How (not) to Pay for Advice: A Framework for Consumer Protection

Roman Inderst* Marco Ottaviani†

October 2009

Abstract

Given the frequent assertion that commissions paid to brokers and other financial intermediaries lead to unsuitable advice, it is puzzling that this practice is pervasive. We offer two rationales for why advice may be compensated indirectly through (disclosed or undisclosed) commissions or other payments made by product providers, rather than directly through a flat fee paid by the customer. Indirect compensation may allow product providers and their agents to maximally exploit customers who are naive about the true conflict of interest at the advice stage. Commissions also arise under efficient contracting with wary customers, even though advice will then be biased. Policy intervention that is directed towards making customers pay directly for advice would thus be beneficial for channels populated (or for products purchased) mainly by naive customers, but it may be counterproductive otherwise.

*Johann Wolfgang Goethe University Frankfurt (IMFS) and Imperial College London. E-mail: inderst@finance.uni-frankfurt.de.

†Kellogg School of Management, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208-2013, USA. E-mail: m-ottaviani@northwestern.edu.

"Impartial advice represents one of the most important financial services consumers can receive. . . . Mortgage brokers often advertise their trustworthiness as advisors on difficult mortgage decisions. When these intermediaries accept side payments from product providers, they can compromise their ability to be impartial. Consumers, however, may retain faith that the intermediary is working for them and placing their interests above his or her own, even if the conflict of interest is disclosed. Accordingly, in some cases consumers may reasonably but mistakenly rely on advice from conflicted intermediaries." *Financial Regulatory Reform. A New Foundation: Rebuilding Financial Supervision and Regulation, US Department of Treasury, June 2009 (page 68)*

1 Introduction

The recent US subprime mortgage debacle has generated an active debate on the role of the advice households receive from brokers and other information intermediaries when purchasing mortgages as well as other financial services such as consumer credit, life insurance, and investment products.[1] According to a common practice in the retail finance industry, customers are not charged directly for advice, but end up paying indirectly through distribution fees, commissions, and other inducements (often called "kickbacks") that flow from product providers to brokers and (supposedly) independent financial advisors.[2]

Across countries, customers seem to trust the financial advice they receive and rely on it when making important decisions.[3] On the other hand, there is growing recognition and concern, also in the industry, that biased advice repeatedly leads to bouts of unsuitable

[1] "Many borrowers whose credit scores might have qualified them for more conventional loans say they were pushed into risky subprime loans. . . . The subprime sales pitch sometimes was fueled with faxes and emails from lenders to brokers touting easier qualification for borrowers and attractive payouts for mortgage brokers who brought in business. One of the biggest weapons: a compensation structure that rewarded brokers for persuading borrowers to take a loan with an interest rate higher than the borrower might have qualified for." Subprime Debacle Traps Even Very Credit-Worthy As Housing Boomed, Industry Pushed Loans To a Broader Market, *Wall Street Journal*, December 3, 2007.

[2] At least in some countries, when a customer pays directly for advice, the advisor is legally bound to pass on to the customer these benefits, implying that for the customer there is an immediate tradeoff. Also, the payments made to intermediaries may be funded by fees that are directly collected from the respective investment vehicles or that are funded from the additional interest ("yield spread") that a customer pays (cf. the discussion in Jackson and Burlingam 2007 and Keith et al. 2008).

[3] For Europe, this is documented by the large-scale survey conducted in 2003 by the European Commission (Eurobarometer 60.2, November-December 2003). In many countries, such as Finland, Germany, and Austria more than 90% of respondents stated that they expect to receive advice from financial institutions, albeit this drops to around 50% for Southern European countries like Spain and Portugal. In this pre-crisis survey, apart from some Southern European countries, the majority of customers reports that they (still) trust advice (e.g., 79% in Finland, 65% in Germany, and 75% in Austria).

advice and "misselling".[4] Would customers be better served if, instead, advisers were paid directly, through an hourly fee, as increasingly advocated by consumer groups, regulators, and politicians?[5] Advisers would then earn the same compensation regardless of the ultimate decision of the customer and would thus no longer be biased towards recommending a particular product or toward recommending a purchase altogether. But if the prevalent practice of compensation structure for advice seriously compromises its value, why would intermediaries and product providers not find a more efficient arrangement?[6]

In our model, the value of advice depends both on the advisor's effort to acquire information about the suitability (or match quality) of a particular product for a specific customer as well as on the advisor's potential bias toward the conclusion of the transaction. We identify two reasons for why, in equilibrium, customers end up paying indirectly for advice that turns out to be biased. The reasons depend on whether the customers are wary of the advisor's strategic incentives or they naively fail to adequately take into account the potentially self-interested nature of advice.

Our model is applicable both to situations in which a customer's choice is simply between purchasing a given product or not purchasing at all and to situations in which the customer must choose between different products. In the latter case, the driving force of a potential bias in advice is the difference in margins that product providers and intermediaries can jointly realize with different products. For an illustration, product providers' margins, as well as commissions, are typically higher with equity-linked or more innovative investment products. Similarly, particular types of mortgage contracts yield higher profits than more standard contracts. Even though our analysis is motivated by examples from the retail financial industry, our model applies to more general settings

[4]According to a pool of the EU members of the CFA Institute (2009), 64% of respondents "believe that the fee structure of investment products drive their sale to customers rather than their suitability to customers."

[5]In a recent consultation document, the UK's financial regulator Financial Services Authority (2009), henceforth FSA, has proposed steps to encourage a *complete* switch towards a regime in which customers pay independent financial advisers directly. The new rules would "require adviser firms to be paid by adviser charges: the rules do not allow adviser firms to receive commissions offered by product providers." As part of a package of sweeping reforms proposed in the wake of the financial crisis, the Obama administration is determined to institute a Consumer Financial Protection Agency which would have authority to write such rules to protect consumers (cf. US Department of Treasury 2009).

[6]Here, the role of advice is key. In other industries, such as the real estate market, a transaction-based fee may clearly be warranted as a way to incentivize agents to become active and search for a counterparty. (See, however, Levitt and Syverson (2008) on how real-estate brokers may not use their information to the best advantage of their clients, when using the outcome from the sales of the properties they privately own as a benchmark.)

in which the compensation structure for advisors affects the quality of their information services.

In equilibrium, naive customers who are offered to pay for advice indirectly end up underestimating the likelihood with which they ultimately purchase a "premium" product (or a product at all) that generates higher profits for the respective product provider and higher commissions for the intermediary than a "basic" offering (or no purchase). As we show, providers are then able to maximally exploit this misperception by charging customers no fee at all for advice, while compensating advisors through commissions (or distribution fees) when a purchase is made. Even though customers do not pay anything directly for advice, they are seriously shortchanged through biased advice. Furthermore, their naive perceptions of the advisors' incentives allow product providers to demand a higher price for the product than would be justified on the basis of the quality of advice customers receive.

In this case, there are two rationales for policy intervention that would make customers pay directly for advice. First, we show how a cap (or, ultimately, a ban) on contingent commissions increases consumer surplus by restricting the extent to which firms can exploit customers' naive beliefs. Second, when advisors have little to lose if customers make unsuitable choices (e.g., because they have limited reputational capital at stake or they do not have much to fear from prosecution), these regulations also increase social efficiency.

Policy intervention can, however, backfire when the practice to pay indirectly for advice arises in the presence of wary customers, who see through the incentives of product providers and intermediaries. We show that with wary customers there can be an efficiency rationale for compensating indirectly for advice. Even though this practice leads to biased advice, the overall quality of advice may be higher. This is due to the fact that in our model the quality of advice depends not only on the advisors' bias when recommending a particular option to customers, but also on the advisor's incentives to acquire information about product suitability, given a particular customer's needs and preferences. As we show, in order to induce more information acquisition, it may be efficient *not* to perfectly align the customer's and the advisor's interest.

Taken together, policy intervention that would make customers pay directly for advice would thus be beneficial for channels populated mainly by naive customers or for products purchased mainly by naive customers, but it would backfire otherwise. We also discuss

the policy option of mandating disclosure of a potential conflict of interest between customers and their (trusted) advisor. In our model, the detailed compensation received by intermediary agents would not need to be disclosed to obtain the efficient outcome, as long as a general "health warning" would make otherwise naive customers wary of the underlying conflict of interest—which is why firms themselves may be reluctant to provide such information.[7]

Our specification that naive customers do not rationally anticipate that advisors may have a conflict of interest is consistent with Malmendier and Shanthikumar's (2007) empirical analysis of security analysts.[8] Experiments with trust games also suggest that many subjects are indeed willing to follow advice (e.g., Gneezy 2005). Most notably, Cain, Loewenstein, and Moore (2005) show that subjects fail to sufficiently de-bias the advice they receive, even when an advisor's incentives are disclosed.[9] In the theoretical literature, Admati and Pfleiderer (2004), Kartik, Ottaviani, and Squintani (2007), and Hong, Scheinkman, and Xiong (2008) also add behavioral dimensions to Crawford and Sobel's (1982) game of strategic information transmission. Inderst and Ottaviani (2008a) employ a related specification in a model of cancellation terms and refunds. There, naive customers lend too much trust to the sales talk of a seller, and thus have inflated perception of a product's value.

In our model, also wary customers have some trust in financial advice because reputational concerns and the threat of prosecution ensure that preferences are at least partly aligned. This feature of the model is similar to Inderst and Ottaviani (2008b), but the focus there is squarely on the agency problem between product providers and an intermediary agent who obtains a rent because of contractual frictions. In contrast, in the

[7] In the US, the Federal Trade Commission (2008) has proposed rules that would require also for brokers to enter with customers in an initial agreement that "must state that the consumer will pay the entire compensation even if all or part is paid directly by the lender, and that a lender's payment to a broker can influence the broker to offer the consumer loan terms or products that are not in the consumer's interest or are not the most favorable the consumer could obtain". According to this staff report, a key argument used to support disclosure is that "many consumers purportedly view mortgage brokers as trusted advisors who shop for the best loan for the consumer". The staff report also highlights that excessively detailed disclosure of compensation can deflect customers' attention from key price information.

[8] Jackson (2005) shows how biased analysts' forecasting, provided it is (naively) followed, may be more informative, as it makes firm managers reveal more information to analysts.

[9] In Cain, Loewenstein, and Moore (2005) subjects are paid for the precision of the estimates of the number of coins in a jar. They can rely on the additional judgement of an "advisor", who can closely inspect the jar. While in a first treatment advisors are paid for the accuracy of the subjects' guesses of the number of coins, in a second treatment they are paid more when the guess is high. The estimate of the subjects is 28 per cent higher in the second treatment.

present model the set of feasible contracts allows a perfect alignment of the interests of the product provider and the intermediary agent. This different specification that abstracts from the agency problem within the supply chain is ideal given our focus on the external agency problem vis-à-vis customers. The key novelty of this paper is the analysis of how the quality of advice depends on *how* customers are charged for advice.[10]

Several recent contributions show that households' savings, investment and credit decisions may be inefficient. As stressed by Campbell (2006), financial decisions made by individuals are taken with only incomplete information and potentially high costs of information acquisition and may be affected by behavioral biases and bounded rationality. On the "supply side", strategies taken by the financial industry, such as excessive complexity to make price comparison more difficult or to cater to investors' different biases and sophistication levels (cf. Carlin 2008 or Christoffersen and Musto 2002), negatively affect the efficiency of households' financial decisions. Here, we focus on the role of advice, a core component of the retail financial industry that is currently coming under regulatory scrutiny.[11]

A number of recent empirical studies have looked at advice in the investment industry. In the US, mutual funds and equities (outside employer-sponsored plans) are overwhelmingly purchased through brokers.[12] Bergstresser, Chalmers, and Tufano (2007), Edelen, Evans, and Kadelec (2008), and Chen, Hong, and Kubik (2007) suggest that mutual funds sold through broker/agent networks tend to underperform and that funds with higher fees improve distribution through higher commissions. Hackethal, Inderst, and Meyer (2009) document how the initiative to reshuffle investment portfolio of the retail customers of a German bank originates overwhelmingly from bank advisors and how those customers who rely more on advice end up trading significantly more, thereby generating significantly

[10] Bolton, Freixas, and Shapiro (2007) analyze the case in which product providers (rather than advisors) fear prosecution when unsuitable advice is given. Inderst and Ottaviani (2009) add advisors, but focus on the internal agency problem that arises from a multi-tasking problem, without consideration of the way in which customers pay for advice. Carlin and Gervais (2009) analyze how the imposition of liability on product providers and intermediaries affects their efforts to collaborate (in a team environment) in the provision of services to customers.

[11] Following the publication of the first Consumer Markets Scoreboard, the Commission of the European Communities (2008) has singled out the provision of precontractual information through advice as one of the three main problem areas for the retail financial sector. In particular, see pages 12–14 of the staff working document of the Commission of the European Communities (2009).

[12] Cf. Bergstresser, Chalmers, and Tufano (2007) and, in more detail, "Equity Ownership in America 2005", http://www.ici.org/pdf/rpt_05_equity_owners.pdf.

higher fees for the bank.

The paper proceeds as follows. Section 2 introduces the model. Section 3 analyzes the advisor's optimal strategy to acquire information and provide advice. This is then used in Sections 4 and 5 to solve for the equilibrium in the presence of naive or wary customers, respectively. Section 6 analyzes the implications for welfare and consumer protection when commissions are banned by requiring that customers directly pay for advice through a fixed fee. Section 7 discusses the implications of a disclosure policy, as well as different liability standards, and Section 8 offers some concluding remarks.

2 Baseline Model

The focus of our analysis is on the quality of advice. For this purpose, we abstract from the specific characteristics of particular products, e.g., investment, savings, or credit products, and, instead, frame our analysis more generally.

Products, Preferences, and Advice. A customer has to choose between two options, $\theta = A, B$, where A always represents the choice of product A, while B may stand for another product or, alternatively, for the option of not purchasing at all. Our analysis applies to both cases.

In case the two options correspond to different products, we may think of B as representing the "basic" (or default) option, while A represents the "advanced" (or premium) option. Our focus lies on the interaction between a customer, an advisor, and the seller of product A. For instance, option B may represent the option of not investing or that of investing in Treasury bills, while option A may represent some equity-linked mutual fund. Alternatively, B could be a plain vanilla mortgage and A a more innovative arrangement.

Advice is needed to provide the right match for a customer's preferences and needs. We capture this by supposing that there are two customer types, $\widehat{\theta} = A, B$, with corresponding utilities $v_{\theta,\widehat{\theta}}$ in case product θ is matched with customer type $\widehat{\theta}$. The key assumption is that a fitting match creates higher utility, $v_{A,A} > v_{B,A}$ and $v_{B,B} > v_{A,B}$. To simplify the exposition, we impose symmetry by supposing that $v_{A,A} = v_{B,B} = v_h$ and $v_{A,B} = v_{B,A} = v_l$, with $v_h > v_l$. Furthermore, we stipulate that the *a priori* probability that the advanced product A represents the better match is $0 < q_0 < 1/2$. We comment below in detail on where this specification plays a role. Note also that q_0 captures the customer's information,

which is common knowledge.

The customer can obtain advice from an intermediary agent. The advisor's prior information about product characteristics already allows him to support the customer's decision-making. However, by spending time and effort the advisor can become more familiar with the customer's specific circumstances and can further improve the quality of information about the match.

To model the informativeness of the seller's advice, we exploit the binary structure of the match quality. Note first that any (additional) information that the advisor observes gives rise to some posterior belief, denoted by q, that product A provides a better match (i.e., that $\widehat{\theta} = A$). Next, denote the advisor's (privatively observed) effort by e, which is obtained at cost $\kappa(e)$, where we stipulate that $\kappa(0) = 0$, $\kappa'(0) = 0$, $\kappa'(e) \geq 0$ for all e, and $\kappa(e) \to \infty$ as $e \to \infty$.[13] The precision of the advisor's information is characterized by the properties of the distribution of the posterior belief that is induced by e. An increase in effort affects the cumulative distribution function of the advisor's posterior belief, $G(q \mid e)$, by inducing a mean-preserving rotation of $G(q \mid e)$, around the prior belief, q_0:

$$\frac{dG(q \mid e)}{de} > 0 \text{ for } q < q_0, \ \frac{dG(q \mid e)}{de} < 0 \text{ for } q > q_0, \ \frac{dG(q \mid e)}{de} = 0 \text{ for } q = q_0. \qquad (1)$$

For convenience, we also suppose that for all feasible effort levels $e \geq 0$ the distribution has full support on $q \in [0, 1]$ and that it is continuously differentiable in both q and e. By the law of iterated expectation, we have $\frac{d}{de} \int_0^1 q dG(q \mid e) = 0$. Given our dichotomous structure with two states, $\widehat{\theta} = A, B$, any signal structure that results in the described rotation of the posterior distribution is more informative in the sense of Blackwell (cf. Ganuza and Penalva 2009, Theorem 2).

Contracting. There are five periods.

At time $t = 1$, product provider A with cost c_A determines the respective price p_A and, at the same time, offers a contract to the advisor. The contract offered to the advisor prescribes a commission t_A that is paid only when subsequently product A is sold and a fixed payment T_A. When the customer purchases product A, the advisor bears a "handling" cost, k_A, associated to the distribution activity. We do not place any sign restrictions on

[13] Even though the time spent with customers is observable and contractible, both in the presence and the absence of a customer it may be difficult to verify how hard the advisor tries to find out the best match.

the payments to the advisor. In particular, the specification that T_A can be negative ("deep pockets" of the advisor) will allow the seller to perfectly resolve the internal agency problem in the distribution chain. This specification allows us to focus on the contracting problem with respect to customers.[14]

When option B represents no purchase, we need not make further specifications. When option B, however, represents another product, we stipulate that both its price, p_B, and the commission that the advisor receives when B is sold, t_B, are exogenous. This could be the case when B is provided competitively. With cost c_B of providing product B and with handling cost k_B for the advisor, we then have $t_B = k_B$ and $p_B = k_B + c_B$: The surplus from B is captured by customers. It is convenient to impose symmetry by setting $k_A = k_B = k$ and $c_A = c_B = c$.

At $t = 2$, provided that the advisor has accepted the seller's offer (T_A, t_A), it is the advisor's turn to stipulate a flat fee $f \geq 0$ for advice.[15] Only when this is accepted by a customer, who arrives next, does the game proceed.

At $t = 3$, the advisor can exert effort e and, thereby, privately obtain additional information. This results in a posterior belief q that A provides a better match.

At $t = 4$, based on this information, the advisor recommends to the customer which option to choose. The game at this stage is one of cheap talk. As we will show, the customer follows the seller's advice in the only informative equilibrium.[16]

At $t = 5$, the purchase decision is made, and then all payoffs are realized. Payoffs are not discounted, all players are risk neutral, and the seller's payoff is additively separable in money and the cost of effort. Note that the ultimate decision between options A and B rests with the customer. However, as will become clear, given that in equilibrium it will always be optimal for the customer to follow the seller's advice, the same outcome would be achieved when the seller makes the decision (as would result under delegated portfolio management when the model is applied to investment decisions).

It is useful to note that without advice, a customer would choose A only if $p_A - p_B \leq$

[14] See Inderst and Ottaviani (2008b) and (2009) for models in which the agent is, instead, liquidity constrained, as explained in the introduction.

[15] From $f \geq 0$, the advisor cannot "bribe" the customer into business with an up-front payment. A standard assumption to rule out such up-front transfers is the presence of a sufficiently large pool of frivolous customers, who would then turn up to cash in the fixed payment while having no intention to make a purchase.

[16] As is well known, any cheap talk game always admits a "babbling" equilibrium, in which no information is conveyed. We abstract from this uninformative equilibrium in which there is no role for advice.

$\Delta_v(2q_0 - 1)$, where $\Delta_v := v_h - v_l$. In what follows, we can indeed restrict consideration to the binary choice between A and B by supposing that $v_l \geq p_B$.

Finally, note that our subsequent results will not be affected by the outside option the intermediary agent obtains when rejecting the offer (T_A, t_A). To stick to the binary choice for the customer, we could specify that the advisor then only handles the sale of product B, without giving advice, and thereby just recovers the handling cost k.

Advisor's Preferences. If, following the advice, the customer ends up realizing v_l, we stipulate that this imposes on the advisor a cost equal to $\rho > 0$ (cf. also Bolton, Freixas, and Shapiro 2007 and Inderst and Ottaviani 2009). These may be reputational costs or they may capture the prospect of legal prosecution by courts or regulators following customer complaints regarding suitability.[17] For the purpose of the present analysis, we can be agnostic about the precise nature of these costs. In particular, in our analysis the liability standard or the nature and frequency of compliance checks by regulators, which should both affect ρ, will not represent relevant policy variables.

In practice, whether an intermediary agent concludes with a customer an explicit contract or not may impose different levels of fiduciary duty on him. In our model, this could be the case whenever $f > 0$ and may thus affect firms' preferences over different forms of contracts. For the purpose of our main analysis we abstract from this possibility. We explore it subsequently, together with a more detailed discussion of the special case where ρ arises from liability.

Customer Rationality. Our analysis distinguishes between two types of customers, wary and naive. Wary customers are perfectly aware of the advisor's incentives arising both from ρ and the commissions t_A, t_B.[18] Naive customers, instead, mistakenly believe that the quality of advice is not affected by the presence and the size of commissions. Thus, naive customers do not take into account the impact of the commissions on the advisor's incentives.

[17] For reputational costs see also Durbin and Iyer (2009). On the other side, as part of their occupational licensing procedures, various US states require mortgage brokers to post a "surety bond" or to maintain a minimum net worth (cf. Pahl 2007). A surety bond is typically posted through a third party (known as surety), who is the first to be liable but is then compelled by regulation to seek redress from the broker.

[18] As we show, because the admissible contractual set (T_A, t_A) allows us to abstract from the agency problem between the seller and the advisor, it will also be inconsequential whether t_A and t_B are disclosed to (or observed by) the customer, provided the customer is wary.

3 Quality of Advice

In this section, we suppose that given some price p_A and a fee f, the customer turns to the advisor and, subsequently, follows his advice. As we show later, this will indeed be the case in equilibrium, both with wary and with naive customers.

(Biased) Advice. Given some posterior q, in $t = 4$ the advisor recommends to purchase product A only if

$$(t_A - k) - (1 - q)\rho \geq (t_B - k) - q\rho, \tag{2}$$

taking into account both the expected private costs of a subsequent mismatch, $(1 - q)\rho$ and $q\rho$, commissions, and the symmetric handling cost k. If interior, then (2) gives rise to a cutoff

$$q^* := \frac{1}{2} - \frac{t_A - t_B}{2\rho}, \tag{3}$$

such that the advisor strictly prefers to propose A when $q > q^*$ and strictly prefers that the customer chooses B if $q < q^*$. As this is a zero probability event, without loss of generality we can specify that he advises to purchase A also when $q = q^*$. Furthermore, by the same token, we can set $q^* = 0$ when (2) holds for all $q \in [0, 1]$ and $q^* = 1$ when the converse holds strictly for all $q \in [0, 1]$.

Proposition 1 *At the advice stage, the advisor follows a cutoff rule by advising the customer to purchase product A if the posterior belief satisfies $q \geq q^*$ and to purchase product B if $q < q^*$. When $t_A - t_B \geq \rho$, then $q^* = 0$; when $t_A - t_B \leq -\rho$, then $q^* = 1$; while, otherwise, $0 < q^* < 1$ is determined by (3).*

Note that from the perspective of maximizing social surplus, the first-best cutoff would clearly be $q^* = 1/2$. This is achieved only when the commissions on both products are the same: $t_A = t_B = k$, thereby just covering the handling cost of the advisor.

Note again that when $q^* = 0$ or $q^* = 1$, then the seller's advice is not informative. Recall also that presently we assume that the customer will follow the seller's advice. Subsequently, we will see that this always holds in equilibrium, though this clearly imposes restrictions on the price-fee pairs (f, p_A) that are feasible.[19]

[19] Throughout our analysis we do not consider the option that the intermediary agent refrains from providing advice and, thereby, also avoids the costs ρ in case of a subsequent mismatch. For instance

Information Acquisition Incentives. Taking into account the outcome of the subsequent advice stage, as characterized in Proposition 1, at $t = 3$ the advisor's optimal choice of e maximizes his expected profits $\pi - \kappa(e)$, where

$$\pi := T_A + \int_0^{q^*} [t_B - k - q\rho]\, dG(q \mid e) + \int_{q^*}^1 [t_A - k - (1-q)\rho]\, dG(q \mid e). \tag{4}$$

As can be seen immediately, when $q^* = 0$ or $q^* = 1$, the advisor has no benefits from exerting effort: $d\pi/de = 0$. When $0 < q^* < 1$ is, instead, determined by (3), then (4) transforms, after partial integration, substituting for q^*, and using also that $\int_0^1 G(q \mid e)dq = 1 - q_0$, to

$$\pi = [T_A + t_A - k - \rho\,(1 - q_0)] + 2\rho \int_0^{q^*} G(q \mid e)dq. \tag{5}$$

Here, the term in brackets equals the expected profits the advisor would obtain by always recommending option A. The second term denotes the benefits, in terms of lower expected mismatch costs, when the customer makes, based on advice, a more informed decision.

From expression (5) and the definition of informativeness in (1), we have immediately the following result.

Proposition 2 *The advisor's marginal benefits from exerting effort, $d\pi/de$, are hump-shaped in the cutoff q^* that is subsequently applied at the advice stage. Precisely:*
i) $d\pi/de = 0$ when $q^* = 0$ or $q^* = 1$, while $d\pi/de > 0$ for all $0 < q < 1$;
ii) $d\pi/de$ is strictly increasing in q^* up to $q^* = q_0$ and strictly decreasing thereafter.

Proof. From (5) we have that $d\pi/de$ only depends on the derivative of $2\rho \int_0^{q^*} G(q \mid e)dq$. Note also that the choice of e does not affect q^*. This term is strictly quasiconcave in q^* due to (1), from which $G(q \mid e)$ is strictly increasing in e for all $0 < q < q_0$ and strictly decreasing for all $q_0 < q < 1$. **Q.E.D.**

When $q^* = 0$ or $q^* = 1$, the advisor does not benefit from the value of information acquired, and thus optimally chooses zero effort in this case. When q^* is not at the

under the Markets in Financial Instruments Directive (MiFID, the EU directive that governs, *inter alia*, the sale of retail financial products), a financial advisor or bank employee could be dispensed from the additional duty of care that is otherwise required by flagging out a transaction as "execution only". When ρ is not too high, refraining from advice will never be optimal when facing naive customers. When, instead, facing wary customers, this will never be optimal if both ρ is not too high and the cost of acquiring information is relatively low.

boundaries, but still far away from the prior belief, q_0, then it is rather unlikely *a priori* that new information will sway advice. In contrast, when $q^* = q_0$, so that at the prior belief the advisor would be indifferent between proposing options A or B, then any new information will (almost always) break this indifference. Proposition 2 asserts that these observations apply monotonically as $|q^* - q_0|$ changes. Proposition 2 implies the following result.

Corollary 1 *Over the range $t_A \in [t_B - \rho, t_B + \rho]$, where q^* is interior, the advisor's incentives to exert effort so as to acquire better information changes as follows in the commission t_A for product A. As long as*

$$t_A < \widetilde{t} := t_B + \rho(1 - 2q_0) \tag{6}$$

holds, the advisor's incentives are increasing in t_A. When the converse of (6) holds, a higher commission t_A reduces the advisor's incentives.

Because it is *a priori* less likely that the advanced product provides a better match, $q_0 < 1/2$, we have that $\widetilde{t} > t_B$. Starting from $t_A = t_B$, where the advisor earns the same irrespective of the option the customer chooses, we thus have that a sufficiently small increase in t_A results in more information acquisition.

Equilibrium Quality of Information. Proposition 3 now characterizes the equilibrium quality of information and advice, still for given contracts.

Proposition 3 *In case the customer follows the advice, then for given contracts the equilibrium at stages $2-4$ is characterized as follows:*
i) $t_A \geq t_B + \rho$: $e^* = 0$ and $q^* = 1$;
ii) $t_A \leq t_B - \rho$: $e^* = 0$ and $q^* = 0$;
iii) $t_A \in (t_B - \rho, t_B + \rho)$: q^* given by (3) and $e^* > 0$ solving the first-order condition

$$2\rho \int_0^{q^*} \frac{dG(q \mid e)}{de} dq = \kappa'(e). \tag{7}$$

Note that e^*, as determined by (7), need not be unique. However, any solution satisfies $e^* > 0$ by the assumption that $\kappa'(0) > 0$ and $\kappa(e) \to \infty$ as $e \to \infty$. For convenience, we assume in what follows that e^* is unique. Note, however, that all our subsequent

results hold also without uniqueness by appealing to standard monotone comparative statics methods. As a measure of the quality of advice, we use the *ex-ante* probability of a mismatch

$$L := \int_0^{q^*} q dG(q \mid e) + \int_{q^*}^1 (1-q) dG(q \mid e) = 1 - q_0 + \int_{q^*}^1 (2q-1) dG(q \mid e),$$

given some exerted effort e and threshold q^*.[20]

The following auxiliary result is helpful to characterize the equilibrium of the full game.

Lemma 1 *Taking into account the thereby induced changes in q^* and e^*, as given by Proposition 3, an exogenous increase in t_A induces an increase in the probability $1-G(q^* \mid e^*)$ with which the customer ends up purchasing product A. This increase is strict when $q^* = 0$ does not already hold.*

Proof. When q^* is interior, we have for $q^* > 0$ that

$$\frac{d}{dt_A}\left[1 - G(q^* \mid e^*)\right] = -\frac{dq^*}{dt_A}\left[g(q^* \mid e^*) + \frac{dG(q^* \mid e^*)}{de^*}\frac{de^*}{dq^*}\right]. \tag{8}$$

To determine the sign of (8), recall first that $dq^*/dt_A < 0$ by (3). Next, from implicit differentiation of (7) we obtain

$$\frac{de^*}{dq^*} = \frac{-2\rho}{SOC}\frac{dG(q^* \mid e^*)}{de^*}, \tag{9}$$

where $SOC < 0$ denotes the second-order condition for e^*. The sign of the second term in (8) is then given by $\left(\frac{dG(q^*|e^*)}{de^*}\right)^2$, which is also strictly positive. Thus, (8) is strictly positive. **Q.E.D.**

4 Serving Naive Customers

Recall that naive customers do not take into account how commissions affect advice. Instead, they invariably interpret advice based on the mistaken belief that the advisor is exclusively motivated by reputational concerns, as captured by ρ. Consequently, naive customers expect the advisor to choose an unbiased cutoff $q^* = \widehat{q}_N := 1/2$. Based on these

[20] In terms of efficiency, when the seller's cost of a mismatch ρ represents also social costs, total surplus is equal to $v_h - (\Delta_v + \rho)L$. When, instead, ρ captures a transfer, total surplus is equal to $v_h - \Delta_v L$.

beliefs, they must, consequently, anticipate that, when acquiring information, the advisor exerts effort $\widehat{e}_N > 0$ solving[21]

$$2\rho \int_0^{1/2} \frac{dG(q \mid e)}{de} dq = \kappa'(e). \tag{10}$$

Contract Design Problem. Recall now that the seller specifies the price p_A and, at the same time, offers the advisor a contract (T_A, t_A). If accepted, the advisor is then free to specify a fee f. Note that for naive customers the value of advice, as captured by the expected choice of $\widehat{q}_N$ and $\widehat{e}_N$, is independent of the contract between the seller and the advisor. The choices of p_A and f must then jointly ensure that, given these expectations, it is optimal for the customer to follow the advice at $t = 5$ and to participate by paying the fee f at $t = 2$.

When not following the advice to purchase product A, the customer's second-best option is to choose, instead, B, thereby expecting to realize

$$v_l + \Delta_v \frac{\int_{\widehat{q}_N}^1 (1-q) dG(q \mid \widehat{e}_N)}{1 - G(\widehat{q}_N \mid \widehat{e}_N)} - p_B, \tag{11}$$

where $\widehat{q}_N = 1/2$. The customer's expected payoff from following the advice in this case equals

$$v_l + \Delta_v \left[\frac{\int_{\widehat{q}_N}^1 q dG(q \mid \widehat{e}_N)}{1 - G(\widehat{q}_N \mid \widehat{e}_N)} \right] - p_A. \tag{12}$$

Comparing (11) with (12), the customer thus follows the advice to purchase product A whenever

$$p_A - p_B \leq \Delta_v \left[\frac{\int_{\widehat{q}_N}^1 (2q-1) dG(q \mid \widehat{e}_N)}{1 - G(\widehat{q}_N \mid \widehat{e}_N)} \right]. \tag{13}$$

On the other hand, when receiving the advice to choose option B, it is always optimal for the customer to follow this advice, as it is immediate to verify.[22]

Turning to the customer's participation constraint at $t = 2$, the customer's outside option equals

$$v_0 := \max \{ v_l + \Delta_v (1 - q_0) - p_B, v_l + \Delta_v q_0 - p_A \} > 0. \tag{14}$$

[21] Note that again, for simplicity but without affecting results qualitatively, we suppose that $\widehat{e}_N$ is uniquely determined.

[22] Strictly speaking, this is only immediate when $p_A \geq p_B$, which will hold when solving the seller's program below.

In what follows, we first suppose that we can neglect the option to still purchase product A even without advice. It turns out that, when solving this relaxed problem, in equilibrium p_A will be sufficiently high such that without advice this is indeed dominated by option B.[23] Comparing the customer's expected utility from participating

$$v_l + \int_0^{\widehat{q}_N} [\Delta_v(1-q) - p_B]\, dG(q \mid \widehat{e}_N) + \int_{\widehat{q}_N}^1 [q\Delta_v - p_A]\, dG(q \mid \widehat{e}_N) - f$$

to now $v_0 = v_l + \Delta_v(1-q_0) - p_B$, the customer will then only participate if

$$p_A - p_B \leq \Delta_v \left[\frac{\int_{\widehat{q}_N}^1 (2q-1) dG(q \mid \widehat{e}_N)}{1 - G(\widehat{q}_N \mid \widehat{e}_N)} \right] - \frac{f}{1 - G(\widehat{q}_N \mid \widehat{e}_N)}. \tag{15}$$

With $f \geq 0$, it follows that the *interim* constraint (13) is implied by the *ex-ante* constraint (15). Furthermore, for given p_A, it is optimal for the advisor to choose $f \geq 0$, so that (15) binds.

Turning to the specification of p_A, as well as the choice of the agency contract (T_A, t_A), observe that by allowing for a fixed transfer T_A that can also be negative, we can abstract from agency problems between the seller and the advisor. That is, under the optimal contract (T_A, t_A) the seller and the advisor will then jointly behave as if they were vertically integrated and thus maximize joint profits

$$\Pi = f + [1 - G(q^* \mid e^*)]\,(p_A - c - k) + G(q^* \mid e^*)(t_B - k) - \rho L - \kappa(e^*), \tag{16}$$

which can be further simplified by using $t_B = k$. Clearly, joint profit maximization is achieved by setting $t_A = p_A - c$ (while from optimality for the seller, T_A is set so as to extract all surplus; cf. the proof of Proposition 4). Furthermore, the seller sets p_A so as to maximize Π. We solve next for the optimal choice of p_A and f.

Characterization. As a starting point, suppose that the commission is set so that the expectations of naive customers about the choice of $q^* = \widehat{q}_N$ and $e^* = \widehat{e}_N$ are correct: $t_A = t_B = k$. By the previous remarks, this would only be optimal when, at the same time, the price for product A is set equal to $p_A = c + k$, just covering the total costs in the vertical supply chain. Then, the seller and the advisor jointly make profits only

[23] Alternatively, one may also suppose that customers have access to product A only through the intermediary agent, while they can always choose option B.

through the sale of advice. From (15) the maximum fee that can then be charged is $f = \Delta_v \int_{\widehat{q}_N}^1 (2q-1) dG(q \mid \widehat{e}_N)$.

Consider now a different contract that charges the customer less for advice, by reducing f, while at the same time charging more for the product. Note again that by optimality an increase in p_A would then be compensated by a one-by-one increase in the commission t_A. Incidentally, this reflects a common practice in the financial industry whereby "kickbacks" given to intermediaries are paid out of the additional charges that are levied to customers. We argue now that such a change, leading to a reduction in f but increases in p_A and t_A, is strictly optimal for the seller. By this argument, the seller will then optimally drive f down to zero.

As long as $t_A = p_A - c$ holds, q^* and e^* are chosen optimally from the perspective of maximizing joint surplus. Hence, by the envelope theorem, we have from (16) that a marginal reduction of f together with a marginal increase of p_A, so that (15) remains binding, leads to a marginal change in joint profits equal to

$$G(\widehat{q}_N \mid \widehat{e}_N) - G(q^* \mid e^*). \tag{17}$$

Note that (17) is zero at $t_A = t_B$, given that then the expectations $q^* = \widehat{q}_N$ and $e^* = \widehat{e}_N$ are still correct. For all higher p_A and t_A, (17) becomes strictly positive by Lemma 1. For all $t_A > t_B$, the likelihood with which option B is chosen, G, is understated by naive customers. While the increase in t_A pushes up the true probability with which A is subsequently sold, taking into account both the change of q^* and the change of e^*, this is not anticipated by naive customers, who still expect $\widehat{q}_N$ and $\widehat{e}_N$ to prevail. Hence, while after an increase in p_A naive customers are clearly only prepared to pay a lower fee f, this required discount is actually too small, because it does not adequately reflect the increase in the likelihood with which they will pay the higher price.

The preceding argument can be applied as long as $f > 0$ holds. Once $f = 0$ is obtained, we then have from (15) the corresponding price

$$p_A = p_N^* := p_B + \Delta_v \left[\frac{\int_{\widehat{q}_N}^1 (2q-1) dG(q \mid \widehat{e}_N)}{1 - G(\widehat{q}_N \mid \widehat{e}_N)} \right], \tag{18}$$

where we can further substitute the price p_B that just covers the costs $c + k$.

Proposition 4 *In equilibrium, naive customers will not be charged for advice directly:* $f = 0$. *The corresponding price of product A is given by (18), while the commission* $t_A = p_A - c$ *for product A strictly exceeds that for product B,* $t_B = k$.

Proof. To complete the characterization, it remains to specify the fixed part T_A of the contract between the seller and the intermediary agent. By optimality, the choice of T_A makes the intermediary just indifferent between acceptance and rejection. For concreteness, we specified that when the intermediary rejects the seller's offer, the intermediary continues to handle the sale of product B without providing advice and so makes zero profits (from $t_B = k$). Consequently, we have that

$$T_A = \kappa(e^*) + \rho L - [1 - G(q^* \mid e^*)]\,(t_A - k).$$

Q.E.D.

With naive customers, Proposition 4 thus offers an explanation for why fee-based compensation for advice, with $f > 0$, should indeed be rare in practice. Naive customers end up still paying for advice, namely through a higher price p_A that goes hand-in-hand with a respective increase in the commission paid to the advisor, t_A. Through this practice sellers can generate higher profits because, in equilibrium, naive customers underestimate the true probability with which they will subsequently be advised to purchase the respective product. Hence, driving f down to zero, while increasing p_A and t_A, is used as a means to exploit customers. Even though naive customers are not asked to pay up-front for advice, their *true* expected surplus with advice ends up being strictly below the surplus that they could achieve when purchasing without advice, v_0.

Note finally that with naive customers advice may remain uninformative, given that $q^* = 0$ and, consequently, also $e^* = 0$ hold. Then, the advisor always recommends product A. Substituting into (3), which determines q^*, the equilibrium characterization, from Proposition 4 together with $t_B = c_B$, $p_B = c_B + k$, and $p_A = t_A + c$, this is the case when

$$\rho \leq \Delta_v \left[\frac{\int_{\widehat{q}_N}^1 (2q-1) dG(q \mid \widehat{e}_N)}{1 - G(\widehat{q}_N \mid \widehat{e}_N)} \right]. \tag{19}$$

Though ρ also affects the naive customers' expectation of $\widehat{e}_N$, condition (19) is always satisfied when ρ is sufficiently small.

5 Serving Wary Customers

In contrast to naive customers, wary customers take into account the bias in advice that results when the advisor earns different margins on A and B as $t_A \neq t_B$.

Contracting Problem. Denote a wary customer's expectations of the advisor's cutoff and effort choices by $\widehat{q}_W$ and $\widehat{e}_W$, respectively. In contrast to the case with naive customers, with wary customers these expectations have to be satisfied in equilibrium: $\widehat{q}_W = q^*_W$ and $\widehat{e}_W = e^*_W$.

As with naive customers, from $f \geq 0$ we only have to consider wary customers' *ex-ante* participation constraint, which is given by

$$p_A - p_B \leq \Delta_v \left[\frac{\int_{\widehat{q}_W}^1 (2q-1) dG(q \mid \widehat{e}_W)}{1 - G(\widehat{q}_W \mid \widehat{e}_W)} \right] - \frac{f}{1 - G(\widehat{q}_W \mid \widehat{e}_W)} \tag{20}$$

and which will bind by optimality for the seller.

Suppose now first that customers do not observe the contract between the advisor and the seller, thus forming expectations that have to be correct in equilibrium. By our previous observations, the true choice of t_A is optimally set so as to maximize joint profits of the seller and the advisor, which for given p_A is the case when $t_A = p_A - c$ ("perfect delegation"). As a consequence, wary customers can correctly infer from p_A the resulting choice of t_A. Recall as well that through T_A the seller can extract all surplus from the advisor, while when f and p_A are set jointly so that (15) binds, the seller also extracts the full value of advice from customers. Taken together, by these observations the seller thus becomes the residual claimant and will thus choose p_A, together with $t_A = p_A - c$, so as to maximize total *ex-ante* surplus.

Before we analyze the implications of this, note that, even more immediately, the equilibrium outcome will also maximize *ex-ante* surplus when wary customers can observe the contract between the seller and the advisor, thereby immediately inferring from t_A the resulting choice of q^* and e^*.

Characterization. Denote the equilibrium outcome with wary customers by q^*_W and e^*_W, where for simplicity we concentrate on the case in which this is uniquely determined by assuming that total surplus is quasi-concave. As a benchmark, when the information quality, as determined by e, was exogenously fixed, then total surplus would be highest

when $t_A = t_B$ and thus $p_A = t_A + c$, resulting in $q^* = 1/2$. Recall now from Corollary 1 that when e is endogenous, then quality of information can be increased by choosing a higher commission t_A, as long as $t_A < \widetilde{t}$ still holds. However, given that this leads to biased advice with $q^* < 1/2$, an increase in t_A has two conflicting effects on the total quality of advice, as expressed by a reduction in L.

Given that at $q^* = 0$ we have $e^* = 0$, we must have that $q^* > 0$. By the same token, we can already conclude that $q^* > q_0$. Whether we have $q^* < 1/2$ and thus biased advice depends now on which of the two effects is stronger. Note here that at $q^* = 1/2$ the first-order effect that a marginal bias, pushing q^* below $1/2$, has on efficiency is zero. On the other hand, from $q_0 < 1/2$ we have that, as long as $\rho > 0$, a marginal reduction of q^*, through a marginal increase in t_A, has a strictly positive first-order effect on e^*. Taken together, we have the following result.

Proposition 5 *With wary customers, the equilibrium outcome, q_W^* and e_W^*, is second-best efficient and satisfies $q_0 < q_W^* < 1/2$, given that the commission for product A strictly exceeds that for product B: $t_A > t_B$.*

Proof. Denote total surplus by

$$\Omega := v_h - L(\Delta_v + \rho) - \kappa(e^*), \tag{21}$$

such that, when (20) is binding, we have that $\Pi = \Omega - v_0$. Recall also that f and p_A do not affect the choice of q^* and e^*, while t_A affects e^* only indirectly through its impact on q^*. The program to choose (T_A, t_A) and (f, p_A) so as to maximize Π thus transforms into that of choosing q^* to maximize Ω, where q^* affects e^* according to (9) and where we have to take into account the constraint $f \geq 0$, which from (20) and after substitution for p_A and t_A becomes

$$\Delta_v \int_{q^*}^{1} (2q-1) dG(q \mid e^*) - [1 - G(q^* \mid e^*)]\, \rho(1 - 2q^*) \geq 0. \tag{22}$$

Differentiating and using the envelope theorem with respect to e^*, we have

$$\frac{d\Omega}{dq^*} = -\frac{dL}{dq^*}(\Delta_v + \rho) - \frac{dL}{de^*}\frac{de^*}{dq^*}\Delta_v.$$

Using next, after partial integration, that

$$\begin{aligned} \frac{dL}{dq^*} &= -g(q^*)(2q^* - 1), \\ \frac{dL}{de^*} &= (2q^* - 1)\frac{dG(q^* \mid e^*)}{de^*} - 2\int_{q^*}^{1} \frac{dG(q \mid e^*)}{de^*} dq, \end{aligned}$$

and substituting for $\frac{de^*}{dq^*}$ from (9), this becomes

$$\begin{aligned}\frac{d\Omega}{dq^*} &= g(q^*)(1-2q^*)\left(\Delta_v+\rho\right) \\ &\quad -\Delta_v\frac{2\rho}{SOC}\frac{dG(q^*\mid e^*)}{de^*}\left[(1-2q^*)\frac{dG(q^*\mid e^*)}{de^*}+2\int_{q^*}^{1}\frac{dG(q\mid e^*)}{de^*}dq\right].\end{aligned} \tag{23}$$

From (1) we have $d\Omega/dq^* > 0$ when $q \leq q_0$ as well as $d\Omega/dq^* < 0$ at $q^* = 1/2$. Given that at $q^* = 1/2$ we have $f > 0$, it is indeed feasible to lower p_A and raise f. This proves the assertion. **Q.E.D.**

The outcome is only second-best efficient because the non-contractible quality of information, e^*, can only be affected indirectly, namely through the cutoff q^*. Thus, biased advice arises in equilibrium even with wary customers because this leads to more information acquisition by the advisor, which implies an overall higher quality of advice, at least when the bias remains sufficiently small.

Compared again to the benchmark where $t_A = t_B$ (and consequently $p_A = c + k$), in equilibrium wary customers thus pay for advice also through a higher price p_A, which is then reflected again in a one-to-one increase of the advisor's commission t_A. Differently to the case with naive customers, however, this does not lead to exploitation because now the customers' *true* individual rationality constraint holds. Further, advice will always remain valuable in equilibrium, as $q^*_W > 0$, in contrast to the case with naive customers, where $q^*_N = 0$ holds when ρ is sufficiently small (condition (19)). Finally, while with naive customers we always had that $f = 0$, implying that customers only paid indirectly for advice, with wary customers this is typically not the case. Though there is no monotonicity of the contractual variables in ρ, we can show that when ρ is sufficiently low, i.e., when the advisor's preferences are not strongly aligned with those of the customer, then $f > 0$ surely holds.[24]

Hence, even though both with wary and naive customers there is biased advice and customers pay for advice also indirectly through higher commissions, at least for low ρ the two cases exhibit the following three differences. First, advice is always informative with wary customers but can become fully uninformative with naive customers. Second, naive customers never pay directly for advice. Third, naive customers would be better off not

[24]Precisely, this follows immediately from inspection of condition (22), which is satisfied strictly for low ρ.

following advice, but instead simply choose option B. In Section 7 we discuss how it may even be possible that fee-based advice to wary customers may coexist with seemingly free advice for naive customers.

Finally, that at $q^* < 1/2$ and thus $t_A > t_B$ the second-best efficient outcome prevails depends on the specification that product A is *a priori* less likely to provide a fitting match. If option B represents the alternative of not buying, then *a priori* less than one half of all potential customers should purchase A. Alternatively, less than one half should purchase the "advanced" product that is offered (only) by A, instead of the "basic" product B that is offered competitively.

6 Banning Commissions

Based on the preceding characterization in Propositions 4 and 5, we can now analyze the implications of a policy that restricts advisors to earn profits *only* through a fixed fee. Note that when $k > 0$, then such a policy could be interpreted as requiring either that there are no transaction-based payments from the seller to the advisor or that these payments do not exceed the advisor's handling costs of administering the purchase, k. Though results hold independently, as in either case advice remains unbiased, in what follows we take the latter specification.

Proposition 6 *Suppose the advisor was restricted to earn profits only through a fixed fee on advice, as commissions can not exceed his cost k of handling a purchase. Then, $t_A = t_B = k$ holds in equilibrium, leading to unbiased advice $q^* = 1/2$ both with wary and with naive customers. Depending on whether customers are wary or naive, this policy has the following additional implications:*
i) With wary customers, this policy would not benefit customers, but would reduce the overall informativeness of advice (higher L) and reduce total surplus.
ii) With naive customers, instead, this policy would always benefit customers. In addition, when ρ is sufficiently low, then this unambiguously increases the overall informativeness of advice (lower L) and increases total surplus.

Proof. Take first assertion i) with wary customers. The claims regarding the surplus Ω, as defined in (21), and L follow immediately from the proof of Proposition 5 (i.e., more precisely, from the fact that the unconstrained contract maximizes Ω, which given that e^*

and thus $\kappa(e^*)$ are strictly higher with $q^* < 1/2$ than with $q^* = 1/2$, immediately implies that L is lower in the former case). Further, note that, by optimality, f and p_A will always be chosen so that the *ex-ante* participation constraint (20) is satisfied with equality.

Take next assertion ii) with naive customers. Using (15), which binds by optimality, we have that the customers' true expected surplus equals

$$\begin{aligned} S &= (p_A - p_B)\left[G(q^* \mid e^*) - G(\widehat{q}_N \mid \widehat{e}_N)\right] \\ &\quad +\Delta_v \left[\begin{array}{c} G(q^* \mid e^*) - 2\int_{q^*}^{1} \left[G(q^* \mid e^*) - G(q \mid e^*)\right] dq \\ -G(\widehat{q}_N \mid \widehat{e}_N) + 2\int_{\widehat{q}_N}^{1} \left[G(\widehat{q}_N \mid \widehat{e}_N) - G(q \mid \widehat{e}_N)\right] dq \end{array} \right]. \end{aligned} \tag{24}$$

We argue now that S is strictly higher after the policy intervention. For this note first that by optimality the restriction $t_A \leq t_B = k$ will be binding, such that in equilibrium $t_A = t_B$ and $q^* = 1/2$. In this case, naive customers expectations are thus correct and we have simply $S = 0$. From (24) we know, instead, that $S < 0$ when $t_A > t_B$ and $p_A > p_B$ (together with $f = 0$). Finally, the asserted impact on Ω with naive customers follows immediately from the observation that $e^*_W \to 0$ as $q^*_W \to 0$ as $\rho \to 0$, such that at least for all sufficiently low values ρ it holds that L is strictly lower when commissions are not constrained. **Q.E.D.**

Incidentally, after the policy intervention, p_A and f are both no longer pinned down uniquely. To see this, take first the case with wary customers. Even when they do not observe the contract with the advisor, they can rationally anticipate that the constraint $t_A \leq t_B = k$ will bind and that even when the seller sets a high price p_A, then can no longer incentive the advisor by more. More formally, a range of values $p_A \geq c + k$ and that and corresponding values f to make the customer indifferent is then feasible. Likewise, because now the naive customers' expectations are correct, the firm can no longer extract a higher surplus by raising p_A and simultaneously lowering f.

With wary customers, the equilibrium outcome after policy intervention is unambiguously worse. Naive customers would, instead, strictly benefit. This is intuitive as the requirement that $t_A = t_B$ (or, as made clear by the proof of Proposition 6, any binding cap on t_A) reduces the scope for exploiting customers. It is immediate that the seller is in this case worse off. On the other hand, generally, the implications for social surplus are ambiguous with naive customers. This is also intuitive, given that we know that some bias $q^* < 1/2$, which is obtained when $t_A > t_B$, is necessary to maximize the total informativeness of advice when the seller's effort e^* is endogenous. Still, as we know that without

policy interference advice will be completely uninformative when ρ is sufficiently small, for low ρ the policy's effect on welfare is unambiguously positive.

From Proposition 6 a ban on commissions would thus only be justified, both in order to improve efficiency and to protect consumers, when one could expect that the particular product or the particular channel through which the product and advice are obtained attracts mainly customers who are naive about the advisor's potential conflict of interest.

7 Discussion

7.1 Disclosure Policy

So far we discussed the policy option of banning commissions. Even though banning commission seems radical, financial service regulators are giving it serious consideration. More traditional policy measures consist in imposing disclosure of conflicts of interest.

Based on similar provision imposed earlier by UK's FSA, since January 2008 the EU's MiFID (see footnote 19) imposes mandatory disclosure for the sale of financial products, with the exclusion of investments plans wrapped into life insurances. For the US mortgage market, by now dominated by third-party brokers, in November 2008 the Department of Housing and Urban Development has strengthened the requirement to disclose to homeowners the payments brokers receive for intermediated mortgage agreements. In addition to informing customers about the level of commissions and other payments that intermediary agents receive, such disclosure policies may have the primary effect of both making customers wary of the existence of such payments and of making their existence a salient piece of information in customers' decision process. If this is the case, then it follows also from our previous analysis that firms facing naive customers would resist such mandatory disclosure.

Proposition 7 *If disclosure of commissions acts as an "eye-opener" to naive customers, the seller will strictly prefer not to disclose commissions. Welfare and consumer surplus would then be higher under mandatory disclosure (and strictly so with naive customers).*

Importantly, the argument that leads to Proposition 7 does not rely on a requirement to disclose precisely the level of commissions. Instead, it could be argued that a general "health warning" on the underlying conflict of interest between customers and advisors

would be a sufficient "eye-opener". Inderst and Ottaviani (2008b) have pointed out that, in a model of common agency and multi tasks, a disclosure of commission levels in the presence of only wary customers may aggravate an inefficiency in the underprovision of incentives for an intermediary agent. Disclosing only the existence of commissions would not have such a drawback. Further, in an experiment for the US Federal Trade Commission, Lacko and Pappalardo (2004) conjecture that commission disclosure leads to information overload and, thereby, diverts customers' attention from more important information. Again, a general health warning should not have such a potential disadvantage. On the other hand, the experiment of Cain, Loewenstein, and Moore (2005) suggests that in some circumstances disclosure of bias may lead advisers to feel morally justified when deviating from professional standards, resulting in a reduction in the quality of advice.

7.2 Liability

In practice, whether an intermediary agent concludes with a customer an explicit contract or not may impose different levels of fiduciary duty on him. In our model, this could be the case whenever $f > 0$ and may thus affect firms' preferences over different forms of contracts. We explore this next.

For this purpose we have to be less agnostic about the origins of ρ. We now stipulate that B represents an alternative product and interpret ρ as the expected costs incurred by the advisor when a purchased product is subsequently verified not to represent an appropriate match, i.e., when $v = v_l$ is realized. Generally, we can now conceive that various fractions of $\rho > v_h - v_l$ represent either deadweight loss, a transfer to some regulator or supervisor, or also compensation for the customer.

A particular clear-cut case is that in which all of ρ, in expectation, is paid to the customer. When customers are wary, in this case it is immediate to show that industry profits and efficiency are strictly increasing in ρ. However, this does not necessarily hold with naive customers. When ρ increases, firm A can still increase its price p_A because the customer expects higher effort by the advisor. The customer, however, underestimates the likelihood of receiving compensation, implying that the firm is unable to extract all of this compensation. As a consequence, a higher ρ, when all of it represents compensation to the customer, increases customer surplus and efficiency, but may decrease industry profits.

Suppose now that two different standards of fiduciary duty apply when $f = 0$ (i.e., for

a "pure broker") and when $f > 0$ (i.e., for a "true advisor"), namely $\rho_{Br} < \rho_{Ad}$. Then even though a higher standard would make higher effort in information provision credible, with naive customers firms may still choose $f = 0$ and, thereby, opt for the lower standard ρ_{Br}. Instead, the imposition of the requirement that also pure brokers (who do not enter with customers into a separate contract for advice) face the same fiduciary duty can lead to higher efficiency and consumer surplus when customers are naive (even though we should still observe that $f = 0$).

When a large fraction of ρ represents deadweight loss, however, then it is intuitive that the imposition of greater fiduciary duty may result in a reduction in efficiency. However, this imposition still increases consumer surplus when customers are naive.

7.3 Discrimination

So far we analyzed the seller's problem when facing either naive or wary customers. Suppose now that a market or a particular sales channel was populated by both naive and wary customers.

Single Offer. We suppose first that the seller has to design a single offer, p_A, and that this also applies to the advisor, who then offers a single fee f. We also assume throughout the following discussion that the commission is not directly observed by either wary or naive customers.

As a starting point, take now again the case where $p_A = k+c$, such that wary customers can rationally anticipate that $t_A = t_B = k$ is optimally chosen and thus have the same expectations as naive customers. They would thus have the same willingness to pay up-front for advice, f. Consider now a marginal increase in p_A. Naive customers then require that the fee is lowered by $\Delta f = \Delta p_A[1 - G(\widehat{q}_N \mid \widehat{e}_N)]$. While wary customers rationally anticipate that the likelihood of paying the higher price for A is actually higher, as the seller optimally increases t_A and as thus q^* decreases, wary customers also anticipate that the quality of information is higher as e^* then strictly exceeds $\widehat{e}_N$. Generally, it is then not clear whether naive or wary customers will accept a lower reduction of f in exchange for a marginal increase in p_A at $p_A = k + c$. In what follows we restrict attention to the case where ρ is sufficiently small. Intuitively, together with $\kappa'' > 0$, this will ensure that the effect on e^* will weigh in less than the shift of q^*.

As an immediate consequence, for all (single) offers, other than that where $p_A = k + c$ and thus $t_A = t_B$, it will be the participation constraint of wary customers that constrains the seller's choice of f, while that of naive customers is slack, provided the offer is made acceptable to all customers. The following result is then immediate.

Proposition 8 *Suppose that only a single offer is made in a market with both naive and wary customers. When ρ is sufficiently low and when all customers participate, then the outcome is the same as that prevailing with only wary customers.*

Proof. Recall from the main text that for naive customers the marginal rate of substitution between p_A and f is obtained from $\Delta f = -\Delta p_A[1 - G(\widehat{q}_N \mid \widehat{e}_N)]$:

$$\frac{\Delta f}{\Delta p_A} = -\left[1 - G(\widehat{q}_N \mid \widehat{e}_N)\right]. \tag{25}$$

For wary customers, we have from their participation constraint that the marginal rate of substitution is obtained from total differentiation of

$$\int_{\widehat{q}_W}^{1} \left[(2q-1)\Delta_v - (p_A - p_B)\right] dG(q \mid \widehat{e}_W) - f,$$

taking into account that $t_A = p_A - c$ and thus that $d\widehat{q}_W/dp_A = -1/(2\rho)$, when interior:

$$\frac{\Delta f}{\Delta p_A} = -\left[1 - G(\widehat{q}_W \mid \widehat{e}_W)\right] - \frac{1}{2\rho}\left[\begin{array}{c} g(\widehat{q}_W \mid \widehat{e}_W)\left[(2\widehat{q}_W - 1)\Delta_v - (p_A - p_B)\right] \\ + \frac{d\widehat{e}_W}{d\widehat{q}_W}\frac{d}{d\widehat{e}_W}\int_{\widehat{q}_W}^{1}\left[(2q-1)\Delta_v - (p_A - p_B)\right] dG(q \mid \widehat{e}_W) \end{array}\right]. \tag{26}$$

As $\frac{d\widehat{e}_W}{d\widehat{q}_W} \to 0$ when$\rho \to 0$ and as we consider values $p_A > p_B$ where also $2\widehat{q}_W - 1 < 0$, we have that (26) is strictly lower than (25) for all considered values of p_A in case ρ is sufficiently small. By the argument in the main text this implies that the participation constraint for wary customers must bind. **Q.E.D.**

Discriminatory Offers. The preceding observations already pave the way for an analysis of the case where the seller and the advisor could (jointly) try to discriminate between wary and naive customers. We suppose that direct discrimination is not possible, implying that the menu of offers must be incentive compatible. Such a menu could prescribe different up-front payments for advice in exchange for different prices for product A. As noted above, in practice this could be implemented such that either the advisor is paid

some (incremental) commission or that the equivalent of this is transferred to the customer through lowering the product price in case he chooses a higher up-front payment for advice.

Intuitively, when they want to ensure participation of all customers, it is then again the participation constraint of wary customers that binds, next to the incentive compatibility constraint of naive customers. Moreover, for sufficiently low ρ, we have that for any increase in p_A wary customers demand a larger reduction in f than naive customers (by single crossing; cf. the proof of Proposition 8). Intuitively, to achieve incentive compatibility at lower cost, in terms of leaving naive customers with a lower (only perceived) information rent, this implies that the contract designed for wary customers now exhibits a still higher up-front fee and, in turn, a still lower product price and thus commission for the advisor. As is also immediate, while now the presence of wary customers still somewhat protects naive customers, this is no longer as complete as in Proposition 8 with a uniform offer, implying that under a discriminatory menu naive customers still realize a negative *true* customer surplus. For brevity's sake we omit a further formal analysis.

8 Conclusion

Advice is frequently paid for not directly, say through an hourly fee, but indirectly: Intermediary agents are paid by product providers when, following their advice, customers decide to purchase a particular product. In the financial industry, in particular, this practice has lead to widespread claims of unsuitable advice. Policy proposals consequently include that of prohibiting or, at least, seriously capping of such commissions, thereby also inducing intermediary agents to charge directly and more transparently for advice. However, these or other policy proposals that are meant to rectify a potential market failure can clearly only be evaluated after having identified the precise reason for why the market does not lead to a more efficient contractual solution.

We argue that when firms face customers who are naive about the true conflict of interest that is induced by commissions, then they can maximally exploit this naiveté through only charging customers indirectly for advice. In this case, banning commissions will protect customers and, at least when advisors' intrinsic preferences are not strongly aligned with those of customers, it will also increase welfare. Alternatively, we suggest that a mandatory disclosure policy could serve as an eye-opener, making also naive cus-

tomers wary of the underlying conflict of interest. Firms would then have themselves no incentives to disclose commissions. Such disclosure would also not risk to have unintended consequences in the presence of wary customers. As we show, with wary customers a ban on commissions, though leading to unbiased advice, tends to decrease the overall quality of advice in case advisors have to exert effort so as to acquire information.

The present analysis can be seen as a first step of a research program that tries to come up with positive and normative conclusions on how customers should pay for advice. In further research one may want to embed the present model into a more competitive environment, where advisors, possibly with different reputation and different contractual arrangements with product providers (e.g., tied vs. untied) interact with customers who are themselves in a position to, at cost, acquire additional information.

9 References

Admati, Anat R., and Paul Pfleiderer. 2004. "Broadcasting Opinions with an Overconfident Sender." *International Economic Review*, 45(2), 467–498.

Bergstresser Daniel, John M.R. Chalmers and Peter Tufano. 2007. "Assessing the Costs and Benefits of Brokers in the Mutual Fund Industry," *Review of Financial Studies*, forthcoming.

Bolton, Patrick, Xavier Freixas, and Joel Shapiro. 2007. "Conflicts of Interest, Information Provision, and Competition in Banking." *Journal of Financial Economics*, 85, 297–330.

Cain, Daylian M., George Loewenstein, and Don A. Moore. 2005. "The Dirt on Coming Clean: Perverse Effects of Disclosing Conflicts of Interest." *Journal of Legal Studies*, 34(1), 1–25.

Campbell, John Y. 2006. "Household Finance." *Journal of Finance*, 61, 1553-1604.

Carlin, Bruce I. 2008. "Strategic Price Complexity in Retail Financial Markets." *Journal of Financial Economics*, forthcoming.

Carlin, Bruce I., and Simon Gervais. 2008. "Legal Protection in Retail Financial Markets." UCLA and Duke mimeo.

Chen, Joseph, Harrison Hong, and Jeffrey D. Kubik. 2006. "Outsourcing Mutual Fund Management: Firm Boundaries, Incentives, and Performance." University of Southern California, Princeton and Syracuse mimeo.

Christoffersen, Susan E.K. and Musto, David. 2002. "Demand Curves and the Pricing of Money Management." *Review of Financial Studies*, 15(5), 1499–1524.

CFA Institute. 2009. "European Union Member Poll on Retail Investment Products." Summary Report.

Commission of the European Communities. 2008. The Consumer Markets Scoreboard: Monitoring Consumer Outcomes in the Single Market.

Commission of the European Communities. 2009. On the Follow Up in Retail Financial Services to the Consumer Markets Scoreboard. Commission Staff Working Document 1251.

Crawford, Vincent P., and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica*, 50, 1431–1451.

Durbin, Erik, and Ganesh Iyer. 2009. "Corruptible Advice." *American Economic Journal: Microeconomics*, 1(2): 220–242.

Edelen, Roger M., Richard Evans, and Gregory B. Kadelec (2008). "What do Soft-dollars buy? Performance, Expense Shifting, Agency Costs." AFA San Francisco Meetings.

Federal Trade Commission. 2008. "Before the Board of Governors of the Federal Reserve System: In the Matter of Request for Comments on Truth in Lending, Proposed Rule Docket No. R-1305".

Financial Services Authority. 2009. "Distribution of Retail Investments: Delivering the RDR." Consultation Paper 09/18.

Ganuza, Juan-Josè, and Josè S. Penalva. 2007. "Signal Orderings Based on Dispersion and the Supply of Private Information in Auctions." *Econometrica*, forthcoming.

Gneezy, Uri 2005. "Deception: The Role of Consequences." *American Economic Review,* 95(1), 384–394

Hackethal, Andreas, Roman Inderst, and Steffen Meyer. 2009. "How Banks Benefit from Investment Advice". Mimeo.

Hong, Harrison, Jose Scheinkman, and Wei Xiong. 2008. "Advisors and Asset Prices: A Model of the Origins of Bubbles." *Journal of Financial Economics*, 89(2), 268–287.

Inderst, Roman, and Marco Ottaviani. 2008a. "Sales Talk, Cancellation Terms, and the Role of Consumer Protection." Northwestern and Frankfurt mimeo.

Inderst, Roman, and Marco Ottaviani. 2008b. "Intermediary Commissions and Kickbacks." Northwestern and Frankfurt mimeo.

Inderst, Roman, and Marco Ottaviani. 2009. "Misselling through Agents." *American Economic Review*, 99(3), 883–908.

Jackson, Andrew R. 2005. "Trade Generation, Reputation, and Sell-Side Analysts." *Journal of Finance*, 60(2), 673–717.

Jackson, Howell E., and Laurie Burlingame. 2007. "Kickbacks and Compensation: The Case of Yield Spread Premiums." *Stanford Journal of Law, Business & Finance*, 12, 289–361.

Kartik, Navin, Ottaviani, Marco, and Francesco Squintani. 2007. "Credulity, Lies, and Costly Talk." *Journal of Economics Theory*, 134 (1), 93–116

Keith, Ernst, Debbie Bocian, and Wei Li. 2008. "Steered Wrong: Brokers, Borrowers, and Subprime Loans." Center for Responsible Lending.

Korobkin, Russell. 2003. "Bounded Rationality, Standard Form Contracts, and Unconscionability." *University of Chicago Law Review*, 70(4), 1203–1295.

Lacko, James M., and Janis N. Pappalardo. 2007. "Improving Consumer Mortgage Disclosures: An Empirical Assessment of Current and Prototype Disclosure Forms." Federal Trade Commission Bureau of Economics Staff Report.

Levitt, Steven D. and Syverson, Chad. 2008. "Market Distortions When Agents are Better Informed: The Value of Information in Real Estate Transactions." *Review of Economics and Statistics*, 90, 599–611.

Malmendier, Ulrike, and Devin Shanthikumar. 2007. "Are Small Investors Naive About Incentives?" *Journal of Financial Economics*, 85(2), 457–489.

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

US Department of Treasury, 2009, Financial Regulatory Reform. A New Foundation: Rebuilding Financial Supervision and Regulation, White Paper, http://www.financial stability.gov /docs/regs/FinalReport_web.pdf