability in case of a "match".

[8] Cf. Bolton et al. (2007) or Inderst and Ottaviani (2008) for a more detailed discussion.

is directed towards the choice between different products, and not towards the decision whether to purchase or not.[9] For this purpose, we specify that it is always efficient for a customer to purchase either of the two goods: $v_l - \max_{n=a,b} c_n > 0$. For instance, it may be common knowledge that a customer must save or must get treatment for a particular illness, while the agent may be better informed about the advantages of certain products or treatments.

2.2 Analysis

We solve the game backwards. Starting at the final period $t = 3$, suppose first that the customer is willing to follow the agent's advice both if he recommends product a and if he recommends product b. In this case, the agent recommends product a only if the commission f is not below $(1-2q)w$, which captures the difference in the agent's expected penalty under the two products. If $f < w$, then this gives rise to an interior cutoff

$$q^* = \frac{1}{2} - \frac{f}{2w}, \tag{1}$$

such that the agent recommends product a if $q \geq q^*$ and product b if $q < q^*$. (Note that the realization of q^* is a zero-probability event.) We capture the corner case with $f \geq w$ by setting $q^* = 0$.

As the customer only has the coarse prior beliefs $\overline{q}$, next to the agent's advice, he calculates his expected utility based on the rationally expected choice of the threshold q^*. The customer's rational expectations for q^* depend, in turn, on his rational expectations for the commission that is offered by firm a. We denote the expected commission by $\widehat{f}$ and the respective threshold by $\widehat{q}^*$. (Again, we set $\widehat{q}^* = 0$ if it is not interior, once we substitute $\widehat{f} \geq 0$ into (1).)

When choosing the prices p_a in $t = 2$, firm a has to take the customers' expectations into account, as they determine their willingness to pay for the product. The maximum feasible price that a can charge is determined from the buyer's outside option, which is to purchase good b (even against the agent's advice).[10] The maximum price p_a that the firm

[9] Cf. Bolton et al. (2007), which focuses also on the role of advice in selecting one out of two products (instead of making a choice between purchasing or not purchasing at all, as in Inderst and Ottaviani (2008)).

[10] We need not specify whether the agent would incur also the penalty w in case the customer realized v_l after purchasing some good, say b, *against* the agent's advice. In case the advice was documented, say by a "letter of suitability", then we should, however, expect that there is no such risk of a future penalty.

can extract is thus given by the following indifference condition for the customer:[11]

$$p_a = c_b + \int_{\widehat{q}^*}^1 [v_a(q) - v_b(q)] \frac{g(q)}{1 - G(\widehat{q}^*)} dq, \tag{2}$$

which can be rewritten as

$$p_a = c' + \int_{\widehat{q}^*}^1 \Delta(q) \frac{g(q)}{1 - G(\widehat{q}^*)} dq.$$

Instead, the firm's profit depends on the true cutoff, q^*: $\pi_a := [1 - G(q^*)][p_a - f - c_a]$, such that

$$\frac{d\pi_a}{df} = g(q^*) \left[\int_{\widehat{q}^*}^1 \Delta(q) \frac{g(q)}{1 - G(\widehat{q}^*)} dq - f \right] \frac{1}{2w} - [1 - G(q^*)]. \tag{3}$$

In what follows, to simplify the derivation of results we make the following standard assumptions:

$$\frac{d}{dq}\left(\frac{g(q)}{1 - G(q)}\right) > 0, \; \frac{d}{dq}\left(\frac{g(q)}{G(q)}\right) < 0. \tag{4}$$

With (4) the firm's objective function is, for given $\widehat{q}^*$, strictly quasiconcave.

In a rational expectations equilibrium, it must hold, in addition to optimality of f, that expectations are fulfilled: $\widehat{f} = f$ and thus also $\widehat{q}^* = q^*$. Once we substitute $\widehat{q}^* = q^*$, the maximum price from (2) generates a strictly increasing mapping from q^* to p_a: The larger is the (anticipated) cutoff q^*, the higher is the customers' willingness to pay for good a if they are advised to do so. It is now helpful to substitute back p_a into the derivative $d\pi_a/df$ in (3). Then, for an interior solution $q^* > 0$, we have from $d\pi_a/df = 0$ a strictly decreasing mapping between q^* and p_a: The higher is p_a, the greater are the firm's incentives to increase f and thus to push down q^*. An interior solution $q^* > 0$ is obtained from the unique intersection of these two mappings (cf. also the proof of Lemma 1).

The following Lemma characterizes the equilibrium outcome without disclosure.

Lemma 1 *In the basic model, the unique equilibrium is characterized as follows.*
i) If

$$w \geq \left(\int_{0.5}^1 \Delta(q) \frac{g(q)}{1 - G(0.5)} dq \right) \frac{2g(0.5)}{1 - G(0.5)}, \tag{5}$$

[11] It should be noted that we abstract from the signaling problem that arises at this stage, given that firm a has already set its commission, which is not observed by customers. Hence, in the language of signaling games, the level of f would, at this stage, represent the firm's "type". With only p_a as the remaining choice variable, however, the firm has no access to a "sorting" instrument. The characterized outcome in (2) would arise, for instance, under so-called passive beliefs: Customers do not update their beliefs about f when observing p_a. The firm could, however, signal its choice of f in case it was feasible to issue warranties to customers. Given our main applications, namely to financial products and health services, we abstract from this possibility.

then it holds that $f = 0$ *and* $q^* = 0.5$.

ii) If the converse of condition (13) holds together with

$$w \left[1 + \frac{2}{g(0)} \right] > \Delta(\overline{q}), \tag{6}$$

then there is a unique cutoff $0 < q^* < 0.5$ *solving*

$$q^* = \frac{1}{2} - \frac{1}{2w} \int_{q^*}^{1} \Delta(q) \frac{g(q)}{1 - G(q^*)} dq + \frac{1 - G(q^*)}{g(q^*)}, \tag{7}$$

which is obtained as firm a *chooses the commission*

$$f = \int_{q^*}^{1} \Delta(q) \frac{g(q)}{1 - G(q^*)} dq - 2w \frac{1 - G(q^*)}{g(q^*)} > 0. \tag{8}$$

iii) Finally, if condition (6) does not hold, then we have that $q^* = 0$, *while* $f = w$.

Proof. See Appendix.

Note that cases ii) and iii), where $q^* < 0.5$ holds, are more likely when w is lower. Moreover, if $0 < q^* < 0.5$ holds, then the cutoff is strictly lower as w decreases. Intuitively, a lower value of w makes it "cheaper" for the firm to expand sales. This raises the equilibrium commission, f, and pushes down q^*.

Here and in what follows, we are mainly interested in whether the outcome is efficient. Note that the outcome is efficient if and only if $q^* = q_{FB}$. In the presently analyzed (single-agency) case, there are two cases to distinguish.

In the first case, it holds that $c_a \geq c_b$. If this holds strictly such that $q_{FB} > 0.5$, then we have immediately from $q^* \leq 0.5$ that product b is sold too infrequently. Intuitively, given that product b is sold competitively, there is no "lobby" (or "marketing") for product b, while from Lemma 1 we have for all sufficiently low values of w that $f > 0$ and, consequently, $q^* < 0.5$. In fact, if product a is from an *ex-ante* perspective more efficient than product b, which from linearity of $\Delta(q)$ holds whenever $\Delta(\overline{q}) > 0$, then we have $q^* = 0$ for all sufficiently low values of w (case iii) in Lemma 1). In summary, in the first case where $c_a \geq c_b$, commissions in the single-agency case can clearly not improve efficiency. For low w, instead, they strictly reduce efficiency. As we show below, however, this result hinges in turn crucially on the present assumption that commissions are not observed by customers.

Take next the second case where $c_a < c_b$, such that $q_{FB} < q^*$. In this case, the outcome without commissions is inefficient: With $f = 0$ it holds that $q^* = 0.5$. While it would be efficient to lower q^*, for firm a this is too expensive. Once w is sufficiently low such that condition (5) no longer holds, we have that $q^* < 0.5$. Moreover, from inspection of Lemma 1, we have that $q_{FB} \leq q^* < 0.5$ holds as long as w is not too low, such that it still satisfies

$$w\left[\left(\frac{1}{2} - q_{FB}\right) + \frac{1 - G(q_{FB})}{g(q_{FB})}\right] \geq \Delta(\overline{q}). \tag{9}$$

(Recall that $\overline{q}$ denotes the prior probability for q.) For lower values of w, where the converse of condition (9) holds strictly, we have instead that $q^* < q_{FB}$: In equilibrium, the commission is then too high, pushing q^* below the efficient cutoff.

Taken together, for the presently analyzed (unregulated) case we have the following results.

Proposition 1 *The equilibrium in the basic model has the following efficiency properties:*
i) If $c_A \geq c_B$, then the equilibrium becomes less efficient as the commission f of firm a increases, which is the case as w decreases.
ii) If $c_A < c_B$, then for low values of f and thus high values of w, the equilibrium becomes more efficient as w decreases and f, consequently, decreases. Instead, for low values of w and thus high values of f, the opposite relationship holds.

Illustration: Uniform Distribution

Suppose for an illustration that q is uniformly distributed, such that $g(q) = 1$ holds for all q. Denote $\Delta_v := v_h - v_l$ and suppose that $\Delta_c := c_b - c_a > 0$. We thus have $\Delta(q) = \Delta_v(2q - 1) + \Delta_c$, while from $\Delta(q_{FB}) = 0$ this yields

$$q_{FB} = \frac{1}{2} - \Delta_c \frac{1}{2}\frac{1}{\Delta_v}.$$

Here, $q_{FB} > 0$ is interior only if $\Delta_v > \Delta_c$.

From (1) an interior solution $q^* < 0.5$ is given by

$$\int_{q^*}^{1} \Delta(q) \frac{g(q)}{1 - G(q^*)} dq = 2w\left(\frac{3}{2} - 2q^*\right),$$

from which we have immediately that $q^* < q_{FB}$ holds for all sufficiently low values of w. Generally, in case of an interior solution we can solve for

$$q^* = \frac{3w - \Delta_c}{\Delta_v + 4w},$$

while the corner solution with $q^* = 0$ applies whenever $w \leq \Delta_c/3$. A comparison with q_{FB} reveals immediately that $q^* \geq q_{FB}$ holds only as long as

$$w \geq \frac{\Delta_v}{2} \frac{\Delta_v + \Delta_c}{\Delta_v + 2\Delta_c}. \tag{10}$$

3 Policy Options

3.1 Liability and Capped Commissions

Recall our interpretation that w may represent a "bond" that the agent must post to obtain a licence and that is forfeited in case v_l is subsequently realized. Observe next that, whenever $f > 0$, we found that $df/dw < 0$ and, consequently, $dq^*/dw > 0$. Also, if product a is a-priori more or equally efficient, i.e., if with $\overline{q} = \int_{q^*}^{1} qg(q)dq$ it holds that $\Delta(\overline{q}) \geq 0$, then we have the corner outcome $q^* = 0$ whenever w becomes sufficiently small. Otherwise, i.e., if product a is a-priori less efficient than product b as $\Delta(\overline{q}) < 0$, then q^* converges to a lower boundary $0 < \overline{q}^* < q_{FB}$ that solves

$$\int_{\overline{q}^*}^{1} \Delta(q) \frac{g(q)}{1 - G(\overline{q}^*)} dq = 0.$$

Taken together, we thus have the following immediate implications.

Proposition 2 *For $c_a \geq c_b$ efficiency in the basic model is highest when either the commission is capped at zero, $f \leq \overline{f} = 0$, or when the agent's liability w is chosen sufficiently large. If instead $c_A < c_B$ holds, then to maximize efficiency commissions should not be capped at zero, but instead at a unique value $\overline{f} > 0$, which could also be achieved by choosing a strictly positive but not too large liability level w for the agent.*

3.2 Disclosure of Commissions

As noted in the Introduction, a common proposal, in particular for financial products, is to impose mandatory disclosure of commissions. In fact, without such a regulation in place, transparent incentives for the agent simply may not be credible.

With disclosed commissions, customers can directly infer the agent's optimal choice of q^*. Consequently, the maximum price p_a that firm a can extract is given by the following indifference condition for the customer:

$$\int_{q^*}^{1} v_a(q)\frac{g(q)}{1-G(q)}dq - p_a = \int_{q^*}^{1} v_b(q)\frac{g(q)}{1-G(q)}dq - p_b.$$

Substituting $p_b = c_b$ and solving for p_a, we thus have that

$$p_a = c_b + \int_{q^*}^{1} [v_a(q) - v_b(q)] \frac{g(q)}{1-G(q)}dq. \tag{11}$$

Then, the firm's respective profits $\pi_a := [1 - G(q^*)][p_a - f - c_a]$ become, after substitution from (11),

$$\pi_a = \int_{q^*}^{1} [\Delta(q) - f] \, g(q)dq.$$

The derivative of profits with respect to f, provided that $0 < q^* < 1$, is given by

$$\frac{d\pi_a}{df} = g(q^*) [\Delta(q^*) - f] \frac{1}{2w} - [1 - G(q^*)]. \tag{12}$$

With assumption (4), the problem of firm a to maximize π_a is strictly quasiconcave. The following result follows then immediately from the firm's program.

Lemma 2 *The unique equilibrium with disclosed commissions is characterized as follows.*
i) If

$$w \geq \Delta(0)\frac{2g(0.5)}{1-G(0.5)}, \tag{13}$$

then it holds that $f = 0$ and $q^* = 0.5$.
ii) If the converse of condition (13) holds, instead, then there is a unique cutoff $0 < q^* < 0.5$ solving

$$q^* = \frac{1}{2} - \frac{\Delta(q^*)}{2w} + \frac{1-G(q^*)}{g(q^*)}, \tag{14}$$

which is obtained as firm a chooses the commission

$$f = \Delta(q^*) - 2w\frac{1-G(q^*)}{g(q^*)} > 0. \tag{15}$$

Observe that both from condition (13) and from the definition of $q^* < 0.5$ in (14) we have again that q^* is lower as w decreases.

We compare now the respective characterizations with and without disclosure from Lemmas 2 and 1. The key difference between the two cases lies in the benefits that firm a derives from raising f and, thereby, pushing down q^*. Without disclosure, an increase in f is not observed by customers, ensuring that the firm can then still charge the same price.

In contrast, with disclosure customers will change their willingness to pay for both goods. Intuitively, this reduces the incentives for firm a to further push down q^* and expand sales, in case f is disclosed. From this observation, we have immediately the following result.

Proposition 3 *In the basic (single-agency) model, in equilibrium firm a's commission without disclosure is always (weakly) larger than with disclosure, while the respective cutoff q^* is always (weakly) lower without disclosure.*

Proof. See Appendix.

In the rest of this Section, we are interested in how the two regimes with and without disclosure compare in terms of efficiency. Hence, we need to compare the respective cutoffs q^*. For ease of exposition, from now on we refer to the equilibrium cutoff with disclosure by q_D^* and to that without disclosure by q_{ND}^*.

Again, the key distinction is that between the case with $c_a \geq c_b$ and that with $c_a < c_b$. For $c_a \geq c_b$ and thus $q_{FB} \geq 0.5$, we have from Lemma 2 that $f = 0$ and thus $q_D^* = 0.5$ hold: Firm a has no incentives to push down q^*, as $\Delta(q) \leq 0$ holds already at $q_D^* = 0.5$, which is achieved without paying a positive commission. Without disclosure, however, we have from Lemma 1 that for low w, where condition (5) does no longer hold, that $q_{ND}^* < 0.5$.

Proposition 4 *For $c_a \geq c_b$ the equilibrium with disclosure is always (weakly) more efficient than that without disclosure. This holds strictly with $q_{ND}^* < q_D^* \leq q_{FB}$ whenever $f > 0$ holds without disclosure, i.e., whenever w is sufficiently small such that (5) does not hold.*

Proposition 4 captures the standard notion that disclosure of commissions is beneficial as it allows customers to truly value an agent's advice, which in turn disciplines product providers. As already noted, the "increased discipline" from disclosure arises from the adjustment of customers' willingness-to-pay for good a, as q^* changes.

However, in case good a can be more efficiently produced as $c_a < c_b$, efficiency can be higher without disclosure. To see this, observe first from Lemma 2 that for $c_a < c_b$ it holds that $q_{FB} < q_D^* < 0.5$, such that the equilibrium cutoff is inefficiently high. The intuition for this is the following. Though the firm now fully internalizes the "value of advice" to consumers, namely through an adjustment in p_a, a sale pays only $p_a - c_a - f$ instead of

$p_a - c_a$, as long as $f > 0$. In addition, when the firm wants to expand sales by lowering q^*, then the incremental commission has to be paid on "all sales", i.e., for all $q > q_D^*$. (Formally, these two effects can be seen directly from inspecting the derivative $d\pi_a/df$ in (12).)[12]

From these observation, together with the results in Lemma 1 and Proposition 1 for the case without disclosure, we have the following comparative result.

Proposition 5 *Suppose that $c_a < c_b$. Then a sufficient condition for $q_{FB} \leq q_{ND}^* < q_D^*$ to hold, such that efficiency is strictly higher without disclosure than with disclosure, is that w lies in an intermediary range, namely between the thresholds from (5) and (9).*

Proposition 5, albeit it only provides sufficient conditions, provides a first warning against the presumption that commissions are invariably too high, unless there is regulatory interference, e.g., through mandatory disclosure. In the presently analyzed case, commissions are necessary to move the advice of the agent, who does not internalize differences in costs, towards product a. As commissions are costly to firm a, however, there will be underprovision, provided that the commission is fully disclosed. Non-disclosure can mitigate this effect, albeit at the risk of an "overshooting", with $q_{ND}^* < q_{FB}$.

Returning to the case with a uniform distribution, note from (14) that an interior solution $q_D^* < 0.5$ with disclosure is given by

$$\Delta(q_D^*) = 2w\left(\frac{3}{2} - 2q_D^*\right).$$

This immediately confirms that $q_D^* > q_{FB}$ holds whenever $w > 0$. Put differently, we have that

$$q_D^* = \frac{3w + \Delta_v - \Delta_c}{2\Delta_v + 4w} = \frac{1}{2} - \frac{\Delta_c - w}{2\Delta_v + 4w}.$$

Hence, together with our previous observations, we have for this example that $q_{FB} \leq q_{ND}^* < q_D^*$ holds indeed whenever w falls into the intermediary range

$$\frac{\Delta_v + \Delta_c}{\Delta_v + 2\Delta_c} \leq w < 4\frac{\Delta_v + 2\Delta_c}{\Delta_v}.$$

[12] The inefficiency, $q_D^* > q_{FB}$ would be absent if the agent had "deep pockets", such that the additional transfer through a higher commission f could be extracted up-front, say by increasing a franchise fee.

3.3 Disclosure vs. Capped Commission and Liability

It is worthwhile to highlight separately the different implications that a (stricter) liability requirement or capped commissions have on efficiency under the two regimes with and without disclosure. With disclosure, the firm internalizes, namely through a changing willingness-to-pay of customers, how a change in commissions affects the agent's advice. Provided that firm a has incentives to influence the agent's advice at all, i.e., if $c_A < c_B$ holds, for $w > 0$ the prevailing commission is alwys too low and the resulting cutoff q_D^* always too high.

Proposition 6 *With disclosure, any policy to cap commissions or to increase the agent's liability can only reduce efficiency.*

An immediate implication of this is the following. Suppose, a change of regulation is considered to increase efficiency. As one moves away from the regime in our "benchmark" model, a change should involve either mandatory disclosure of commissions or, alternatively, capped commissions and stricter liability, but not both. With disclosure, the market outcome is most efficient when $w \to 0$.

3.4 Price Caps

As noted in the Introduction, in health care but also for some financial products price caps may be part of some regulatory regimes. Albeit this is less visible, price caps for financial products may involve a cap on the "total yield reduction" that is admissible for providers of tax-advantaged pension products. A price cap, $p_a \leq \overline{p}_a$, affects the equilibrium outcome with and without disclosure as follows.

Proposition 7 *With disclosure of commissions, a cap on the price of good a has no impact if $c_a \geq c_b$ (and thus $f = 0$ and $q_D^* = 0.5$), while for $c_a < c_b$ the imposition of a binding cap has a non-monotonic impact: If the cap is binding but not too low, then q_D^* strictly decreases, while a further reduction of a sufficiently low cap increases q_D^*. In contrast, without disclosure the reduction of a binding price cap always increases the respective cutoff q_{ND}^*.*

Proof. See Appendix.

Arguably, the most interesting case in Proposition 7 is that where, with disclosure, a reduction in $\overline{p}_a$ leads to an increase in the commission paid by firm a and, thereby, to a reductin of the respective equilibrium cutoff q_D^*. The intuition for this is the following. Without a binding cap, the firm's incentives to expand sales through paying commissions are dampened by the resulting change in customers' willingness-to-pay, as captured by $\Delta(q_D^*)$. When the cap becomes binding, however, this is no longer the case. Intuitively, a further reduction of $\overline{p}_a$ then lowers q_D^* as long as this sthill exeeds the equilibrium cutoff without disclosure, q_{ND}^*. As the cap is further decreased, however, the smaller margin that the firm earns leads now to a lower optimal commission and thus a higher cutoff. From these observations, together with Proposition 7, we thus have the following Corollary.

Corollary 1 *With $c_a \geq c_b$, a price cap $p_a \leq \overline{p}_a$ has an impact only on the outcome without disclosure, where it increases efficiency. With $c_a < c_b$, as the price cap $\overline{p}_a$ decreases, this has the following implications: If the price cap is still sufficiently high, may only have an impact on the outcome with disclosure, namely through reducing the prevailing cutoff, until it reaches the equilibrium cutoff under non-disclosure, q_{ND}^*; while from there onwards, a still lower price cap leads to a higher cutoff both with and without disclosure.*
For $c_a < c_b$, efficiency is thus affected as follows by a price cap:
i) If w is not too low such that $q_{ND}^* \geq q_{FB}$, a sufficiently high price cap increases efficiency, while a lower price cap decreases efficiency (both with and without disclosure).
iii) If w is sufficiently low such that $q_{ND}^* < q_{FB}$, then there are three regimes: A high price cap improves efficiency, namely under disclosure; an intermediate price cap reduces efficiency, again under disclosure; a still lower price cap, which then binds under both regimes, first improves efficiency and then, if it is still further reduced, decreases efficiency under both regimes.

4 Common Agency

4.1 The Case with Disclosure

We suppose now that both products $n = a, b$ are controlled each by one firm. Both firms, $n = a, b$, have thus an incentive to choose positive commissions, which we denote by f_n for $n = a, b$.

It is now convenient to solve first for the equilibrium outcome in case commissions are disclosed. In case $0 < q^* < 1$ holds, the cutoff solves

$$q^* = \frac{1}{2} + \frac{f_b - f_a}{2w}. \tag{16}$$

Instead, the agent will always recommend a if $f_a \geq f_b + w$, which we again capture by setting $q^* = 0$, while he will always recommend b if $f_b \geq f_a + w$, which is captured by setting $q^* = 1$. Note next that now also the price of good b is determined strategically. By optimality, both firms a and b will choose the respective prices as high as possible. A key implication of this is the following. When a customer who is advised to purchase a decides not to follow this advice, then given the strategically chosen price p_b he will now optimally refrain from any purchase. From optimality for firms, prices are thus given by

$$\begin{aligned} p_a &= \overline{v}_a(q^*) := E[v_a(q) \mid q \geq q^*] = \int_{q^*}^{1} v_a(q)\frac{g(q)}{1 - G(q^*)}dq, \\ p_b &= \overline{v}_b(q^*) := E[v_b(s) \mid q \leq q^*] = \int_{0}^{q^*} v_b(q)\frac{g(q)}{G(q^*)}dq. \end{aligned} \tag{17}$$

Profits of firm a are still $\pi_a = [1 - G(q^*)][p_a - f_a - c_a]$, while those of firm b are given by $\pi_b := G(q^*)][p_b - f_b - c_b]$. Substituting for prices from (17), this yields

$$\begin{aligned} \pi_a &= \int_{q^*}^{1} [v_a(q) - f_a - c_a]\, g(q)dq, \\ \pi_b &= \int_{0}^{q^*} [v_b(q) - f_b - c_b] g(q)dq. \end{aligned} \tag{18}$$

From (4) both profit functions are strictly quasiconcave in the respective strategic variable, f_a or f_b. In case $f_n > 0$ holds for $n = a, b$, we obtain from the first-order conditions the best-response functions

$$\begin{aligned} f_a &= [v_a(q^*) - c_a] - 2w\frac{1 - G(q^*)}{g(q^*)}, \\ f_b &= [v_b(q^*) - c_b] - 2w\frac{G(q^*)}{g(q^*)}. \end{aligned} \tag{19}$$

Note that assumption (4) ensures that the best-response for a in (19) is strictly increasing in q^* and thus in f_b. As this holds symmetrically for f_b, commissions are strategic complements.

In case there is an "interior equilibrium" with $0 < q^* < 1$ and $f_n > 0$ for $n = a, b$, then the respective commissions are given by (19), while after substitution into (16) we have that q^* is pinned down by

$$q^* = \frac{1}{2} - \frac{\Delta(q^*)}{2w} + \frac{1 - 2G(q^*)}{g(q^*)}. \tag{20}$$

Note that from (4) the right-hand side of (20) is strictly decreasing in q^*, implying that an equilibrium with $f_n > 0$ and $0 < q^* < 1$ is unique. Lemma 3 establishes when such an interior equilibrium exists.

Lemma 3 *With common agency, if commissions f_n are disclosed, then there exists a unique equilibrium with the following characteristics:*
i) If

$$\begin{aligned} v_a(0) - c_a &\leq 2w\frac{1 - G(0)}{g(0)}, \\ v_b(0) - c_b &\leq 2w\frac{G(0)}{g(0)}, \end{aligned} \tag{21}$$

then the equilibrium prescribes $f_n = 0$ for $n = a, b$, while $q^* = 0.5$. If both conditions do not hold, then this is a sufficient condition for that $f_n > 0$ holds for both $n = a, b$.
ii) If the equilibrium prescribes $f_n > 0$ for both $n = a, b$, then the unique cutoff $0 < q^* < 1$ solves (20), while the respective values of f_n are given by (19).

Proof. See Appendix.

4.2 The Case without Disclosure

Without disclosure, customers' willingness to pay depends again on the rationally anticipated cutoff $\widehat{q}^*$, which in turn depends on the respective anticipated commissions $\widehat{f_n}$. Profits are thus given by

$$\begin{aligned} \pi_a &= [1 - G(q^*)] \left[\overline{v}_a(\widehat{q}^*) - f_a - c_a\right], \\ \pi_b &= G(q^*)[\overline{v}_b(\widehat{q}^*) - f_b - c_b]g(q)dq, \end{aligned} \tag{22}$$

yielding, in case of $f_n > 0$ and $0 < q^* < 1$, the best-response functions

$$\begin{aligned} f_a &= [\overline{v}_a(\widehat{q}^*) - c_a] - 2w\frac{1 - G(q^*)}{g(q^*)}, \\ f_b &= [\overline{v}_b(\widehat{q}^*) - c_b] - 2w\frac{G(q^*)}{g(q^*)}. \end{aligned} \tag{23}$$

As for (19) we can observe that, using assumption (4), firms compete for the agent's advice in strategic complements. It is now useful to define

$$\overline{\Delta}(q) := [\overline{v}_a(q^*) - c_a] - [\overline{v}_b(q^*) - c_b] \, .$$

Using finally that $\widehat{f}_n = f_n$ and thus that $\widehat{q}^* = q^*$ must hold in equilibrium, we can plug this together with the best responses from (23) into (16) to obtain

$$q^* = \frac{1}{2} - \frac{\overline{\Delta}(q^*)}{2w} + \frac{1 - 2G(q^*)}{g(q^*)}. \tag{24}$$

Condition (24) is analogous to condition (20) without disclosure. The key difference is again that the *marginal* expected utilities $v_a(q^*)$ and $v_b(q^*)$ have been replaced by the *average* expected utilities $\overline{v}_a(q^*)$ and $v_b(q^*)$. Lemma 4 characterizes the equilibrium.

Lemma 4 *With common agency, if commissions f_n are not disclosed, then there exists a unique equilibrium with the following characteristics:*
i) If

$$\begin{aligned} \overline{v}_a(0.5) - c_a &\leq 2w\frac{1 - G(0.5)}{g(0.5)}, \\ \overline{v}_b(0.5) - c_b &\leq 2w\frac{G(0.5)}{g(0.5)}, \end{aligned}$$

then the equilibrium prescribes $f_n = 0$ for $n = a, b$, while $q^* = 0.5$. If both conditions do not hold, then this is a sufficient condition for that $f_n > 0$ holds for both $n = a, b$.
ii) If the equilibrium prescribes $f_n > 0$ for both $n = a, b$, then the unique cutoff $0 < q^* < 1$ solves (24), while the respective values of f_n are given by (23).

Proof. See Appendix.

4.3 Policy Analysis

In this Section, our main interest lies in a comparison of the cases with and without disclosure. Recall that, as argued previously, each firm's incentives to increase f_n are higher without disclosure, as the resulting shift in q^* has no effect on the maximum feasible price p_n that the firm can charge. In addition, this effect is now strenghtened as firms compete in strategic complements. The following result thus follows immediately from Lemmas 3 and 4. (and the respective proofs).

Proposition 8 *Commissions are higher under non-disclosure than under disclosure. Formally, we have to distinguish between the following cases. If $f_n > 0$ holds for $n = a, b$ in case of non-disclosure, then both commissions are strictly higher than under disclosure. In case $f_a > f_b = 0$ ($f_b > f_a = 0$) holds under non-disclosure, then also $f_b = 0$ ($f_a = 0$) holds under disclosure, while the respective value of f_a (f_b) is strictly lower. Finally, if $f_a = f_b = 0$ holds under non-disclosure, then this holds also under disclosure.*

As in the single-agency case, however, our main interest lies in comparing the efficiency of the two regimes. For a comparison we again refer to the respective cutoffs by q^*_D and q^*_{ND}, respectively. To isolate the different effects that are at work with common agency, we proceed stepwise, starting first with the case of full symmetry, such that $c_n = c$ for $n = a, b$, while $g(q)$ is symmetric around $q_{FB} = 0.5$. Moreover, for expositional clarity we now focus on the case where under either regime the equilibrium is "interior": $0 < q^* < 1$ for both $q^* = q^*_D$ and $q^* = q^*_{ND}$, while the respective commissions $f_n > 0$ for $n = a, b$ are pinned down by the respective first-order conditions. (A sufficient condition for this case is that w is sufficiently low.)

The following result for the case of symmetry, where also $\overline{v}_a(0.5) = \overline{v}_b(0.5)$, follows immediately from Lemmas 3 and 4.

Proposition 9 *Take the case with symmetric costs $c_n = c$ and symmetric distribution $G(q)$. Then the efficient cutoff q_{FB} prevails both with and without disclosure.*

In case of symmetry, the efficient outcome is thus also obtained without disclosure, albeit from Proposition 8 it leads to higher transfers to the agent. What ensures efficiency is that the firms have equally strong incentives to influence the agent. Even when a customer can not observe the choice of commissions, he can still rely on the fact that the firms' competition for the agent's "favour" creates just the right balance.

Proposition 9 thus illustrates a key difference between the single-agency case and the common-agency case. In the former case, only the sale of product a was possibly "pushed" by commissions, while this did not apply to the competitively provided product b. In contrast, in the case of Proposition 9 the agent's incentives to advise in favour of either product are balanced in equilibrium, given that both firms have themselves the same incentives to influence the agent. This finding is complementary to a key observation in Bolton et al. (2007), who noted that a multi-product firm may have less incentives to give wrong advice on any given product. This follows as when any given product does not provide a good fit, it is likely that the firm can, instead, sell another, more suitable product.

We now proceed to compare more generally the performance of the two regimes with and without disclosure. For ths it is helpful to consider first in more detail the case with disclosure. The following result is again immediate from Lemma 3.

Corollary 2 *For low w, where $f_n > 0$ holds for $n = a, b$, a sufficient and necessary condition for the outcome with disclosure to be efficient, such that $q_D^* = q_{FB}$, is that*

$$q_{FB} - \frac{1}{2} = \frac{1 - 2G(q_{FB})}{g(q_{FB})}. \tag{25}$$

Suppose now first that $c_a < c_b$, such that $q_{FB} < 0.5$, while the distribution is still symmetric, satisfying in particular $G(0.5) = 0.5$. Then condition (25) no longer holds. In fact, in case of an interior solution with $f_n > 0$, we have from Lemma 3 that with disclosure $q_{FB} < q_D^* \leq 0.5$ holds, at least as long as $w > 0$. For $w \to 0$, we have instead that $f_a \to v_a(q_{FB}) - c_a$ and $f_b \to v_b(q_{FB}) - c_b$, such that $q_D^* \to q_{FB}$. Efficiency thus holds in the limit. As long as $w > 0$ and $q_D^* < 0.5$, however, the outcome is inefficient. The intuition is as follows. From symmetry of $G(q)$ we have that from $q_D^* < 0.5$ firm $n = a$ pays the commission more often than firm b: $1 - G(q_D^*) > G(q_D^*)$. Ceteris paribus, this reduces the incentives of firm a to increase its commission, relative to the incentives of firm b. In equilibrium, while $c_a < c_b$ still implies $f_a > f_b$, the difference in commissions is not sufficiently large so as to push q_D^* all the way down to q_{FB}.

Suppose next that $c_a = c_b$ holds, implying $q_{FB} = 0.5$, while the distribution $G(q)$ is no longer symmetric and satisfies, in particular, $G(0.5) < 0.5$. Hence, from an ex-ante perspective it is more likely that product a provides the better fit. Again, this asymmetry implies that condition (25) no longer holds. Using Lemma 3, the equilibrium outcome is asymmetric with $q_D^* > q_{FB} = 0.5$, as long as $w > 0$. The intuition is again as follows. At the efficient cutoff $q^* = q_{FB} = 0.5$, i.e., with $f_a = f_b$, firm a would pay the commission more often than firm b, which induces a to lower f_a. In equilibrium, it will thus hold that $f_a < f_b$, such that indeed $q_D^* > q_{FB} = 0.5$.

The key question is now whether and when the case of non-disclosure can perform better in terms of efficiency. Recall that the difference between the two regimes is that without disclosure, prices do not adjust following a change in commissions and thus in the prevailing cutoff. As a consequence, incentives to choose f_n depend on the difference in the expected willingness-to-pay for both products, $\overline{v}_a(q^*) - \overline{v}_b(q^*)$, and not on the difference in the marginal willingness-to-pay, $v_a(q^*) - v_b(q^*)$. For future reference, note that with $\Delta_v := v_h - v_l$ we have that

$$v_a(q) - v_b(q) = \Delta_v(2q-1), \tag{26}$$

while we have that

$$\overline{v}_a(q) - \overline{v}_b(q) = \Delta_v \left(\frac{\int_{q^*}^1 qg(q)dq}{1-G(q^*)} + \frac{\int_0^{q^*} qg(q)dq}{G(q^*)} - 1 \right) \tag{27}$$

Comparison: Symmetric Distribution (but $c_A < c_B$)

Recall our notation for the cost difference $\Delta_c = c_b - c_a > 0$. For an illustration, we first compare the two regimes for the case of a uniform distribution: $g(q) = 1$ for all $q \in [0,1]$, such that $G(q) = q$. With disclosure, for an interior solution we obtain

$$\Delta(q_D^*) = 6w \left(\frac{1}{2} - q_D^* \right),$$

which immediately confirms the result that $q^* > q_{FB}$ holds whenever $w > 0$. Put differently, we have that

$$q_D^* = \frac{1}{2} - \Delta_c \frac{1}{2} \frac{1}{\Delta_v + 3w},$$

which compares with

$$q_{FB} = \frac{1}{2} - \Delta_c \frac{1}{2} \frac{1}{\Delta_v}.$$

Without disclosure, we obtain the requirement that

$$\overline{\Delta}(q_{ND}^*) = 6w \left(\frac{1}{2} - q_{ND}^* \right).$$

It is now helpful to observe that

$$\overline{\Delta}(q) := \overline{v}_a(q) - \overline{v}_b(q) = \Delta_v \frac{1}{2}(2q-1)$$

is everywhere strictly flatter than the difference $\Delta(q)$ in (26). From this it is already immediate that $q_{ND}^* < q_D^*$. More explicitly, we have

$$q_{ND}^* = \frac{1}{2} - \Delta_c \frac{1}{2} \frac{1}{\Delta_v/2 + 3w},$$

such that

$$q_{FB} \leq q_{ND}^{*} < q_{D}^{*} \text{ if } w \geq \frac{\Delta_v}{6}.$$

Otherwise, i.e., for still lower values $w < \frac{\Delta_v}{6}$ we have that $q_{ND}^{*} < q_{FB}$. Note also that $q_{ND}^{*} \to 1 - \Delta_c/\Delta_v$ as $w \to 0$.

Hence, the preceding illustration generalizes a key insight from the single-agency case: If costs are asymmetric, disclosure is more efficient if and only if w is sufficiently small.

For more general distributions, the preceding discussion already reveals the following insights (cf. the proof of Proposition 10 for a formalization). In case w is sufficiently small, then disclosure (generically) outperforms non-disclosure: While q_D^* converges monotonically towards $q_{FB} > 0$ as w decreases, q_{ND}^* converges monotonically to some generically different cutoff $\overline{q}_{ND}$. The latter cutoff is either defined uniquely from $\overline{\Delta}(\overline{q}_{ND}) = 0$, provided that such a value $\overline{q}_{ND} > 0$ exists, or it is equal to zero.

For a further insight observe that the difference in average utilities $\overline{\Delta}(q)$ strictly exceeds the difference in marginal utilities $\Delta(q)$ whenever q is sufficiently small. Consequently, whenever q_{FB} is sufficiently small, there exists again some intermediary range of values w such that $q_{FB} \leq q_{ND}^{*} < q_{D}^{*}$.

The following Proposition summarizes some of the key general insights. (Note for this also that $\Delta(0.5) = \overline{\Delta}(0.5)$.

Proposition 10 *Suppose the distribution is symmetric, but that $c_a < c_b$ holds. Then a comparison of the cases with and without disclosure reveals the following results:*
i) Both with and without disclosure, a decrease in w reduces the prevailing cutoff, where q_D^* converges to q_{FB} and q_{ND}^* to $\overline{q}_{ND}$ (or $q_{ND}^* = 0$ for all sufficiently low values of w).
ii) The disclosure regime becomes more efficient as w decreases. Instead, while for high w and thus high q_{ND}^* a reduction of w also improves efficiency without disclosure, in case $\overline{q}_{ND} < q_{FB}$ holds, then a lower w and, consequently, a lower q_{ND}^* reduces efficiency.
iii) For high w both regimes are equally efficient as commissions are zero, such that $q_{ND}^* = q_D^* = 0$. Provided that $\overline{q}_{ND} \neq q_{FB}$, then disclosure is more efficient for all sufficiently low values of w.
iv) For intermediary values of w, non-disclosure can be more efficient. Sufficient conditions are the following:
- $G(q)$ is uniformly distributed (as in the example).

- q_{FB} *is sufficiently low;*
- $\overline{\Delta}(q) - \Delta(q) > 0$ *holds for all* $q < 0.5$*, for which a sufficient condition is again that the difference is monotonic.*

Proof. See Appendix.

Comparison: Symmetric Cost (But Symmetric Distribution)

Suppose now that there is cost symmetry, $c_n = c$ for $n = a, b$, while the distribution is no longer symmetric. Suppose further that $G(0.5) < 0$, implying that from an ex-ante perspective it is more likely that product a provides the better fit (and thus, from $c_n = c$, is also more efficient). Note that in this case it holds that $q_{FB} = 0.5$.

Take again first the case with disclosure, from which for all sufficiently low w (where $f_b > 0$) it holds that $q_D^* > 0.5$. To see why this is intuitive, take the case where commissions are symmetric such that $q^* = 0$. Recall now that as commissions are disclosed, in terms of capturing additional sales both firms face the same trade-off with respect to the ensuing reduction in costumers' willingness to pay. (Note that this holds regardless of whether $p_a > p_b$ or of whether the opposite holds, which in turn depends on the distribution over all q.) Crucially, however, at $q^* = 0.5$ we have from $G(0.5) > 0$ that firm a pays any given commission with higher probability. This makes an increase in f_a less profitable, leading in equilibrium to $f_a < f_b$, provided that $f_b > 0$, and thus to $q_D^* > q_{FB}$. This inefficiency again decreases as w becomes smaller.

The outcome without disclosure depends, in contrast, again on the difference in average utilities, $\overline{\Delta}(q)$. At $q^* = 0.5$, the fact that $G(0.5)$ is clearly compatible with both $\overline{\Delta}(0.5) > \Delta(0.5)$ or the opposite. That is, when evaluating firm's marginal incentives to increase their commissions f_n at $q^* = \widehat{q}^* = 0.5$, without further specifications of $G(q)$ the incentives of a relative to those of b can be higher or lower without disclosure than with disclosure. Intuitively, again for intermediate values of w this implies that the regime with disclosure can perform more or less efficiently.

5 Commissions and the Provision of Other Services

In the single-agency case, the commission of firm a served the purpose of steering the agent's advice towards the firm's own product. Without any commission the agent's advice only maximizes the "fit" with respect to consumers' preferences, but it does not take into account cost differences between products. In the common-agency case, provided that the other firm's commission is positive, a firm's own commission provides a countervailing force. This was most apparent in the symmetric case, where the efficient outcome obtains both with and without disclosure.

The preceding analysis ignored, however, other functions that commissions can perform, most notably to incentivize the agent along other dimensions. In this Section, we explore this along the particular task of locating customers. We thus suppose that the agent has to exert non-contractible effort e at cost $k(e)$ in order to locate a customer with probability e. To generate an interior solution, suppose that $k(0) = 0$, $k'(0) = 0$, $k'' > 0$, while $k'(e)$ becomes sufficiently large as $e \to 1$. Importantly, this task is not firm or product specific: If a customer is located, then he may either purchase good a or good b.

It turns out that the main results that we wish to isolate in this Section are obtained most clearly in the common agency case. Furthermore, we first take the case with symmetry, i.e., where $c_n = c$ holds for $n = a, b$ and where $g(q)$ is symmetric around $q = q_{FB} = 0.5$. As we are again mainly interested in the efficiency of an equilibrium, we need some further specifications. When a customer, who has previously unaware of the products, has been located by the agent, this leads to a surplus of $v_a(q) - c_a$ or $v_b(q) - c_b$, depending on which product is ultimately sold. In case some of the "liability cost" w represents a deadweight loss, this has, however, to be taken into account additionally. Most generally, we may thus suppose that when a "surety bond" of value w is forfeited, a fraction β_C is made available for compensation of customers, while a fraction β_{DWL} represents deadweight loss. The residual fraction, $1 - \beta_C - \beta_{DWL}$ may represent a (penalty) transfer to the regulator.

Note that with this specification, if the advisor chooses a cutoff q^*, the expected surplus realized with a given customer equals

$$\omega = \int_0^{q^*} [u_b(q) - q\beta_{DWL}w]\, g(q)dq + \int_{q^*}^1 [u_a(q) - (1-q)\beta_{DWL}w]\, g(q)dq.$$

Holding q^* fixed, the respective first-best effort level is then given by $e^*_{FB} := \arg\max_e (e\omega - k(e))$. Instead, as the agent realizes with a customer the expected profits

$$u = \int_0^{q^*} [f_b - qw]\, g(q)dq + \int_{q^*}^1 [f_a - (1-q)w]\, g(q)dq,$$

his privately optimal effort level is determined by $e^* := \arg\max_e (eu - k(e))$.[13]

As a first simplification, suppose that all of w is dead-weight loss: $\beta_{DWL} = 1$. In this case, the agent's private incentives to exert effort would be equal to the socially optimal incentives if and only if commissions f_n are chosen sufficiently high such that the agent extracts all surplus, while both firms $n = a, b$ realize zero profits. With disclosure, there are *three* separate factors that are responsible for why the equilibrium effort level e_D^* will be strictly lower. To see this, note first that given symmetry a symmetric equilibrium with $f_n = f$ for $n = a, b$ and, consequently, $q_D^* = q_{FB} = 0.5$ exists.[14] Denote expected profits, including now the probability of customer acquisition, by $\widetilde{\pi}_n := e^* \pi_n$.

With disclosure, the incentives of some firm, say a, to marginally increase its commission above the symmetric level are determined by the derivative

$$\begin{aligned}\frac{d\widetilde{\pi}_a}{df_a} &= \pi_a \frac{de^*}{df_a} + e^* \frac{d\pi_a}{df_a} \\ &= \pi_a \frac{1 - G(q^*)}{k''(e^*)} + e^* \frac{1}{2w} \left[g(q^*)(p_a - f_a - c_a) - [1 - G(q^*)] \frac{dp_a}{dq^*} \right],\end{aligned} \tag{28}$$

where we substituted $\left|\frac{dq^*}{df_a}\right| = 2w$ and obtained $\frac{de^*}{df_a}$ from implicit differentiation of the first-order condition for e^*. The first term in (28) captures the additional incentives that the introduction of the customer acquisition problem generates. Importantly, firm a only takes into account its own profits, but not those of firm b. This free-riding problem dampens incentives to induce more effort. Next, the second term in (28) captures, as previously without the effort problem, the benefits from shifting advice, while holding the price p_a constant. Note here that the reduction in the margin of firm a, through a strictly positive commission $f_a > 0$, reduces the firm's incentives to further push up e^* by increasing f_a.[15]

Tthese two effects, which both work towards reducing the equilibrium level of effort below the first-best level, are also present without disclosure. The third effect that we isolate, however, is only present with disclosure. Note that the final term in (28) captures

[13] Note that for $f_n = 0$ the agent would not have any incentives to acquire a customer. However, the model could be simply extended by introducing some (future) benefits $B > 0$ from "cross-selling" other products or services to the newly acquired customer. If these benefits are also part of the overall surplus, then our analysis survives without changes.

[14] To see why assumption (4) no longer ensures uniqueness, see the proof of Proposition 12.

[15] Admittedly, the given level of $f_a = f > 0$, from which we consider a marginal deviation, already incentivizes the agent to exert effort. Still, in the language of Industrial Organization, even absent a common agency setting we would observe here a "double marginalization" problem: The optimal choice of e^* by the agent does not internalize the remaining margin of firm a, while the optimal choice of f ensures that firm a still retains a strictly positive margin.

the price impact of a change in q^* under disclosure. Hence, firm a is more reluctant under disclosure to further increase f_a as this will reduce the price that it can charge. This is not the case without disclosure. We refer to the respective equilibrium effort levels under the two regimes by e_D^* and e_{ND}^*. The following results hold irrespective of the choice of β_C and β_{DWL}.

Proposition 11 *Suppose the agent has to exert non-contractible and privately costly effort e in order to locate a customer. In case of symmetry, i.e., where $c_n = c$ and where $g(q)$ is symmetric, and for sufficiently low w such that $f_n > 0$ holds under both regimes, it still holds that $q_D^* = q_{ND}^* = q_{FB}$, while the fact that commissions are higher without disclosure implies $e_{FB}^* > e_{ND}^* > e_D^*$, such that the non-disclosure regime is strictly more efficient.*

With disclosure firms are less willing to pay commission as customers will rationally expect a change in advice and will, consequently, pay less for the product that is "pushed" more in this way. In the presently analyzed setting, however, commissions serve the additional purpose of incentivizing the agent to locate customers in the first place. Both with and without diclosure there is underprovision of this service. For the case considered in Proposition 11 this implies that mandatory disclosure of commissions reduces efficiency. More generally, such a policy chills firms' incentives to induce services that enhance welfare.

It should be noted that these observations are not restricted to the common-agency case, albeit the symmetric outcome under both disclosure and non-disclosure in Proposition 11 allows to clearly isolate the additional role of commissions. In particular, while in the single-agency case there is no free-rider problem amongst firms, still the incentives provided by firm a through its commission represent a public good, namely in this case for customers, who in the absence of monopolistic pricing power on good b realize strictly positive consumer surplus.[16]

The insights from Proposition 11 also extend to the cases where the common agency problem is asymmetric. In case non-disclosure then leads to less efficient advice, which generically holds for all sufficiently low values of w, from an efficiency perspective this would imply a trade-off: lower quality of advice without disclosure at the benefit of higher effort by the agent and thus higher probability of customer acquisition.

[16] Note that also the case with common agency could be extended so as to ensure that customers, at least in expectations or on average, realize a strictly positive rent. For instance, any given customer could have private information about his valuation, say whether a match results in utility v_h' or in utility $v_h'' > v_h'$. This would give rise to a (standard) monopolistic pricing problem.

Finally, as noted above, Proposition 11 holds irrespective of the particular specifications of β_C and β_{DWL}. These specifications affect, however, the equilibrium outcome, both in terms of efficiency of advice and in terms of effort provision. In particular, as a higher value of β_C allows firms to charge a higher price, this will lead to higher commissions both with and without disclosure. The impact on efficiency is summarized by the following results.

Proposition 12 *As a larger fraction of w represents a transfer to customers (higher β_C), commissions increase both with and without disclosure. Taking the case where $G(q)$ is symmetric but $c_a \neq c_b$, this is always beneficial with disclosure, both as advice becomes more efficient, through a reduction of q_D^*, and as more effort is provided. Without disclosure, the fact that this also reduces q_{ND}^* has an ambiguous impact on efficiency. In particular, it the conditions of case iv) in Proposition 10 are satisfied, advice becomes more efficient when it still holds that $q_D^* > q_{FB}$, while otherwise advice becomes less efficient.*

Proof. Presently omitted.

If $q^* < 0.5$ holds, note that together with symmetry of $G(q)$ we have both with and without disclosure that the incremental price that firm a can charge as β_C increases is strictly higher than that of firm b. This explains why, for the case considered in Proposition 12, the respective equilibrium cutoffs q_D^* or q_{ND}^* decrease as β_C increases. The remaining assertions are then immediate from our previous observations.

6 Discussion

6.1 Discussion of Contingent Commissions

In this Section we suppose that firms, in the common agency setting, could condition their commission payments on whether the agent subsequently has to pay w, after realization of v_l, or whether this is not the case, as the customer realized v_h. Hence, we consider separate payments in the two states, f_n^l and f_n^h.[17] For the present discussion, we also abstract from the provision of other services by the agent.

[17] It should be noted that we still do not allow the firm to extract a payment from the agent when its own product was not sold. While in this case the agent may earn a commission on the other good, thereby relaxing his "zero-wealth" constraint, we may suppose that it is not verifiable whether, following the agent's advice, a customer purchased another product.

In case of an interior cutoff, we now have more generally that

$$q^* = \frac{w + f_b^h - f_a^l}{2w + (f_a^h - f_a^l) + (f_b^h - f_b^l)}. \tag{29}$$

Apart from this, the previous programs and conditions with and without disclosure remain, in principle, unchanged. When firms are now able to pay such contingent commissions, we obtain that optimally, both with and without disclosure, it holds that $f_n^h = 0$: Firms only pay commissions when v_l is subsequently realized. In this case, the cutoff (29) simplifies to

$$q^* = \frac{f_a^l - w}{(f_a^l + f_b^l) - 2w},$$

which, interestingly, for $w \to 0$ converges to the "Tullock contest function" with

$$q^* = \frac{f_a^l}{f_a^l + f_b^l}.$$

The optimality of $f_n^h = 0$ is immediate. As presently commissions only serve the purpose of affecting the cutoff, any firm strives to do this at least cost, i.e., by ensuring that the agent obtains the lowest possible rent. Through setting $f_n^h = 0$ and $f_n^l > 0$ it is ensured that, for a given payment, the marginal effect on q^* is maximal, given that at q^* the probability that the customer realizes v_l is highest and the probability that he realizes v_h is lowest among the set of all values q for which the agent advises the customer to purchase product n.

As noted previously, however, such contracts that partially indemnify the agent in case v_l is realized may not be feasible, e.g., as the firm then runs the risk of being held vicariously liable. With the restriction to $f_h^l \geq f_n^h$, however, paying $f_n^h = f_n^l = f_n$ is thus indeed optimal from the previous observations.

Finally, it should be noted that when firms can make contracts contingent on the customers' realization of a high or low utility, then this may also allow firms to issue warranties, i.e., to make the price of their product contingent on the ultimately realized state, a or b. With warranties in place, we can show that the outcome with disclosure becomes less efficient if firms differ in costs: The low-cost firm has even less incentives to push q_D^* towards the efficient cutoff q_{FB}. Without disclosure, instead, firms may use warranties to commit not to raise their commissions and, thereby, to inefficiently shift sales towards their own product.

6.2 Incentives to Invest in Advice: Some Observations

In the present model, the agent obtains the more informative signal, i.e., q, without additional costs. In an extended model, the precision of this information (cf. footnote 3) could depend, instead, on some effort that is undertaken by the agent. This effort could be exerted either after a customer was located and thus, in particular, after firms offered commissions or, alternatively, before a customer was located, e.g., through investment in general training. In the latter case, this investment may be observable to firms, while in the former case this should not be the case.

Take the scenario where effort is exerted only once a customer was located. In the common-agency setting, take further the case of symmetry. The agent's effort could result in a mean-preserving spread of $G(p)$ around $q_{FB} = 0.5$. As long as firms offer symmetric commissions, such that $q^* = 0.5$, the level of commissions does not affect the agent's incentives, which are only determined by the objective to reduce the payment of v_l, given $q^* = q_{FB} = 0.5$. Suppose a firm deviates and offers a commission of, say, $f_a > f$. This shifts q^* to the left. The further f_a increases, the more the agent's payoff profile, as a function of q, resembles a line instead of being "V-shaped" around $q = 0.5$. In the extreme case where $f_a > f$ is such that $q^* = 0$, the agent's payoff profile is a straight line (of slope w). It is immediate that in this case there are no longer incentives to acquire information to provide good advice.

From these observations, we can conjecture that the more asymmetric the outcome, e.g., given that $c_a < c_b$ and thus $q^* < 0.5$ under both regimes, the less the agent will invest in information acquisition. To compare the overall efficiency of outcomes, e.g., with and without disclosure, this must be considered in addition.

Take next the second scenario where the agent invests initially in his quality of advice. Suppose this is observable to firms as it represents some general training or qualification. Importantly, as this is done before commissions are set, it affects firms' incentives to subsequently compete in terms of commissions. Consider for a further illustration the case of non-disclosure and symmetry. If more information results in a mean-preserving spread and thus, in particular a reduction of $g(0.5)$, then more information leads to lower commissions. For the agent there is thus a trade-off: As he invests more in observable training, this reduces his subsequent likelihood of losing w, while making it less attractive for firms to influence his advice, given that it is less likely that the "intermediary" values of q are realized, where neither of the two products is markedly more likely to provide a good match.

6.3 Naive Customers

In the presently analyzed setting, with disclosure customers can not only perfectly observe commissions, but they can also perfectly reason through the implications that any combination of commissions has on the quality of advice.

Clearly, also without disclosure, i.e., in the rational expectations equilibrium, customers need a great deal of procedural rationality in case this was a one-shot environment. The outcome may, however, also be thought of as prevailing in the long run, through adjustments, observations and learning.

Even with disclosure, only a fraction α of customers may, also in the long run, be sufficiently "wary" (or sophisticated) to draw inferences from a change in commissions on a change in q^*. When deviating, firms then face the decision of whether to set prices p_n sufficiently low so as to still serve all customers or whether they only target the "naive" customers. Intuitively, we have that for high α the outcome is exactly as in the disclosure regime where all customers are "wary", while for low α the outcome "flips" and becomes identical to that in the non-disclosure regime.

6.4 Product Innovation

In our present analysis, the characteristics of firms' products were taken as given. This reduced consideration to static efficiency only. A firm that, for instance, invests in the provision of a more cost-effective investment product (e.g., through a more efficient administering of payment collections for a savings product) will only do so if it can reap a sufficiently large fraction of the resulting benefits. If expanding sales is more costly as it requires to pay much higher commissions, which is the case when w is high, then welfare-improving investments will not be undertaken.

More generally, this observation suggests a trade-off between "consumer protection", i.e., particularly the quality of advice in the case where commissions are not or can not be credibly disclosed, and innovations.

7 Conclusion

Commissions are abundant in many industries that rely on third-party agents such as brokers in order to locate customers and, which is the focus of this paper, to provide advice. This paper introduces a modeling framework to explore the various roles of commissions and their interaction, as well as the implications of different policy options, most notably that of disclosing commissions.

The main result is that disclosing commission may not necessarily increase efficiency. It may stifle the provision of other services that are incentivized through commissions and it may also lead to inefficient advice in case disclosure is imposed, while other regulations, most notably liability for the agent, are still in place.

Future research may, by relying on this model, extend the analysis to questions concerning the Industrial Organization of the industry. This concerns, most notably, the choice of single- vs. common-agency relationships, i.e., of different forms of distribution channels.

8 Appendix: Proofs

Proof of Lemma 1. We ask first when $f = 0$ and, consequently, $q^* = 0.5$ is an equilibrium outcome. Using strict quasiconcavity of π_a, this holds if and only if $d\pi_a/df \leq 0$ is satisfied at $f = 0$ and $q^* = \widehat{q}^* = 0.5$. This gives rise to condition (5).

Suppose next that there is an "interior" equilibrium where $f > 0$ and (still) $q^* > 0$ both hold. Uniqueness follows immediately from the argument in the main text, which are now present more formally. Characterizing an equilibrium by the choice of (p_a, q^*), we thus have from (2) and $d\pi_a/df = 0$ the two conditions that

$$\begin{aligned} p_a &= c_b + \int_{q^*}^{1} [v_a(q) - v_b(q)] \frac{g(q)}{1 - G(q^*)} dq, \\ p_a &= \frac{c_a - c_b}{2w} + \frac{1}{2} - q^* + \frac{1 - G(q^*)}{g(q^*)}, \end{aligned}$$

where the first equation defines a continuous and strictly increasing mapping between p_a and q^*, while from (4) the second equation defines a continuous and strictly decreasing mapping between p_a and q^*. At an intersection, we have from $d\pi_a/df = 0$ that f satisfies (8) and q^* satisfies (7). To ensure that indeed $q^* > 0$, from the previous observations a necessary and sufficient condition is that $d\pi_a/df < 0$ holds at $q^* = \widehat{q}^* = 0$, which gives rise to condition (6), where we have used that, from linearity in q, $\int_0^1 \Delta(q)g(q)dq = \Delta(\overline{q})$. Note that condition (6) is strictly weaker than condition (5). Finally, it is immediate that an equilibrium with $q^* = 0$ exists only if (6) does not hold, while then $f = w$ is obtained from optimality for firm a, together with the (corner) condition that $q^* = 0$. **Q.E.D.**

Proof of Proposition 3. For the case where both equilibria are "interior", i.e., where $f > 0$ and $q^* > 0$ holds in either case, the result holds strictly and follows immediately from comparing (14) with (7), while using that for any given $q^* < 1$ it holds that

$$\int_{q^*}^1 \Delta(q)\frac{g(q)}{1 - G(q^*)}dq > \Delta(q^*).$$

Next, given that $q^* > 0$ holds always with disclosure, the result also holds strictly when case iii) of Lemma 1 applies, such that $q^* = 0$ holds without disclosure. Note finally that condition (13) is weaker than (5). Hence, taken together, we have shown that the result applies strictly unless (5) holds, in which case $f = 0$ and $q^* = 0.5$ hold both with and without disclosure. **Q.E.D.**

Proof of Proposition 7. Take the case with disclosure. Given the equilibrium cutoff q_D^* and substituting for the equilibrium price p_a, the price cap $p_a \leq \overline{p}_a$ binds whenever

$$\overline{p}_a - c_a < \int_{q_D^*}^1 \Delta(q)\frac{g(q)}{1 - G(q_D^*)}dq. \tag{30}$$

Note next that with a binding price cap, where a marginal change in the cutoff q^* does no longer affect the price, we have from the first-order condition to maximize π_a that, at an interior solution $f > 0$,

$$f_{cap} = \overline{p}_a - c_a - 2w\frac{1 - G(q^*)}{g(q^*)}, \tag{31}$$

implying for the resulting cutoff

$$q_{cap}^* = \frac{1}{2} - \frac{\overline{p}_a - c_a}{2w} + \frac{1 - G(q_{cap}^*)}{g(q_{cap}^*)}. \tag{32}$$

Suppose now first that $\overline{p}_a$ is not too small such that $q_{cap}^* < q_D^*$, given that $\overline{p}_a - c_a > \Delta(q_D^*)$. (Note that this condition is from monotonicity of $\Delta(q)$ weaker than condition (30)). Recall next that p_a must still satisfy the customer's "participation constraint" with

$$p_a \leq c_b + \int_{q^*}^1 [v_a(q) - v_b(q)]\frac{g(q)}{1 - G(q^*)}dq. \tag{33}$$

Take now $p_a = \overline{p}_a$. Then this gives rise to some value $q^* = q^*_{cap2}$ that just solves (33) with equality:

$$\overline{p}_a - c_a = \int_{q^*_{cap2}}^{1} \Delta(q) \frac{g(q)}{1 - G(q^*_{cap2})} dq. \quad (34)$$

(The respective commission is then given by $f = w(1-2q^*_{cap2})$.) Given strict quasiconcavity of the firm's program, we thus have that as long as $q^*_{cap2} \geq q^*_{cap}$, firm a sets $p_a = \overline{p}_a$ together with $f = w(1 - 2q^*_{cap2})$, while the resulting cutoff q^*_{cap2} is indeed strictly lower as $\overline{p}_a$ decreases. Instead, if $q^*_{cap2} < q^*_{cap}$ holds, then the optimal solution specifies $p_a = \overline{p}_a$ and $f = f_{cap}$, as in (31), implying that the resulting cutoff q^*_{cap} is now strictly higher as $\overline{p}_a$ decreases.

The preceding characterization and comparative result extend to the case where $\overline{p}_a - c_a \leq \Delta(q^*_D)$ and thus $q^*_{cap} \geq q^*_D$. (More precisely, it extends strictly as long as $q^*_{cap} < 0.5$ is interior.)

Finally, the case without disclosure is immediate. The key difference is here that at the unconstrained equilibrium values of p_a and q^*_{ND} the firm does not want to increase f and, thereby, lower the cutoff even though p_a is not affected. By our previous observations from the case with disclosure, this leaves us only with the following two cases: i) The cap does not bind; ii) or the cap is binding and the cutoff equals $q^*_{cap} > q^*_{ND}$, where q^*_{ND} represents the unconstrained equilibrium outcome. **Q.E.D.**

Proof of Lemma 3. We first establish that the firms' objective functions are strictly quasiconcave. Taking the case of $n = a$, note that

$$\frac{d\pi_a}{df_a} = g(q^*) \left[\frac{1}{2w} \left[v_a(q^*) - f_a - c_a \right] - \frac{1 - G(q^*)}{g(q^*)} \right],$$

which from (4) thus indeed ensures strict quasiconcavity, given that the derivative can change sign (from positive to negative) at most once. We next distinguish between three different cases, depending on whether $f_n > 0$ or $f_n = 0$ for $n = a, b$.

Take first the case where in equilibrium $f_n = 0$ holds for both $n = a, b$. From strict quasiconcavity of profits and using the derivatives of (18), we thus obtain that condition (21) is both necessary and sufficient. (Note, however, that this does not yet imply uniqueness in case (21) holds.) Take next the case where $f_a > f_b = 0$. In this case, the unique (candidate) equilibrium specifies f_a from (19) and q^* from

$$q^* = \frac{1}{2} - \frac{v_a(q^*) - c_a}{2w} + \frac{1 - G(q^*)}{g(q^*)},$$

provided that this is still interior with $0 < q^* < 0.5$. (Otherwise, we have that $q^* = 0$.) Necessary and sufficient conditions to support this equilibrium are that, first, (21) does not hold for $n = a$ and that, second, it holds for $n = b$ at the respective cutoff $q^* < 0.5$ that

$$v_b(q^*) - c_b \leq 2w\frac{G(q^*)}{g(q^*)}.$$

Note that this condition is stricter than that in (21). The case with $f_b > f_a = 0$ is symmetric, while the characterization of the "interior" equilibrium follows from the main text. Taken together, this characterization also establishes uniqueness. **Q.E.D.**

Proof of Lemma 4. We first establish again that the firms' objective functions are strictly quasiconcave. Taking the case of $n = a$, note that

$$\frac{d\pi_a}{df_a} = g(q^*)\left[\frac{1}{2w}\left[\overline{v}_a(\widehat{q}^*) - f_a - c_a\right] - \frac{1 - G(q^*)}{g(q^*)}\right],$$

such that (4) indeed ensures strict quasiconcavity, given that the derivative can change sign (from positive to negative) at most once. The case of π_b is analogous. The remaining argument is then analogous to that in the proof of Lemma 3. **Q.E.D.**

Proof of Proposition 10. Take first the case with disclosure and $f_n > 0$ for $n = a, b$. From (20), using also that the left-hand side is strictly increasing and the right-hand side strictly decreasing in q^*, we have that the equilibrium value q_D^* indeed satisfies $q_{FB} < q_D^* < 0.5$. (Note that we take the distribution to be symmetric with, in particular, $G(0.5) = 0.5$.) As thus $\Delta(q_D^*) > 0$, we also have that $dq_D^*/dw > 0$.

Take next the case without disclosure and $f_n > 0$ for $n = a, b$. Observe again first that the left-hand side of (24) is strictly increasing and the right-hand side strictly decreasing in q^*. Note next that $\overline{\Delta}(0.5) > 0$ holds from symmetry of $G(q)$ together with $c_a < c_b$. Consequently, as long as $q_{ND}^* > \overline{q}_{ND}$, which holds surely for all sufficiently high values of w, we have that $dq_{ND}^*/dw > 0$. As now $w \to 0$, we have that $q_{ND}^* \to \overline{q}_{ND}$. (Note that when such an interior value $\overline{q}_{ND}$ does not exist, then for sufficiently low values of w it holds that $f_b = 0$ and $q_{ND}^* = 0$.)

The remaining assertions follow from the discussion in the main text. **Q.E.D.**

9 References

Basu, A. K., Rajiv L., Srinivasan, V., and R. Staelin, 1985, Salesforce Compensation Plans: An Agency Theoretic Perspective, Marketing Science 4, 267–291.

Bolton, P., Freixas, X., and J. Shapiro, 2007, Conflicts of Interest, Information Provision, and Competition in Banking, Journal of Financial Economics 85, 297–330.

Crawford, Vi. P. and J. Sobel, 1982, Strategic Information Transmission, Econometrica 50, 1431–1451.

Darby, M.R. and E. Karni, 1973, Free Competition and the Optimal Amount of Fraud, Journal of Law and Economics 16, 67–88.

Dixit, A. K., 1996, The Making of Economic Policy: A Transaction-Cost Politics Perspective. Cambridge, MA: MIT Press.

Dulleck, U. and R. Kerschbamer, 2006, On Doctors, Mechanics, and Computer Specialists: The Economics of Credence Goods, Journal of Economic Literature 44, 5–42.

Inderst, R. and F. Hoffman, 2008, How to Sell Advice.

Inderst, R., and M. Ottaviani, 2008, Misselling through Agents, forthcoming American Economic Review.

Pahl, C., 2007, A Compilation of State Mortgage Broker Laws and Regulations, 1996-2006, Federal Reserve Bank of Minneapolis, Community Affairs Report No. 2007-2.

Tirole, J., 2001, Corporate Governance, Econometrica, 69, 1-36.