How (Not) to Pay for Advice: A Framework for Consumer Financial Protection[*]

Roman Inderst[†]   Marco Ottaviani[‡]

December 2010

ABSTRACT

This paper investigates the determinants of the compensation structure for brokers who advise customers regarding the suitability of financial products. Our model explains why brokers are commonly compensated indirectly through contingent commissions paid by product providers. While biasing the broker's recommendation, commissions can enhance efficiency by improving the broker's incentives to acquire information. When customers are naive about the broker's conflict of interest, product providers exploit the customers' incorrect perceptions by increasing product prices and commissions, while charging no direct fee for advice. We analyze the effectiveness of various consumer financial protection regulations depending on the rationality of customers.

Journal of Economic Literature Classiffication Codes: D18, D83, G24, G28.

Keywords: Brokers, financial advisers, commissions, consumer financial protection, disclosure.

"Impartial advice represents one of the most important financial services consumers can receive. . . . Mortgage brokers often advertise their trustworthiness as advisers 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)*

Across countries, customers rely on recommendations from brokers and other financial advisers when making important decisions about purchasing financial services such as mortgages, consumer credit, life insurance, and investment products.[1] In many instances, however, the recommendations may be biased, because often the advising intermediaries are not paid directly by customers but, instead, receive commissions and other distribution fees from the providers of financial products.[2] These payments may tilt their recommendations towards particular financial products.[3] Likewise, when the payments from product providers are proportional to the size of transactions (or when the adviser is compensated only when

---

[1] A large-scale survey conducted in 2003 by the European Commission (Eurobarometer 60.2, November-December 2003) documented that in many European countries such as Finland, Germany, and Austria more than 90% of respondents responded that they expect to receive advice from financial institutions. In Southern European countries like Spain and Portugal this percentage drops to around 50%. In this pre-crisis survey, apart from some Southern European countries, the majority of customers report to trust advice (e.g., 79% in Finland, 65% in Germany, and 75% in Austria). Also for the US, the role of professional financial advice for the purchase of investment products (outside employer-sponsored plans) has been much documented. For instance, in a survey conducted by the Investment Company Institute (ICI 2007), over 80 percent stated that they obtained financial advice from professional advisors or other sources (cf. also *Equity Ownership in America* 2005, http://www.ici.org/pdf/rpt_05_equity_owners.pdf).

[2] 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."

[3] "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.

a transaction is made), customers may be induced to take larger positions (or to make more frequent transactions).[4]

There is growing concern among government regulators that this compensation structure may lead to unsuitable advice.[5] Would customers of retail financial services be better served if, instead, intermediaries were paid directly, through an hourly fee, as increasingly advocated by consumer groups, regulators, and politicians?[6] Brokers or financial 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 service. But if the prevalent compensation structure for advice seriously compromises its value, why would intermediaries and product providers not find a more efficient arrangement?[7]

This paper offers a rationale for the prevailing compensation structure and then investigates the need for policy intervention from a normative perspective. To this end, we propose a model that jointly endogenizes the payments that product providers make to intermediaries such as financial advisers as well as the way customers pay for financial products and advice. In equilibrium, lower up-front fees for advice but higher product prices (e.g., higher loads for investment products or higher interest rates) are associated with higher commissions or other inducements that are paid to advisers or brokers.

Payments from product providers bias the intermediary's recommendation. However, the quality of advice depends not only on the intermediary's use of any given information

---

[4]Hackethal, Inderst, and Meyer (2010) document how branches of a large German bank make considerably higher revenues from security transactions when retail customers report to strongly rely on the bank's advice.

[5]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).

[6]In a recent consultation document, the UK 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 enacted in the wake of the financial crisis, the US Consumer Financial Protection Act of 2010 has instituted a Bureau of Consumer Financial Protection which has authority to write such rules to protect consumers (see the Dodd–Frank Wall Street Reform and Consumer Protection Act, Title X).

[7]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.

but also on the quality of information that the intermediary acquires in the first place. The intermediary may spend time and effort both to learn about different products and to become familiar with the preferences and needs of a particular customer. We show that unbiased advice may entail less information acquisition. Then, the overall quality of financial advice is jointly determined by the up-front payment and the commissions intermediaries receive respectively from customers and product providers.

As we show, the resulting outcome is second-best efficient only when all customers are wary about the conflict of interest. Even when they do not observe the specific commissions paid by product providers, in equilibrium wary customers rationally anticipate how the higher price that they pay to the product provider is passed through into higher commissions to intermediaries, and how these commissions ultimately affect recommendations and choices. However, when some customers naively fail to adequately take into account the potentially self-interested nature of advice (as in the opening quote from the US Treasury), the fee structure that prevails in equilibrium is no longer second-best efficient. As we show, product providers are able to better exploit the mis-perceptions of naive customers by inducing a compensation structure involving a lower up-front charge for advice and a higher final price. In fact, when all customers are naive in this way, our model predicts that customers are not asked to pay any up-front charges for advice. Then, intermediaries are only compensated indirectly through the commission payments they receive from product providers.

Our model is applicable not only to settings in which a customer's choice is simply between purchasing a given product or not purchasing at all, but also to settings 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 can realize with different products. Fees paid for investment products, as well as commissions, are typically higher when they are actively managed or more innovative, as in the case of complex structured investments that are marketed to retail investors. Similarly, as noted previously, particular types of mortgage contracts yield higher margins than more conventional contracts.

In equilibrium, naive customers underestimate the likelihood with which they ultimately purchase a "premium" product (or a product at all) that generates higher profits for the respective financial institutions and for the intermediary than a more "basic" offering (or no purchase). As we show, this misperception is maximally exploited by charging customers no fee at all for advice, while advisers are paid exclusively through commissions or distribution fees. Even though customers appear not to pay for advice, in reality they are seriously short-

3

changed through biased advice and higher product prices, in the form of higher management fees on investment products or higher interest rates on mortgages.

In the presence of naive customers, there are two benefits of policy intervention that requires firms to make customers pay directly for advice. First, we show that a cap (or, ultimately, a ban) on commissions or other inducements increases consumer surplus by restricting the extent to which customers' naive beliefs can be exploited. Second, when advisers have little independent interest that customers make more suitable choices (e.g., because they have limited reputational capital at stake), overall social efficiency is higher when such inducements are banned and intermediaries must then be compensated directly by customers.

Policy intervention can, however, backfire when the practice of paying "indirectly" for advice arises in the presence of wary customers, who see through the incentives of financial institutions and intermediaries. With wary customers, we highlight an efficiency rationale for compensating intermediaries also through commissions paid by product providers. Even though indirect pay for advice leads to biased advice, the overall quality of advice that results may be higher because the adviser's incentives to acquire information are improved. It may thus be efficient *not* to perfectly align, at the recommendation stage, the interests of advisers with those of wary customers. Specifically, even when customers are wary of the conflict of interest and presence of commissions, we show that high commissions result for products that are likely ex ante to suit the preferences and needs of only a small fraction of customers. Intuitively, in the absence of commissions, the adviser would have little incentive to learn whether such products are indeed suitable for a customer.

We also discuss the policy option of mandating disclosure of a potential conflict of interest between customers and their adviser. 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 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.[8]

Overall, our analysis suggests that policy intervention should be sensitive to the types of customers that purchase particular retail financial products through a particular distribution

---

[8]In the US, the Federal Trade Commission (2008) has proposed rules that would require that brokers enter in an initial agreement with customers 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."

channel. To what extent can customers be reasonably expected to be sufficiently wary of the conflict of interest when their advisers are paid through commissions or other inducements? The form of naiveté about incentives that we posit is similar to the one documented empirically by Malmendier and Shanthikumar (2007) in the context of recommendations made by security analysts to investors. Experiments with games of trust and cheap talk also suggest that many subjects are willing to follow advice more than they should even when payoffs and incentives are revealed (e.g., Gneezy (2005) and Cain, Loewenstein, and Moore (2005)).[9] Using data from the Survey of Consumer Finances, Bergstresser and Beshears (2010) show that borrowers who were less able to comprehend financial questions and who were less suspicious in interviews were more likely to purchase adjustable-rate mortgages in the period 2004–2007; these mortgages then exhibited higher rates of foreclosure than fixed rate mortgages during the mortgage crisis. Hackethal, Inderst, and Meyer (2010) document that retail customers of a German bank who report to strongly rely on advice purchase a substantially larger fraction of securities for which their advisers were incentivized and which generate higher revenues.

To the fledgling literature on consumer financial protection, we contribute a positive and normative analysis of the compensation structure for advice. Other recent contributions in the area focus on different aspects relevant to the provision of non-verifiable information to customers.[10] Bolton, Freixas, and Shapiro (2007) analyze how incentives for information provision depend on competition among banks. Carlin and Gervais (2009) focus on the effect of the legal rules determining how liability is split between the product provider and the broker. Inderst and Ottaviani (2009) focus on the multi-task agency problem a seller faces when hiring an agent to find as well as to advise customers. Inderst and Ottaviani (2010a) analyze competition through commissions as well as through prices among multiple product providers in a common agency framework.

The quality of advice in our model depends both on the quality of information that advisers gather and on the use that they make of this information to provide potentially

---

[9] In Cain, Loewenstein, and Moore (2005) subjects are paid for the precision of the estimates of the number of coins in a jar. Subjects 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% higher in the second treatment.

[10] Earlier papers, such as Admati and Pfleiderer (1986), analyze how a seller should optimally charge for information when its quality can be verified by customers.

compromised recommendations. In the context of security analysts, Lim (2001) argues that analysts who are favorably biased are able to obtain more accurate information from the management of the firm they cover; thus analysts acquire accuracy through bias. In our model, instead, the bias induced by commissions increases the adviser's incentives to acquire costly information.

In an early contribution cast in the context of insurance markets, Gravelle (1994) also analyzes the compensation structure of brokers. In Gravelle's (1994) model, however, brokers truthfully reveal to customers the valuation for the product, so that the choice between up-front payment and commission trades off two monopoly-pricing problems; the up-front payment reduces the number of customers who become informed, whereas the commission charge reduces the number of informed customers who actually purchase the insurance product. Gravelle (1993) captures the activity of insurance brokers with respect to unsophisticated customers through an upward shift in demand. This is reminiscent of the recent analysis by Stoughton, Wu, and Zechner (2008), where intermediaries can be incentivized to market more aggressively investment products to unsophisticated investors. In their analysis of delegated investment management, kickbacks paid by portfolio managers to intermediaries enable investment fund managers to price discriminate across investors with more or less wealth.

In our model, indirect compensation for advice is at once a hero (improving the seller's incentives to acquire information) and a villain (allowing firms to exploit naive customers). Our exploitation result is reminiscent of DellaVigna and Malmendier (2004). While in their model customers are naive about their future demand, in our model customers are naive about the incentives behind the advice received. Given that incentives are endogenously determined in our model, firms exploit the customers' naiveté by enhancing the conflict of interest through commissions. Finally, also Carlin (2010) considers customers with varying degrees of sophistication; in his model, however, sophisticated customers are able to observe individual prices, while non-sophisticated customers purchase randomly.

The paper proceeds as follows. Section I introduces the model. Section II analyzes the adviser's optimal strategy to acquire information and provide advice. Sections III and IV solve for the equilibrium in the presence of naive and wary customers, respectively. Section V analyzes the implications for consumer surplus and social welfare when commissions are not allowed, so that advisers are required to charge customers directly for advice, or when there is mandatory disclosure of commissions. Section VI analyzes the effect of competition.

Section VII extends the analysis to allow for a mixed population of naive and wary customers. Section VIII concludes.

I. Model

We are interested in analyzing some generic features of the market for many retail financial services, such as investment products, pension plans, mortgages, and life insurance policies. Thus, as in the theoretical work by Bolton, Freixas, and Shapiro (2007), Carlin (2010), and Carlin and Gervais (2009), we abstract from specific features of markets for particular products and services. Instead, we frame our analysis more generally in terms of a customer's choice between two options. This choice is based on an adviser's recommendation regarding the suitability of the characteristics of either option to the customer's specific needs and circumstances, such as the customer's wealth, earnings prospects, age, risk attitude, and tax status. For instance, the attractiveness of a fixed rate mortgage (FRM) relative to an adjustable rate mortgage (ARM) depends on the indexation of the borrower's income stream, among many other factors reviewed by Campbell, Jackson, Madrian, and Tufano (2010). A household's optimal choice of pension scheme, in terms of risk and liquidity, depends on factors such as age to retirement and risk tolerance given the composition of the household's asset portfolio.[11] Similarly, the tax implications of different investment vehicles, such as stocks and municipal bonds, depend on an investor's tax bracket. To focus our analysis, we largely abstract from specific characteristics of particular markets for retail financial services, such as the degree of competition (to which we turn in Section VI). We consider a setting with three strategic players: the provider of a "premium" product, an intermediary who provides advice, and a customer.

Products and Customer Preferences.   We denote the customer's options by $\theta = A, B$, 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 the customer, an adviser, and the seller of product

---

[11]In a recent review of the advice provision for personal pension plans, the UK Financial Service Authority (2010) reported many instances of advised pension switches that were unsuitable given customers' attitude to risk, often in addition to involving an inappropriate loss of benefits from the ceding scheme.

A. For instance, option *B* may represent the option of not investing or that of investing in Treasury bills, while option *A* may represent a mutual fund or a structured product. Alternatively, *B* could be a plain vanilla mortgage (such as an FRM) and *A* a more innovative arrangement.

The price of product *A*, p_A, is chosen by the respective product provider. It may represent management fees or interest. Instead, what the customer pays when choosing the alternative option, *B*, is exogenously given. In case this represents an alternative product, rather than the option of not purchasing at all, then we may suppose that the respective price p_B is determined competitively, in which case it is equal to the respective cost of providing the product. To streamline the notation we set equal to zero all costs, i.e., both the cost of providing each product and the cost of administering a purchase.

The value realized by a customer depends on the match between the customer's preferences and needs with the characteristics of the options available. We capture the importance of the match 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 $v_{B,B} > v_{A,B}$. We impose symmetry by supposing that $v_{A,A} = v_{B,B} = v_h$ and $v_{A,B} = v_{B,A} = v_l$, with $v_h > v_l$, and we define $\Delta_v := v_h - v_l$.

The initial (or prior) public probability that choice *A* is more suitable is equal to q_0. The customer's expected gross payoff is then $v_l + q_0\Delta_v$ when choosing *A*, and it is $v_l + (1 - q_0)\Delta_v$ when choosing *B*. We assume that the basic option is more suitable for the average customer, $q_0 < 1/2$, so that the advanced option constitutes a niche market.

Advice. The customer can obtain advice from an intermediary agent (the adviser) who, by spending time and effort, can become more familiar with the customer's specific circumstances and needs. Denote the adviser's (privatively observed) effort by $e \geq 0$, which incurs costs $\kappa(e)$, where we stipulate that $\kappa(0) = 0$, $\kappa'(0) = 0$, $\kappa'(e) \geq 0$ for all e, and $\kappa(e) \to \infty$ as $e \to \infty$.[12] To model the resulting quality of the adviser's information, we exploit the binary structure of the match quality. Note first that any (additional) information that the adviser observes gives rise to some posterior belief, denoted by q, that product *A* provides a better match (i.e., that $\widehat{\theta} = A$). We characterize the quality of the adviser's information by

---

[12]Even when the time spent with customers was observable and contractible, it would be difficult to verify how hard the adviser tries to find out the best match.
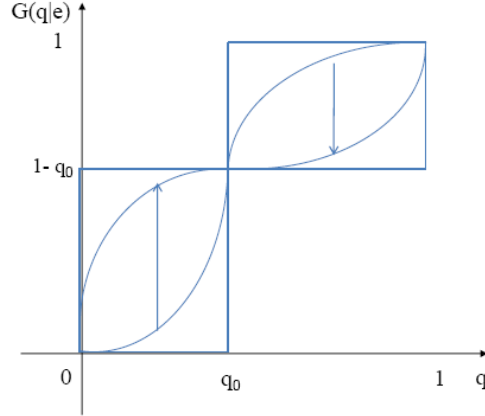
Figure 1. Information Quality. An increase in information acquisition effort e rotates the distribution function of the adviser's belief $G(q|e)$ clockwise. The distribution is shifted upward (respectively, downward) for beliefs below (respectively, above) the prior probability q_0.

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 adviser's posterior belief, $G(q \mid e)$, by inducing a mean-preserving rotation of $G(q \mid e)$, around the prior belief, q_0:

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

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. Note that we have $\frac{d}{de} \int_0^1 q dG(q \mid e) = 0$ because the expected posterior is equal to the prior by the martingale property of beliefs.

To understand the importance of condition (1), consider the extreme cases with no information and perfect information. When the adviser has access to no information, the adviser's posterior belief is always equal to the prior q_0; in this case the distribution is equal to zero for $q < q_0$ and to one for $q \geq q_0$. When the adviser has access to perfect information, the adviser's posterior belief is equal to $q = 0$ with probability $1 - q_0$ and to $q = 1$ with probability q_0; in this case the distribution is equal to $1 - q_0$ for $q < 1$ and to one for $q = 1$. As can be seen in Figure 1, the perfect information distribution is a clockwise rotation of the no-information distribution. According to condition (1), an increase in information quality results in a clockwise rotation of the distribution. Given our dichotomous structure with two states, $\widehat{\theta} = A, B$, any signal structure that results in the described rotation of the posterior

9

distribution is more informative in the sense of Blackwell, as shown by Ganuza and Penalva's (2009) Theorem 2. This way of capturing the quality of the adviser's information is thus both intuitive and general.

Adviser Preferences. The adviser is concerned about the suitability of the option chosen by the customer. We capture this concern by stipulating that the adviser's payoff is reduced by $\rho > 0$ when the customer ultimately realizes low utility v_l instead of high utility v_h. Even though the respective levels of the adviser's payoff is inessential for our analysis, for concreteness we specify that the adviser's payoff, gross of payments received from product providers, is equal to τ_l when v_l is realized and equal to τ_h when v_h is realized, so that $\rho = \tau_h - \tau_l$. This simple way of modeling the suitability concern follows Bolton, Freixas, and Shapiro (2007) and Inderst and Ottaviani (2009).[13]

The adviser's concern for suitability may have different origins. The adviser may simply have professional concerns about a customer's well-being. There may also be reputational costs, e.g., through the loss of future business with this or other customers.[14] Further, ρ may capture the prospect of prosecution by courts or regulators following customer complaints regarding suitability or a review of past sales by supervising authorities. To be specific, we suppose that ρ represents a fine paid to regulators.[15]

To obtain a unique solution for the choice of information quality we further assume that

$$k''(e) > \rho \max_{q \in [0,1]} \left| \frac{d^2 G(q \mid e)}{de^2} \right| \tag{2}$$

for all e, so that concavity of the maximization program is guaranteed. Without this additional assumption the equilibrium information quality need not be unique. However, standard monotone comparative statics methods can be used to extend our results also when this additional concavity assumption does not hold.

---

[13] While this model takes ρ as exogenous, Inderst and Ottaviani (2010a) endogenize the suitability concern through a simple dynamic model in which ρ captures the expected value of future income lost following an unsuitable sale. Alternatively, see Inderst and Ottaviani (2010b) for a model in which the penalty for selling to an unsuitable buyer is endogenized set by the seller in a commitment stage in which the terms for contract cancellation and return are determined.

[14] For reputational costs see also Durbin and Iyer (2009).

[15] 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 (see 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.

Contracting.   There are five periods.

At time $t = 1$, product provider A determines the respective price $p_A = p$ and, at the same time, offers a contract to the adviser. The contract offered to the adviser prescribes two elements, a fixed payment T and a conditional payment t that is paid only when subsequently product A is sold. We do not place any sign restrictions on the payments to the adviser. In particular, the specification that T can be negative ("deep pockets" of the adviser) allows the product provider to perfectly solve the internal agency problem in the distribution chain. This specification allows us to focus on the contracting problem with respect to customers.[16] We return to this issue after characterizing the equilibrium.

Note also that the adviser does not receive additional payment when option B is chosen, so that either no purchase is made or the basic (and competitively provided) product is purchased. It is, however, straightforward to extend the analysis to allow for payments to the adviser that would be made to cover any administrative or handling costs. After all, what will matter for our analysis is the *difference* in the payments that the adviser receives when the customer makes the respective choice.

The contingent payment t may take different forms in practice. For some investment products, the respective broker or independent financial adviser may receive all or a fraction of the "load" that the customer initially pays to the product provider. More generally, the intermediary may receive a commission. With credit products, brokers' compensation is often tied to the interest rate (through the so-called "yield spread," see Jackson and Burlingame (2007)). Sellers of life insurance plans may be paid both up-front or via a trail-commission over the duration of the contract (see Cummins and Doherty (2006)).

At $t = 2$, provided that the adviser has accepted the product provider's offer (T, t), it is the adviser's turn to stipulate a flat fee f for advice. We restrict this fee to be nonnegative, $f \geq 0$ according to a standard "no free lunch" condition that prevents the adviser from bribing the customer into business with a positive 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. Only when the adviser's offer is accepted by a customer, who arrives next, does the game proceed.

---

[16]See Inderst and Ottaviani (2009) and (2010a) for models in which the agent is subject to wealth liability constraints and thus obtains a rent in equilibrium. In our model, instead, we abstract from the agency problem by allowing efficient contracting between the product provider and the adviser.

At $t = 3$, the adviser 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 adviser recommends to the customer which option to choose. The game at this stage is one of cheap talk (cf. Crawford and Sobel (1982)). As we show below, the customer follows the adviser's recommendation in the only informative equilibrium.[17]

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.

Customer Rationality. Our analysis distinguishes between two types of customers, wary and naive. Wary customers are perfectly aware of the adviser's incentives arising both from his concern for suitability ρ and from the contingent payment t that is made by provider A. To be specific, we suppose that the contract between the adviser and the product provider A is not disclosed to the customer (cf., however, the discussion in Section V). A wary customer, nevertheless, forms rational beliefs. On the other hand, naive customers mistakenly believe that the quality of advice is not affected by the presence and the size of payments made by product providers. Consistent with evidence discussed in the introduction, naive customers do not take into account the impact of such payments on the adviser's incentives.

In the analysis that follows, we solve the game backwards. We begin by supposing that at $t = 5$ the customer follows the adviser's recommendation, which we then show to be consistent with equilibrium behavior. In Section II we turn to stages $t = 3$ and $t = 4$ of the game and solve for the adviser's optimal choice of information acquisition and recommendation; this analysis applies both to the scenario with wary and naive customers. Customer rationality plays a key role in the contracting stages at $t = 1$ and $t = 2$. Thus we proceed in Section III by solving the game under the specification that the customer is (known to be) naive, and then turn to the case with wary customer in Section IV. Our specification of general two-part contractual transfers (f, p) between the product provider and the adviser allows the product provider to extract the adviser's rent, thus ensuring that what drives behavior is the overall payoff in the vertical supply structure. Similarly, by allowing for general two-part contractual transfers (T, t) between the vertical structure

---

[17] 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.

and the customer we obtain that the consumer surplus is fully extracted in equilibrium. In Section VII we analyze the general case of a population composed of a fraction of wary customers and a fraction of naive customers.

II. Provision of Advice

Given the realization of a posterior belief q (that product A is more suitable), at $t = 4$ it is optimal for the adviser to recommend product A whenever the adviser obtains a higher expected payoff when the customer purchases purchases A rather than product B, i.e., when $t + q\tau_h + (1 - q)\tau_l \geq q\tau_l + (1 - q)\tau_h$. The adviser thus considers not only the monetary inducement t in case of recommending product A, but also the expected private costs of a subsequent mismatch, which are equal to $(1 - q)\rho$ for A and $q\rho$ for B after substitution of $\rho = \tau_h - \tau_l$ If interior, the recommendation is characterized by a cutoff

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

so that the adviser strictly prefers to recommend A when $q > q^*$ and strictly prefers to recommend B when $q < q^*$. Note that the cutoff is not interior when $t \geq \rho$, in which case the adviser always recommends product A; for this case, we specify $q^* = 0$.

Before providing a recommendation at $t = 4$, at $t = 3$ the adviser can exert effort to improve the quality of information. The adviser optimally chooses effort e to maximize the expected payoff $\pi - \kappa(e)$, where we have

$$\pi := f + T + \tau_l + \int_0^{q^*} \rho(1 - q)dG(q \mid e) + \int_{q^*}^1 [t + \rho q] \, dG(q \mid e), \tag{4}$$

using again $\rho = \tau_h - \tau_l$. When $q^* = 0$, so that the adviser always recommends A, then clearly $d\pi/de = 0$, so that the adviser has no incentive to exert effort. When, instead, $q^* > 0$ is determined by (3), expression (4) transforms to

$$\pi = [f + T + t + \tau_l + \rho q_0] + 2\rho \int_0^{q^*} G(q \mid e)dq, \tag{5}$$

after integrating by parts, substituting for q^*, and using the martingale property of beliefs, $\int_0^1 G(q \mid e)dq = 1 - q_0$. Expression (5) has a simple interpretation. The first term, which is put in brackets, is equal to the expected payoff the adviser would obtain by *always* recommending option A. Note that this would allow the adviser to obtain for sure the commission t. The second term in (5) denotes the benefits, in terms of lower expected mismatch costs, when the customer makes a more informed decision based on the advice received.

13

When q^* is interior, then from (5) an optimal choice of effort solves the first-order condition

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

(6)

For all interior q^* the left-hand side of (6) is clearly strictly positive, because the adviser cares about suitability ($\rho > 0$). The maximizing level of effort e^* is unique by our concavity assumption (2), and it is strictly positive by $\kappa'(0) = 0$.

Proposition 1 summarizes the characterization of the adviser's decision problem.

PROPOSITION 1: *In case the customer follows the adviser's recommendation, then for given contracts the equilibrium at stages 3 and 4 is characterized as follows. If the payment t that the adviser receives from the product provider when the customer purchases product A is large ($t \geq \rho$), then the adviser exerts no effort to acquire information and always recommends product A. Instead, when $t < \rho$ holds, the adviser recommends B when $q \leq q^*$ and A when $q > q^*$, where the cutoff $0 < q^* < 1/2$ is determined by (3); and the adviser chooses strictly positive effort $e^* > 0$, which solves (6).*

III.   Exploiting Naive Customers

Naive customers do not take into account how the adviser's recommendation is affected by payments made by product providers. Instead, they invariably interpret advice based on the mistaken belief that the adviser is exclusively motivated by suitability, ρ. Consequently, naive customers expect the adviser to choose the symmetric (and thus unbiased) cutoff, $q^* = 1/2$. Based on these beliefs and using the adviser's expected payoff in (5), naive customers also anticipate that the adviser exerts acquisition information effort $\widehat{e}_N > 0$ solving

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

(7)

Note that we use here the notation $\widehat{e}_N$ to denote naive customers' *beliefs,* in contrast to the truly chosen level of effort, which we characterize in what follows. To avoid confusion, it is convenient to denote by $\widehat{q}_N = 1/2$ the recommendation cutoff that naive customers expect the adviser to apply. Note also that naive customers only possibly err with respect to the incentives of the adviser. Based on these beliefs, however, they form rational beliefs with respect to the adviser's actions: $\widehat{q}_N$ and $\widehat{e}_N$.

A. The Contracting Problems

Recall that the strategic product provider A chooses both the price p that is charged to the customer and, at the same time, the two-part contract that is offered the adviser, (T, t). After accepting this contract, the adviser is free to specify a fee f that customers have to pay before receiving advice and possibly purchasing a product. For naive customers, the expected value of advice depends only on their beliefs about the adviser's choice, $\widehat{q}_N$ and $\widehat{e}_N$. These beliefs are in turn independent of any observed contractual choices, notably p and f.

Customer Participation Constraint. For a customer who chooses not to obtain advice it is optimal to always choose option B and, thereby, realize net utility

$$v_0 := v_l + \Delta_v(1 - q_0) > 0. \tag{8}$$

For this, we use our simplification that in case option B consists of buying an alternative product B, this product is competitively provided at a price equal to its zero cost. From $v_0 > 0$ we also have that the customer will always follow a recommendation to purchase B: For any cutoff $q^* > 0$ the expected utility *conditional* on $q > q^*$ is strictly higher than v_0. Note that this holds, in particular, when a naive customer expects the adviser to apply the cutoff $\widehat{q}_N = 1/2$. Instead, the customer's incentives to follow the recommendation to purchase product A depend on the prevailing price p. Note that at this stage, the customer conditions on the information contained in the recommendation, based on the conjecture that $q > \widehat{q}_N$. Thus, the naive customer follows a recommendation to purchase A if the corresponding conditional payoff is higher than the one obtained from product B

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

which simplifies to the requirement that

$$p \leq \Delta_v \int_{\widehat{q}_N}^1 (2q - 1) \frac{dG(q \mid \widehat{e}_N)}{1 - G(\widehat{q}_N \mid \widehat{e}_N)}. \tag{9}$$

Next, a customer will optimally only be willing to pay a fee $f \geq 0$ up-front if the respective expected payoff exceeds that from not obtaining advice:

$$v_l + \Delta_v \int_0^{\widehat{q}_N} (1 - q) dG(q \mid \widehat{e}_N) + \int_{\widehat{q}_N}^1 [q\Delta_v - p] \, dG(q \mid \widehat{e}_N) - f \geq v_0. \tag{10}$$

Substituting for the customer's outside option v_0 from (8), the ex ante participation con-

straint (10) becomes

$$p + \frac{f}{1 - G(\widehat{q}_N \mid \widehat{e}_N)} \leq \Delta_v \left[\int_{\widehat{q}_N}^{1} (2q - 1) \frac{dG(q \mid \widehat{e}_N)}{1 - G(\widehat{q}_N \mid \widehat{e}_N)} \right]. \tag{11}$$

Given that $f \geq 0$, we thus conclude that this ex ante constraint implies the ex post constraint (9). As is intuitive, a customer who would optimally not follow the recommendation to purchase product A would clearly not be willing to pay a fee f to receive such advice. Hence, we need only consider for the customer the ex ante participation constraint (11).

Adviser's Contract Design Problem. Next, turn to the optimal contracting problem of the adviser. Provided that the adviser has accepted the contract offered by the product provider, at $t = 2$ the adviser must specify the up-front fee that the customer will have to pay. If a positive fee $f \geq 0$ exists for which the customer's ex-ante participation constraint (11) is satisfied, the adviser will optimally set the fee at the highest possible level. The binding constraint (11) then pins down a unique fee, which for naive customers only depends on p. We denote it by $f_N(p)$.

Product Provider's Contracting Problem. Product provider A determines both the price p and the bilateral contract with the adviser (T, t). When the adviser accepts, the expected profits of A equal

$$\Pi = [1 - G(q^* \mid e^*)] (p - t) - T, \tag{12}$$

where q^* and e^* are the respective unique optimal choices of the adviser, as determined in Proposition 1. As long as p satisfies (9), the adviser's optimal choice $f = f_N(p)$ will ensure that the customer participates by paying the fee. Subsequently, the customer purchase product A with probability $1 - G(q^* \mid e^*)$.

In addition, product provider A must ensure that also the adviser accepts the contract. When rejecting the contract, the adviser realizes zero payoff. Thus, the adviser's participation constraint is given by $\pi - \kappa(e^*) \geq 0$, where e^* denotes the optimal choice of effort (cf. Proposition 1) and π is determined from (4) once we substitute $f = f_N(p)$. By optimality the fixed part T is chosen so as to make the adviser just indifferent between accepting and rejecting. Denote the resulting value for T that satisfies with equality $\pi = \kappa(e^*)$ by $T_N(p)$. After substitution, profits (12) become

$$\Pi = f_N(p) + \tau_l + [1 - G(q^* \mid e^*)] p + \rho L - \kappa(e^*) \tag{13}$$

where

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

denotes the ex ante probability of a *suitable* choice.

To interpret expression (13), note again that through the fixed part T the product provider can extract all profits from the adviser, so that Π in (13) comprises the total surplus that is realized jointly by the product provider and the adviser. As noted previously, by allowing for a fixed part T we can thus abstract from contracting imperfections between product providers and the adviser.

There is an immediate implication from this observation. When we take as given some value of p, then to maximize profits Π in (13) it must hold that $t = p$. Likewise, any change in the price p would be optimally reflected in a one-for-one change of the commission t, as then the incentives of the adviser and that of the product provider are still perfectly aligned. 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. For example, the payments made to mortgage brokers are typically 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; see the discussion in Jackson and Burlingame (2007) or Keith, Bocian, and Li (2008). In the insurance market, Cummins and Doherty (2006) provide evidence that premium-based and contingent commissions paid to brokers and other intermediaries are passed on to policyholders in the premium.

Summarizing the observations reported above, we have the following result:

LEMMA 1: *The contracting problem when customers are naive is characterized as follows. The product provider chooses a price p and a corresponding commission $t = p$ to maximize profits Π in (13). The subsequently determined fixed fee for advice $f = f_N(p)$, where we use naive customers' beliefs $\widehat{q}_N$ and $\widehat{e}_N$, makes customers just indifferent between paying the fee or not receiving advice, as given by constraint (11). Finally, the adviser's expected payoff is reduced to zero through the fixed part $T = T_N(p)$.*

It should be noted that our analysis does not hinge on the specification that under the optimal contract the adviser receives a payoff of zero. In fact, with the exception of the specification of the fixed part T, the characterization in Lemma 1 is not affected by how the surplus is shared between product providers and the adviser. Formally, instead of

having product provider A make a take-it-or-leave-it offer, we could also specify a game of bargaining. As noted previously, the *contingent* payment t would then still be chosen to align the interests of the product provider and the adviser, while the fixed part T would be adjusted to reflect the distribution of bargaining power.

B. Characterization

We now turn to solve the remaining contracting problem that consists in the choice of product A's price p, from which the commission $t = p$ and the respective fixed fee that naive customers pay for advice, $f = f_N(p)$, follows from Lemma 1.

As a starting point, suppose that firms extract customer surplus *only* through a fixed fee, by setting the product price for A equal to cost: $p = 0$. By optimality, this implies that also the commission paid to the adviser is zero (cf. Lemma 1) and that, consequently, we have the recommendation cutoff $q^* = 1/2$ by Proposition 1. This is equal to the cutoff that naive customers anticipate: $q^* = \widehat{q}_N = 1/2$. Also the adviser's choice of effort is then equal to that anticipated by customers: $e^* = \widehat{e}_N$. The respective fee for advice that can then be charged is $f_N(p = 0) = \Delta_v \int_{\widehat{q}_N}^1 (2q - 1) dG(q \mid \widehat{e}_N)$, at which the participation constraint (11) binds, and naive customers' expected surplus is v_0, as given by (8). However, we now show that firms can increase profits at the expense of reducing naive customers' true surplus by raising the price p, together with the commission t, while lowering the fixed fee for advice f. To this end, the following result is instrumental:

LEMMA 2: *Taking into account the thereby induced changes in q^* and e^*, as given by Proposition 1, an increase in the commission t induces an increase in the probability $1 - G(q^* \mid e^*)$ with which the customer is recommended to purchase product A. This increase is strict until $q^* = 0$.*

Proof of Lemma 2. When q^* is interior, we have for $q^* > 0$ that

$$\frac{d}{dt}[1 - G(q^* \mid e^*)] = -\frac{dq^*}{dt}\left[g(q^* \mid e^*) + \frac{dG(q^* \mid e^*)}{de^*}\frac{de^*}{dq^*}\right]. \tag{14}$$

To determine the sign of (14), recall first that $dq^*/dt < 0$ by (3). Next, from implicit differentiation of (6) we obtain

$$\frac{de^*}{dq^*} = \frac{-2\rho}{SOC}\frac{dG(q^* \mid e^*)}{de^*}, \tag{15}$$

where $SOC < 0$ denotes the second-order condition for e^*. (Recall that we stipulated that

18

the advisor's program to choose e^* yields a unique solution, which for $0 < q^* < 1$ is strictly positive.) The sign of the second term in (14) is then given by $\left(\frac{dG(q^*|e^*)}{de^*}\right)^2$, which is also strictly positive. Thus, (14) is strictly positive. **Q.E.D.**

An increase in the commission for selling product A results in a reduction of the cutoff q^*. From Lemma 2 the respective probability with which product A is recommended, $1 - G(q^* \mid e^*)$, is even higher when we take into account the adjustment of the information acquisition effort e^* optimally chosen by the adviser.

Suppose now that product provider A increases the product price and, by optimality, also the respective payment to the adviser. Once we substitute $f_N(p)$ into the firm profits Π in (13), we obtain the marginal change

$$\frac{d\Pi}{dp} = G(\widehat{q}_N \mid \widehat{e}_N) - G(q^* \mid e^*), \tag{16}$$

which for $p = t = 0$ (such that $\widehat{q}_N = q^* = 1/2$ and $\widehat{e}_N = e^*$) is zero, but it is strictly positive for all $p = t > 0$. Hence, the considered marginal increase in the product price and in the commission, together with a reduction in the direct fee for advice, increases profits. The unique optimal choice will then imply that customers are charged *no* direct fee for advice, $f = 0$. We now offer some intuition for this result.

When naive customers observe a higher price for product A, they do *not* rationally anticipate that product provider A will also increase its commission to the adviser *and* that the adviser will then optimally adjust the information acquisition effort and recommendation strategy. In particular, according to Lemma 2, a naive customer underestimates the probability of receiving a recommendation to buy the now more expensive product A. In fact, as the customer still expects that the recommendation to buy A happens only with probability, $1 - G(\widehat{q}_N \mid \widehat{e}_N)$, the difference in purchase probabilities (i.e., the statistical error that is made) is exactly equal to the difference $G(\widehat{q}_N \mid \widehat{e}_N) - G(q^* \mid e^*)$ in expression (16). This observation is key. Profits thus strictly increase whenever the up-front payment for advice is reduced, provided that the participation constraint of the naive customer is still satisfied. This strict monotonicity holds because of the exploitation of the naive customer's beliefs, which are wrong whenever $t > 0$.

More formally, expression (16) uses also the fact that the interests of the product provider and the adviser are aligned through the use of the two-part contract (T, t). This ensures that, following a change in $p = t$, both q^* and e^* are still chosen optimally to maximize Π. Formally, this allows us to invoke the envelope theorem, so that the marginal impact of

the considered change is indeed equal to the difference between the true probability and the naively anticipated probability that product A is recommended.

Once we substitute $f = f_N^* = 0$ into the naive customer's binding ex-ante participation constraint (11), we obtain for the corresponding equilibrium product price

$$p = p_N^* = \Delta_v \left[\int_{\widehat{q}_N}^{1} (2q - 1) \frac{dG(q \mid \widehat{e}_N)}{1 - G(\widehat{q}_N \mid \widehat{e}_N)} \right]. \tag{17}$$

We have established the following result.

PROPOSITION 2: *In equilibrium, naive customers are not charged directly, so that $f = f_N^* = 0$. The corresponding price p_N^* of product A is given by (17), and the respective equilibrium choices of the adviser, the adviser's information acquisition effort e_N^* and the advice cutoff q_N^*, are obtained from substituting $t_N^* = p_N^*$ into the characterization in Proposition 1.*

Discussion. With naive customers, Proposition 2 thus offers a possible rationale for why frequently retail financial customers do not pay directly for financial advice. Firms generate higher profits when, in equilibrium, naive customers underestimate the true probability with which they will subsequently be advised to purchase the respective product. This makes it optimal to reduce the up-front fee as much as possible, while raising the price p and the commission t.

Note finally that with naive customers advice may become completely uninformative, namely when $q_N^* = 0$ and, consequently, also $e_N^* = 0$ result. Then, the adviser always recommends product A and does not acquire any information. From Proposition 1 this is the case when

$$\rho \leq \Delta_v \left[\int_{\widehat{q}_N}^{1} (2q - 1) \frac{dG(q \mid \widehat{e}_N)}{1 - G(\widehat{q}_N \mid \widehat{e}_N)} \right]. \tag{18}$$

Even though ρ also affects the naive customers' expectation of $\widehat{e}_N$, which is strictly increasing in ρ, condition (18) is always satisfied when the adviser's concern for suitability ρ is sufficiently small.

IV. Serving Wary Customers

In contrast to naive customers, wary customers take into account the bias in advice that results when the adviser receives contingent commission payments from product providers. Even though these commissions are not transparent, wary customers hold rational beliefs

and thus correctly anticipate that a higher price p for product A will translate into a higher respective commission t. In this section, we analyze the contracting problem when firms face wary customers.

A.   Contracting

Denote a wary customer's expectations of the adviser'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 $f \geq 0$ it is immediate that we only have to consider wary customers' *ex ante* participation constraint:

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

Given that the following argument is similar to that with naive customers, we can be brief. Note first that by optimality it still holds that $t = p$. To align interests, the product provider will optimally set the commission perfectly in line with the product price paid by customers. Consequently, the rationally anticipated choices $\widehat{e}_W$ and $\widehat{q}_W$ are equal to the respective optimal choices e^* and q^*, as given by Proposition 1. Second, the participation constraint (19) binds because otherwise the adviser would find it optimal to increase the required fixed payment f. Denote the respective value by $f_W(p)$, where importantly, in contrast to a naive customer, a change in the price p and thus also in the anticipated commission leads to a change in the anticipated choices $\widehat{e}_W$ and $\widehat{q}_W$. Further, the optimal fixed part $t = T_W(p)$ of the adviser's contract is obtained again from the requirement that $\pi = \kappa(e^*)$, where we use the respective up-front fee $f_W(p)$. In complete analogy to expression (13), after substitution of $T_W(p)$ we obtain the provider's profits

$$\Pi = f_W(p) + \tau_l + [1 - G(q^* \mid e^*)] p + \rho L - \kappa(e^*). \tag{20}$$

We can compare these profits resulting with wary customers with profits (13) resulting with naive customers. As is intuitive, the only difference is in the respective fees $f_W(p)$ and $f_N(p)$, which in turn are determined by the respective participation constraints. In summary:

LEMMA 3: *The contracting problem when customers are wary is characterized as follows. The product provider chooses a price p and a corresponding commission $t = p$ to maximize profits Π in (20). The subsequently determined fixed fee for advice $f = f_W(p)$, obtained*

using the wary customer's rational belief $\widehat{q}_W = q^*$ for the adviser's recommendation cutoff and $\widehat{e}_W = e^*$ for the information acquisition effort, makes customers just indifferent between paying the fee and not receiving advice, according to the binding constraint (19). Finally, the fixed part $T = T_W(p)$ in the contract reduces the adviser's expected payoff to zero.

B. Characterization

Once we substitute for $f = f_W(p)$, which makes wary customers just indifferent between accepting and rejecting, into the product provider's profits in (20), we obtain

$$\Pi = (\tau_l + v_l - v_0) + (\rho + \Delta_v)L - \kappa(e^*). \tag{21}$$

Recall that L is the probability of a suitable choice, given q^* and e^*. Expression (21) is thus simply the total surplus that is realized jointly by the supply vertical chain (product provider and adviser) and the consumer on the demand side. With wary customers, we conclude that the equilibrium choice of contracts maximizes total surplus. This is our first key observation for the case with wary customers

Suppose now for a moment that the quality of *advice* depended only on the adviser's recommendation but not on the effort choice. This would be the case when the quality of the adviser's *information* e was exogenously given, instead of being determined by his effort. Then, maximization of total surplus (21) leads to the uniquely optimal choice of $p = 0$ and $t = 0$, and thus to unbiased advice, $q^* = 1/2$. More formally, holding e fixed, this unbiased recommendation cutoff maximizes L, the likelihood of a suitable choice. Next, we show that this choice of contracts is no longer optimal when the quality of advice depends also on the endogenous quality of information, as captured by e^*.

To this end, it is useful to focus first on the adviser's incentives to acquire information. According to Proposition 1 the optimal choice e^* solves the first-order condition (6). From the rotation ordering of $G(q \mid e)$ in (1) we immediately have the following result.

LEMMA 4: *The adviser's incentives to acquire information and thus also the uniquely chosen effort e^* are hump-shaped as a function of q^* and thus also as a function of the commission t. Incentives are lowest at $q^* = 0$, which holds when $t \geq \rho$. Starting from $t = 0$ and thus $q^* = 1/2$, as t increases also incentives increase up to $t_0 := \rho(1 - 2q_0)$, where $q^* = q_0 < 1/2$. For all higher $t > t_0$, for which $q^* < q_0$, incentives are lower.*

This result is intuitive. When the adviser is *a priori* relatively sure to recommend one

option, say product A, the adviser has little incentive to acquire information, because this information is not likely to sway the recommendation and thus the customer's decision. This is the case, in particular, when q^* is close to zero. At the opposite extreme, when at the prior beliefs the adviser is exactly indifferent between recommending either option, i.e., when $q^* = q_0$, any additional information will break this indifference almost surely. The adviser's incentives to acquire information are then highest.

From Lemma 4 there are now two countervailing effects when advice becomes biased ($q^* < 1/2$) because of the payment of $t > 0$. The immediate effect is that this bias makes it less likely that the customer's choice is suitable, i.e., L decreases. The second effect is that, at least as long as still $t < t_0$, L increases as the adviser's information becomes more precise. Note now that at the unbiased recommendation cutoff, $q^* = 1/2$, the first-order effect that a reduction of the cutoff has on L is, however, zero, given that then both options are equally likely to result in a suitable choice. For all $q_0 < q^* \leq 1/2$ and thus, in particular also for $q^* = 1/2$, the effect on the adviser's quality of information is, however, strictly positive. Taken together, we thus know that L is highest when $q^* < 1/2$: Advice is most informative when it is biased. In turn, this bias requires that $p = t > 0$, so that also wary customers always pay the adviser indirectly, namely through the contingent payment that the adviser receives from product providers.

For brevity of exposition, we now assume that the program to choose q^* and thus e^* so as to maximize total surplus (21) is strictly quasiconcave. Denote the respective contract choices by $t_W^* = p_W^*$, which induce e_W^* and q_W^*.

PROPOSITION 3: *The equilibrium outcome with wary customers maximizes the total surplus of firms and customers*

$$\omega = (\tau_l + v_l - v_0) + (\rho + \Delta_v)L - \kappa(e^*). \tag{22}$$

This outcome is achieved when the adviser obtains a positive commission, $t_W^* = p_W^* > 0$, and leads to biased advice (with $q_W^* < 1/2$) but also to an overall higher quality of advice because then the adviser acquires more information than would result with zero commissions and unbiased advice ($q^* = 1/2$).

Proof of Proposition 3. From the discussion in the main text and Lemma 3, it remains to choose q^* so as to maximize the surplus ω, as given by (22), where q^* affects e^* according to (15) and where we have to take into account the constraint $f \geq 0$. Using (19), the latter

23

constraint becomes

$$\Delta_v \int_{q^*}^1 (2q-1)dG(q \mid e^*) - [1 - G(q^* \mid e^*)]\, \rho(1-2q^*) \geq 0. \tag{23}$$

Using the expression ω for the surplus in (22), we can also write the optimization problem with respect to the cutoff q^* as follows:

$$\frac{d\omega}{dq^*} = (\rho + \Delta_v) \left[\frac{dL}{dq^*} + \frac{de^*}{dq^*} \frac{dL}{de^*} \right] - \kappa'(e^*) \frac{de^*}{dq^*} = 0.$$

As e^* maximizes the adviser's payoff, so that $\rho \frac{dL}{de^*} = \kappa'(e^*)$, this becomes

$$(\rho + \Delta_v) \frac{dL}{dq^*} + \Delta_v \frac{de^*}{dq^*} \frac{dL}{de^*} = 0. \tag{24}$$

Using next, after integration by parts, that

$$\frac{dL}{dq^*} = g(q^* \mid e^*)(1 - 2q^*),$$

$$\frac{dL}{de^*} = (1 - 2q^*)\frac{dG(q^* \mid e^*)}{de^*} + 2 \int_0^{q^*} \frac{dG(q \mid e^*)}{de^*} dq,$$

and substituting for $\frac{de^*}{dq^*}$ from (15), expression (24) becomes

$$\frac{d\omega}{dq^*} = g(q^* \mid e^*)(1 - 2q^*)\,(\Delta_v + \rho) \tag{25}$$

$$-\Delta_v \frac{2\rho}{SOC} \frac{dG(q^* \mid e^*)}{de^*} \left[(1 - 2q^*)\frac{dG(q^* \mid e^*)}{de^*} + 2 \int_0^{q^*} \frac{dG(q \mid e^*)}{de^*} dq \right].$$

From (1) we have $d\omega/dq^* > 0$ when $q \leq q_0$ as well as $d\omega/dq^* < 0$ at $q^* = 1/2$. As we stipulated that the program is strictly quasiconcave, there is a unique solution $q_0 < q^* < 1/2$ and a corresponding value t from (3). However, this may not be feasible when after substituting the respective values $p = t$ into the binding constraint (19) we have $f > 0$. Then, from strict quasiconcavity the unique value q^* is the lowest value satisfying $f = 0$. Finally, $q^* < 1/2$ holds also then because at $q^* = 1/2$ we have $f > 0$, together with $t = p = 0$, so that from (19) it is indeed feasible to increase p and reduce f. **Q.E.D.**

Discussion. Given that the adviser also receives a commission when recommending product A, advice is biased also with wary customers. However, there are a number of important differences to the case with naive customers.

Because wary customers correctly anticipate the increased likelihood with which they are recommended product A, their *true* participation constraint always holds. Thus, the commission payment to the adviser and the biased recommendation do not lead to customer exploitation when customers are wary.

In addition, with wary customers the recommendation cutoff is always strictly positive $q_W^* > 0$ so that advice is always informative in equilibrium. With naive customers, instead, we observed that $q_N^* = 0$ holds when ρ is sufficiently small. Finally, while with naive customers we always have the corner outcome with $f_N^* = 0$, so that customers do not pay directly for advice, this is not the case with wary customers. Even though the contractual variables are not monotonic in ρ, we can show that $f_W^* > 0$ surely holds when the adviser is not concerned too much for suitability (low ρ).[18] This is intuitive because for low ρ even a small commission leads to a large bias in the recommendation, while we know that $q_W^* > q_0$ must always hold by optimality. To ensure this, $p_W^* = t_W^*$ must then remain sufficiently small, and the adviser charges, instead, a high direct fee for advice. In the presence of both naive and wary customers, we further show in Section VII that firms can better price discriminate by charging a higher up-front payment for advice to wary customers.

Example.   To illustrate the model and to obtain some additional comparative statics results we turn to a simple parametric example. This example also allows us to show how the distribution of the adviser's posterior beliefs, $G(q \mid e)$, can be derived from a noisy signal technology.

Suppose that the adviser privately observes a signal $s \in [0, 1]$ with conditional distributions $H_A(s \mid e) = s^{e+1}$ and $H_B(s \mid e) = 1 - (1 - s)^{e+1}$ parametrized by $e \geq 0$. The adviser's posterior belief as a function of the observed signal is then equal to

$$q = \widetilde{q}(s) := \frac{q_0 s^e}{q_0 s^e + (1 - q_0)(1 - s)^e}$$

by Bayes' rule. Note that $\widetilde{q}(0) = 0$ and $\widetilde{q}(1) = 1$. Also, we may now, alternatively to the specification of a cutoff q^*, define a cutoff on the signal s^* with

$$\frac{q_0}{1 - q_0} \frac{1 - q^*}{q^*} = \left(\frac{1 - s^*}{s^*}\right)^e,$$

so that the adviser recommends A if $s \geq s^*$ and B if $s < s^*$. After some transformations, the likelihood of a suitable choice as a function of s^* is then given by

$$L = 1 - (1 - q_0)(1 - s^*)^{e+1} - q_0(s^*)^{e+1}.$$

Given that the signal has the unconditional cumulative distribution function $q_0 H_A(s \mid e) +$
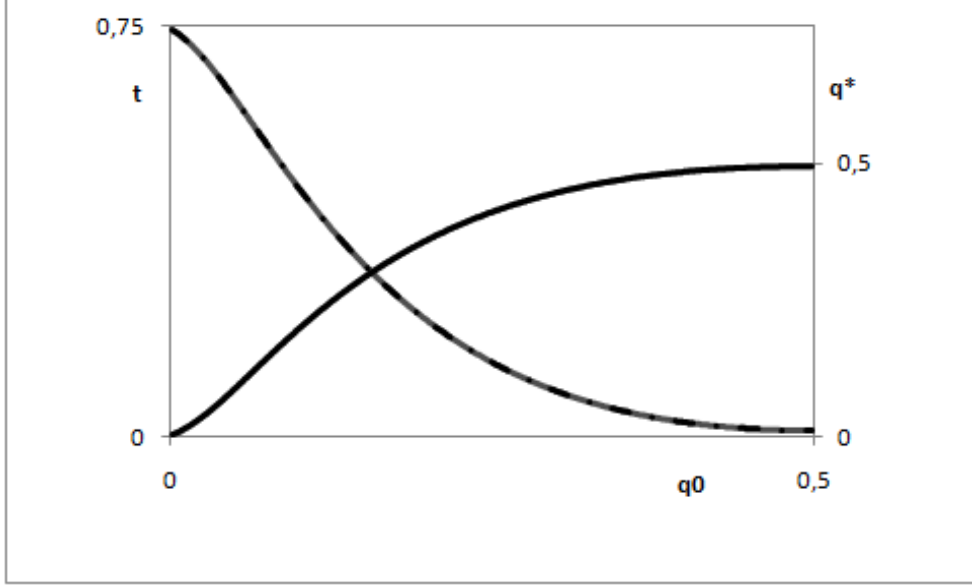

Figure 2. Commissions and Bias. For the parametric example discussed in the text, this figure reports the equilibrium level of commission t (the decreasing curve) and the equilibrium recommendation cutoff (the increasing curve) as a function of the initial probability q_0 that product A is suitable.

$(1 - q_0)H_B(s \mid e)$, we further obtain

$$G(q \mid e) = q_0 \left[\widetilde{q}^{-1}(q)\right]^{e+1} + (1 - q_0) \left[1 - \left[1 - \widetilde{q}^{-1}(q)\right]^{e+1}\right]. \tag{26}$$

It is straightforward to show that this $G(q \mid e)$ satisfies the rotation ordering (1).

For a comparative analysis we specify that the information acquisition cost is quadratic, $\kappa(e) = e^2/(2c)$ with $c > 0$. With this specification, we now analyze how the outcome depends on the likelihood with which the advanced product A is ex ante more suitable, q_0. For Figure 2 we specify $\rho = 0.75$ for the adviser's preferences, $\Delta_v = v_h - v_l = 2$ for the incremental benefits of a suitable choice, and $c = 0.65$ for the adviser's cost of effort function. As q_0 decreases, the basic option (or, equivalently, the option of not buying) is ex ante more likely to be suitable; alternatively, product A is targeted more to a niche market. As illustrated in the figure, under the optimal contractual arrangement with wary customers, the commission t paid to advisers increases and the recommendation cutoff q^* decreases when the initial probability q_0 is reduced from $1/2$ to 1. While a recommendation becomes thus more and more biased, in this example the loss in the quality of advice generated by the bias is more than compensated by the higher level of information acquisition that is thereby induced.

V. Policy Intervention

Prohibiting Commissions. Policy makers could prohibit product providers from paying commissions or other contingent payments to advisers. When $t = 0$ irrespective of the prices and thus the margins that product providers earn, advisers would have to charge customers directly for advice, so that $f > 0$. This policy represents a drastic change in some markets for retail financial services, where customers are not asked to pay directly for advice and were contingent payments made by product providers to intermediaries are common. However, a radical policy along these lines is currently being implemented in some jurisdictions, most notably by the UK's Financial Service Authority.

Mandating Disclosure of Conflicts of Interest. A policy option that is commonly adopted consists in mandating disclosure of conflicts of interest between intermediaries and customers. 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. Similarly, since January 2008 the European Union's MiFID directive imposes mandatory disclosure for the sale of many financial products.

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 making customers wary in the first place. Disclosure of a conflict of interest, even without spelling out the details of payments, may then act as an "eye opener" to previously naive customers.

To analyze the impact of these two policy interventions, we begin in Section A by considering the benchmark case in which the adviser's quality of information is exogenously fixed at a given level $e = \tilde{e} > 0$. In this benchmark, we obtain some first, clear-cut observations on the effectiveness of the two policies. Next, in Section B we derive a more nuanced set of policy recommendations for the case in which the quality of advice depends also on the adviser's incentives to generate information in the first place.

A. Benchmark: Exogenous Quality of Information

To analyze the case in which the quality of information is exogenous, let us revisit our characterization results for fixed $e = \tilde{e}$. When customers are naive, our previous results hold

without qualification. To maximally exploit naive customers, firms set $f = 0$, and advice is biased as $t_N^* > 0$ and thus $q_N^* < 1/2$. When customers are wary, recall that contracts are designed so as to maximize the joint surplus of firms and customers, which is now clearly the case when there is no bias in the adviser's recommendation, so that $t_W^* = 0$ and thus $q_W^* = 1/2$.

Prohibiting Commissions. In the absence of commissions that are contingent on the sale of a particular product, the adviser's recommendation is only affected by the concern for suitability. The resulting recommendation cutoff is then always equal to $q^* = 1/2$. When customers are wary, such a policy has no impact because $t_W^* = 0$ holds also without policy intervention when the adviser's effort is exogenous. When customers are naive, instead, a ban on contingent payments to the adviser affects the equilibrium outcome. When advisers do not receive commissions, then also naive costumers' expectations are correct in equilibrium. Their expected surplus is then equal to their reservation value v_0. Recall that naive customers' true expected surplus is strictly below v_0 without policy intervention, as their expectations are wrong in equilibrium. Hence, by protecting naive customers from exploitation, a ban on commissions strictly increases their true expected surplus. In addition, also total surplus is strictly higher, given that presently the adviser's quality of information is exogenous, the value of advice is highest when $q^* = 1/2$.

Mandating Disclosure of Conflicts of Interest. Clearly disclosure of commissions would not affect results with wary customers, given that in this case product providers pay no commissions in equilibrium, $t_W^* = 0$. With naive customers, we posit that the disclosure of a potential conflict of interest turn them into wary customers. When the size of the commission is disclosed directly, customers thus immediately respond to commissions whenever $t > 0$ by adjusting their beliefs. Consequently, with mandatory disclosure the outcome with naive customers is the same as the one resulting with wary customers. Firms facing naive customers make higher profits by exploiting customers' misperceptions and thus strictly prefer *not* to disclose their conflict of interest, even though commissions and the resulting bias in recommendations leads to strict reduction in total surplus.[19]

We now summarize our conclusions for the benchmark case.

---

[19] On firms' disincentives to educate customers see also Gabaix and Laibson (2006).

PROPOSITION 4: *Consider the benchmark case where the quality of the adviser's informa-tion is exogenous ($e = \widetilde{e}$). With wary customers, prohibiting commissions or mandating disclosure of commissions has no effect on the equilibrium outcome. With naive customers, instead, both policies lead to the same outcome in which customer surplus and total surplus are both strictly higher than without policy intervention.*

B. Policy Analysis with Endogenous Quality of Information

Part of the modelling innovation in this paper is to endogenize the quality of the adviser's information. The overall quality of advice then depends on both the adviser's incentives to generate information and the incentives to provide unbiased recommendations. We analyze now the impact of policy intervention in this more realistic scenario.

Wary Customers. The case with wary customers is still clear cut, with some notable differences to the analysis of the benchmark case in Proposition 4. As previously, mandatory disclosure does not affect the equilibrium outcome, even though we know from Proposition 3 that $t_W^* > 0$. This is immediate as we observed above that even when commissions are not observed, still equilibrium contracts maximize the joint surplus ω that is realized by the product provider, the adviser, and the customer. This outcome cannot be improved upon, which is why the same outcome prevails even when commissions become observable through mandatory disclosure.

However, when the quality of the adviser's information is endogenous, a ban on commis-sions changes the equilibrium outcome. In fact, we know from Proposition 3 that when the advisor has to choose effort e, paying a positive commission is necessary to ensure that total surplus ω is maximized, even though this leads to biased advice, $q_W^* < 1/2$. A ban on com-missions leads to $q^* = 1/2$ and consequently leads to a strict reduction in total surplus and well as in the probability that the more suitable product is bought, as we show below in the proof of Proposition V. Consumer surplus, instead, is not affected, because wary customers are kept indifferent, according to the binding participation constraint (19). Intuitively, the regulation of the *form* of the contract between the adviser and product providers does not affect the market power of firms and thus their ability to extract consumer surplus (cf. also the discussion of competition in Section VI).

Naive Customers. The impact of policy intervention in the case of naive customers is more intricate when the quality of the adviser's information is endogenous. As we observed in the benchmark case, a policy that bans commissions protects naive customers from the exploitation of their misperceptions. While naive customers do not anticipate the bias in the adviser's recommendation, however, they also fail to anticipate how commissions affect the adviser's effort. From the analysis with wary customers we know that a marginal increase in commission can, however, lead to strictly higher effort, although this is the case only when the resulting bias is still sufficiently small as $q^* > q_0$. Instead, when the bias is large, the adviser's effort is strictly smaller than when there is a ban on commissions. This is intuitively the case when the adviser's suitability concern ρ is low. Then, naive customers underestimate the quality of advice *both* as they do not anticipate a biased recommendation *and* as they overestimate the adviser's effort.

When ρ is low, then also disclosure protects naive customers from exploitation and thus unambiguously leads to higher consumer surplus. Our analysis of the case with wary customers also implies that total surplus is also increased. Furthermore, when ρ is low it is not in the interest of firms to disclose commissions and, thereby, make naive customers wary of a conflict of interest.

PROPOSITION 5: *When the quality of the adviser's information is endogenous, policy intervention has the following impact:*

i) With wary customers, mandatory disclosure of commissions has no impact on the equilibrium outcome, while a ban on commissions strictly reduces the likelihood of a suitable purchase (lower L) without affecting consumer surplus.

ii) With naive customers, at least when the adviser's concern for suitability is not too large (low ρ), both mandatory disclosure and a ban on commissions strictly increase consumer surplus and the likelihood of a suitable purchase (higher L).

Proof of Proposition 5. Regarding assertion i) with wary customers, observe first that from Proposition 3 total surplus ω is strictly lower with the constraint $t = 0$. Also, from Lemma 4 we know that the choice of e^* is strictly lower under regulation. Given that thus both ω and $\kappa(e^*)$ are strictly lower when commissions are banned, also L must be lower.

For naive customers, note from condition (18) that without regulation we have $q_N^* = 0$ and thus $e_N^* = 0$ for all sufficiently low ρ, implying that L is then strictly lower than with policy intervention, i.e., with either mandatory disclosure (when $q_0 < q^* < 1/2$) or a

ban on commissions (when $q^* = 1/2$). Note next that for either policy intervention expected consumer surplus is equal to v_0, given that in either case customers' expectations are correct. To conclude that true expected consumer surplus is strictly lower than v_0 without policy intervention we use the binding participation constraint (11) together with the fact that for low ρ the true value of L is strictly lower than what naive customers anticipate, as observed above.[20] **Q.E.D.**

From the perspective of maximizing total surplus, from Proposition 5 mandatory disclosure is strictly preferable to a ban on commissions, provided that it indeed acts as an "eye opener" to naive customers. Recall that in the benchmark case with exogenous information quality both mandatory disclosure and a ban on commissions lead, instead, to the same outcome. In terms of consumer surplus, with wary customers policy intervention has no effect. Instead, under either policy intervention naive customers realize exactly v_0, which is more than without policy intervention at least for low ρ. Taken together, in our model mandatory disclosure has the same implication in terms of consumer protection but leads to higher total surplus than a ban on commissions, and is thus the preferred policy option.

VI. Fostering Competition

Given our focus on the structure of payments between customers, product providers, and financial advisers, our analysis abstracts from the institutional details of particular markets for retail financial services, such as investments or mortgages. Even in a particular class of financial products and services, there are large differences in the organization of the industry across different countries. In what follows, we analyze the effect of competition in a way that does not require spelling out the detail of the market structure that prevails in a particular industry. We analyze the impact of competition by performing comparative statics with respect to the customer's reservation value. We denote this reservation value by V_0, which so far was set at $V_0 = v_0$, according to (8).

---

[20]Incidentally, when the true effort level is e^* and the true recommendation cutoff is q^*, the true expected surplus is given by

$$S = v_0 + \Delta_v \left[\int_{q^*}^1 (2q - 1) dG(q \mid e^*) - \frac{1 - G(q^* \mid e^*)}{1 - G(\widehat{q}_N \mid \widehat{e}_N)} \int_{\widehat{q}_N}^1 (2q - 1) dG(q \mid \widehat{e}_N) \right].$$

Benchmark: Exogenous Quality of Information. For the benchmark case in which the quality of the adviser's information is exogenous, recall that we have $t_W^* = 0$ when customers are wary. Then, an increase in V_0 affects consumer surplus, namely through a reduction in the fixed fee for advice $f_W^* > 0$, but it has no effect on the suitability of advice and thus overall surplus. For naive customers, instead, we have $f_N^* = 0$ and advice is biased given the strictly positive commission $t_N^* > 0$. As the reservation value V_0 increases, the maximally feasible product price is reduced, and with it the commission (given that $t = p$) and also the bias in the adviser's recommendation are reduced. There are then two effects. First, the overall social welfare increases. Second, as the misperception of the naive customers is reduced, given that the wedge $\widehat{q}_N - q_N^*$ decreases, there is a second channel through which the true expected surplus of naive customers increases. In other words, when increased competition leads to a higher reservation value for customers and, thereby, to lower margins for product provider A, naive consumers benefit also indirectly, because their misperceptions are exploited to a lesser extent.

Endogenous Information Acquisition. Next, turn to our main scenario with endogenous information quality. With naive customers, the respective participation constraint (11) is easily modified to account for the possibility that the reservation value satisfies $V_0 \geq v_0$. Also, like in the benchmark case, a change in V_0 does not affect the form of the equilibrium contract with customers, because we always have $f_N^* = 0$, even though p_N^* must decrease and thus also $t_N^* = p_N^*$, resulting in a higher recommendation cutoff q_N^*. When the suitability concern ρ is sufficiently low, the end result is an increase in the likelihood of a suitable choice L. Also, naive customers' misperception is exploited less, now both in terms of the adviser's recommendation threshold and in terms of the effort exerted. Again, naive customers then benefit twofold, both directly from a reduction in the product price and indirectly from a reduction in exploitation.

With wary customers, the impact of a change in their reservation value V_0 is the same as in the benchmark case, provided that from $f_W^* > 0$ it is still feasible to marginally adjust downwards customers' direct fee for advice. Note from our previous observations that when ρ is low, the commission that the adviser receives with wary customers optimally remains low, implying also that customers pay relatively more directly for advice. Thus, when ρ is sufficiently low, the constraint $f \geq 0$ does not bind before and after the considered marginal

increase in V_0.[21] The following result then holds:

PROPOSITION 6: *When the adviser does not care too much about suitability (sufficiently low ρ), an increase in the customer's reservation value, V_0, does not affect total surplus when customers are wary, but it strictly increases total surplus (as well as consumer surplus) when customers are naive.*

VII.   Heterogeneous Customer Base

We now extend the analysis to consider a more general market composed of a fraction μ of wary customers and a fraction $1 - \mu$ of naive customers. When meeting a customer the adviser does not observe directly whether the customer is naive or wary.[22] We begin by considering the case in which the seller makes a single offer, and then allow the seller to indirectly price discrimination by offering a menu of offers, across which different customer types self select.

Single Offer.   We suppose first that the seller has to design a single offer, p. For some retail financial services this may be a reasonable assumption. For instance, in a given "share class" that is targeted to retail investors, mutual funds typically entail a fixed load and management fee. In our model, when there is a single price p, also the adviser then offers a single fee f. We assume throughout the following discussion that the commission is not directly observed by either wary or naive customers.

As a starting point, consider again the case without commissions ($t = 0$), where we also have $p = 0$, given that we set the cost to zero. Wary customers then have the same expectations as naive customers and have thus also the same willingness to pay up-front for advice. Consider now an increase in p. Naive customers then require that the fee is lowered by $df = dp[1 - G(\widehat{q}_N \mid \widehat{e}_N)]$. Instead, wary customers rationally anticipate that the likelihood of being recommended product A is actually higher, as the seller optimally increases the commission t. But wary customers also anticipate that the adviser's effort and thus the quality of information changes. In fact, when the change in commission and thus also in the true cutoff q^* is still not too large, we know that the true level of effort e^* strictly exceeds the level that is still anticipated by naive customers $\widehat{e}_N$ (Lemma 4). However, in analogy

---

[21]Instead, an increase in V_0 leads to lower total surplus when the constraint $f \geq 0$ binds.

[22]Instead, the analysis in Sections III and IV applies to the case in which the adviser directly observes whether the customer is naive or wary.

to our previous results, we can show that when the adviser's own concern for suitability is not too large (low ρ), in equilibrium with a single offer the participation constraint of wary customers must be binding.[23]

PROPOSITION 7: *Suppose that only a single acceptable offer is made in a market with both naive and wary customers. When the adviser's concern for suitability is not too large (low ρ) and when all customers participate, the outcome is the same as that prevailing with only wary customers.*

Proof of Proposition 7. For naive customers the marginal rate of substitution between p and f is obtained from the binding participation constraint as follows:

$$\frac{df}{dp_A} = -\left[1 - G(\widehat{q}_N \mid \widehat{e}_N)\right]. \tag{27}$$

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

$$v_0 + \Delta_v \int_{q^*}^{1} (2q - 1) dG(q \mid e^*) - p\left[1 - G(q^* \mid e^*)\right] - f = 0,$$

where q^* and e^* are the respective optimal choices (cf. Lemma 3). Using $t = p$ and thus that $dq^*/dp = -1/(2\rho)$, the marginal rate of substitution for wary customers is

$$\frac{df}{dp} = -\left[1 - G(q^* \mid e^*)\right] - \frac{1}{2\rho} \left[\begin{array}{c} -g(q^* \mid e^*)\left[(2q^* - 1)\Delta_v - p\right] \\ +\frac{de^*}{dq^*}\frac{d}{de^*}\int_{q^*}^{1}\left[(2q - 1)\Delta_v - p\right] dG(q \mid e^*) \end{array} \right], \tag{28}$$

when interior. From (15) we have $\frac{de^*}{dq^*} \to 0$ when $\rho \to 0$. Given that $2q^* - 1 < 0$, we conclude that (28) is strictly lower than (27) if ρ is sufficiently small. Recall now that the anticipated (and true) consumer surplus of naive and wary customers is the same at the corner $p = t = 0$. When both customers find the offer acceptable, we conclude that only the participation constraint of wary customers binds for low ρ. Thus in this case the profit maximization problem is then identical to the one resulting with only wary customers. **Q.E.D.**

Note that Proposition 7 presumes that a uniform offer is acceptable to all customers, which is clearly optimal when the fraction of naive customers is not too large (i.e., when μ is not too small). In addition, note that the outcome of Proposition 7 once again holds irrespective of the level of ρ when the adviser cannot affect (or can affect only marginally) the

---

[23]Clearly, in the benchmark case with exogenous information quality, $e = \widetilde{e}$, the participation constraint of wary customers binds more generally for any level of ρ.

quality of information. Then, under uniform offers naive customers are adequately protected by the presence of wary customers.

Discriminatory Offers. Even when direct (first-degree) discrimination is not possible when the rationality of the customer is not publicly observed, what would happen if the product provider and the adviser could make discriminatory offers through a menu? Such a menu could prescribe different up-front payments for advice together with different prices for product A. Clearly, the menu of offers must be incentive compatible.

When customers choose different contracts for low ρ, it follows from the arguments leading to Proposition 7 that the respective contracts (p_W, f_W) for wary customers and (p_N, f_N) for naive customers must satisfy $f_W > f_N$ and $p_W > p_N$. Furthermore, in an incentive compatible menu that is acceptable to both customer types the participation constraint for wary customers must bind, at least for low ρ. By optimality, when the participation constraint of naive customers is slack, their incentive compatibility constraint must bind. From standard arguments, given that for low ρ there is a single-crossing property with respect to customers' payments p and f, to extract higher surplus from naive customers firms will then optimally distort the offer made to wary customers by raising f and lowering p. Instead, for naive customers we still have $f_N = 0$. The formal analysis in the Appendix confirms that it is optimal for the firm never to charge naive customers up-front for advice, consistent with practice that prevails in markets for retail financial services. This final result confirms the robustness of one of our main theoretical insights.

VIII. Conclusion

Rather that being compensated directly by customers, advisers and salespeople in the financial industry are often paid indirectly by product providers when customers decide to purchase the product offered. This practice has led to widespread claims of unsuitable advice. Policy proposals include prohibiting or, at least, seriously capping 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.

When firms face customers who are naive about the true conflict of interest that is induced by commissions, we have argued that firms can maximally exploit this naiveté by

only charging customers indirectly for advice. In this case, banning commissions protects customers and, at least when advisers' intrinsic preferences are not strongly aligned with those of customers, also leads to more suitable purchase decisions. Alternatively, we suggest that mandating the disclosure of the presence of commissions could serve as an eye-opener, making naive customers wary of the underlying conflict of interest. Firms would then have themselves no incentives to disclose commissions. Such disclosure would also risk generating unintended consequences in the presence of wary customers. A ban on commissions, in contrast, leads to unbiased advice but can decrease the overall quality of advice in case advisers have to exert effort to acquire information.

The present analysis is a first step of a research program that aims at reaching positive and normative predictions on the compensation structure in the retail financial industry, with special emphasis on the role of advice. Our model allows for compensation from product providers to advising intermediaries in combination with payments made by customers to both product providers (through a price contingent on the transaction) and intermediaries (through a up-front fixed fee). Future work could add more structure by analyzing the separate channels through which advisers could be disciplined, such as liability or reputational concerns. Also, allowing also customers to acquire information, possibly to a different degree depending on their financial capability, may add more realism when applying the model to specific markets.

IX.   Appendix: Analysis with Discriminatory Offers

We denote the respective triples under menu offers by (f_W, p_W, t_W) and (f_N, p_N, t_N), respectively, where $t_. = p_.$ from firms' optimality. Recall that the respective ex ante participation constraints are given by (11) and (19). Further, the incentive compatibility constraint of naive customers is

$$f_W - f_N \geq [1 - G(\widehat{q}_N \mid \widehat{e}_N)] [p_N - p_W]. \tag{29}$$

For ease of exposition, now denote the advisers' optimal choices under the wary customers' contract by q_W^* and e_W^*. Then, wary customers' incentive compatibility constraint becomes

$$
\begin{aligned}
f_W - f_N \ \leq \ & \left[\int_0^{q_W^*} [\Delta_v(1-q)] \, dG(q \mid e_W^*) + \int_{q_W^*}^1 [q\Delta_v - p_W] \, dG(q \mid e_W^*) \right] \\
& - \left[\int_0^{q_N^*} [\Delta_v(1-q)] \, dG(q \mid e_N^*) + \int_{q_N^*}^1 [q\Delta_v - p_N] \, dG(q \mid e_N^*) \right].
\end{aligned}
\tag{30}
$$

Note first that the ex ante participation constraint for naive customers must be slack for low ρ. This follows immediately from the observation that naive and wary customers' expected surplus is the same when $p = 0$, together with the argument in Proposition 7. When contracts are different, from this argument and incentive compatibility we also have immediately that $p_W < p_N$. By optimality, the incentive compatibility constraint for naive customers must bind. As naive customers thus expect to realize the same surplus under either contract, when contracts are different, we have from $p_W < p_N$ together with the argument in Proposition 7 that the incentive compatibility constraint for wary customers must be slack.

Summing up, in case firms want to ensure that all customers participate, denoting the fraction of wary customers by μ, the objective is to maximize

$$\mu\left[f_W + [1 - G(q_W^* \mid e_W^*)]\,p_W + \rho L - \kappa(e_W^*)\right]$$
$$+(1 - \mu)\left[f_N + [1 - G(q_N^* \mid e_N^*)]\,p_N + \rho L - \kappa(e_N^*)\right]$$

subject to (19) and (29). Substituting for f_W and f_N from the two binding constraints, the objective function becomes

$$\mu\omega_W + (1 - \mu)\omega_N \tag{31}$$
$$+(1 - \mu)p_N\left[G(q_N^* \mid e_N^*) - G(\widehat{q}_N \mid \widehat{e}_N)\right]$$
$$-(1 - \mu)\left[\begin{array}{c}\left[\int_0^{\widehat{q}_N} \Delta_v(1-q)dG(q \mid \widehat{e}_N) + \int_{\widehat{q}_N}^1 [q\Delta_v - p_W]\,dG(q \mid \widehat{e}_N)\right] \\ -\left[\int_0^{q_W^*} \Delta_v(1-q)dG(q \mid e_W^*) + \int_{q_W^*}^1 [q\Delta_v - p_W]\,dG(q \mid e_W^*)\right]\end{array}\right].$$

From this we obtain that p_N is chosen as high as possible, i.e., until (29) binds, so that $f_N = 0$. Further, by our previous arguments (cf. Proposition 7) the derivative of the last line in (31) with respect to p_W is strictly negative, implying thus that an optimal choice of p_W and thus of $t_W = p_W$ is now strictly lower than that obtained when only wary customers are present. We conclude that f_W is strictly higher with a menu of offers than with a single offer, and thus with the offer that is made when only wary customers are in the market.

X.   References

Admati, Anat, and Paul Pfleiderer, 1986, A monopolistic market for information, *Journal of Economic Theory* 39, 400–438.

Bergstresser, Daniel, and John Beshears, 2010, Who selected adjustable-rate mortgages? Evidence from the 1989-2007 Surveys of Consumer Finances, Working Paper 10-083, Harvard Business School.

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* 22, 4129–4156.

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, 1–25.

Campbell, John, Howell Jackson, Brigitte Madrian, and Peter Tufano, 2010, Consumer Financial Protection, *Journal of Economic Perspectives*, forthcoming.

Carlin, Bruce I., and Simon Gervais, 2008, Legal protection in retail financial markets, mimeo, UCLA and Duke University.

Carlin, Bruce I., 2009, Strategic price complexity in retail financial markets, *Journal of Financial Economics* 91, 278–287.

Chen, Joseph, Harrison Hong, and Jeffrey D. Kubik, 2006, Outsourcing mutual fund management: firm boundaries, incentives, and performance, Working paper, University of Southern California, Princeton, and Syracuse.

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.

Cummins, David, and Neal Doherty, 2006, The economics of insurance intermediaries, *Journal of Risk and Insurance* 73, 359–339.

DellaVigna, Stefano, and Ulrike Malmendier, 2004, Contract design and self-control: theory and evidence, *Quarterly Journal of Economics* 119, 353–402.

Durbin, Erik, and Ganesh Iyer, 2009, Corruptible advice, *American Economic Journal: Microeconomics* 1, 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.

Financial Services Authority, 2010, *Quality of Advice on Pension Switching: An Update*, Report April 2010.

Gabaix, Xavier, and David Laibson, 2006, Shrouded attributes, consumer myopia, and information suppression in competitive markets, *Quarterly Journal of Economics* 121, 505–540.

Ganuza, Juan-Josè, and Josè S. Penalva, 2010, Signal orderings based on dispersion and the supply of private information in auctions, *Econometrica* 78, 1007–1030.

Gneezy, Uri, 2005, Deception: the role of consequences, *American Economic Review* 95, 384–394.

Gravelle, Hugh, 1993, Product price and advice quality: implications of the commission system in life assurance, *Geneva Papers on Risk and Insurance Theory* 18, 31–53.

Gravelle, Hugh, 1994, Remunerating Information Providers: Commissions versus Fees in Life Insurance." *Journal of Risk and Insurance* 61, 425–457.

Hackethal, Andreas, Roman Inderst, and Steffen Meyer, 2010, Trading on advice, mimeo, University of Frankfurt.

Hong, Harrison, Jose Scheinkman, and Wei Xiong, 2008, Advisors and asset prices: a model of the origins of bubbles, *Journal of Financial Economics* 89, 268–287.

Inderst, Roman, and Marco Ottaviani, 2009, Misselling through agents, *American Economic Review* 99, 883–908.

Inderst, Roman, and Marco Ottaviani, 2010a, Sales talk, cancellation terms, and the role of consumer protection, mimeo, Northwestern University and University of Frankfurt.

Inderst, Roman, and Marco Ottaviani, 2010b, Competition through commissions and kickbacks, mimeo, Northwestern University and University of Frankfurt.

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.

Keith, Ernst, Debbie Bocian, and Wei Li, 2008, Steered wrong: brokers, borrowers, and subprime loans, Working paper, Center for Responsible Lending.

Lacko, James M., and Janis N. Pappalardo, 2007, Improving consumer mortgage disclosures: an empirical assessment of current and prototype disclosure forms, Bureau of Economics Staff Report, Federal Trade Commission.

Lim, Terence, 2001, Rationality and analysts' forecast bias, *Journal of Finance* 56, 369–385.

Malmendier, Ulrike, and Devin Shanthikumar, 2007, Are small investors naive about incentives? *Journal of Financial Economics* 85, 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.

Stoughton, Neal M., Youchang Wu, and Josef Zechner, 2008, Intermediated investment management, *Journal of Finance* forthcoming.

US Department of Treasury, 2009, *Financial Regulatory Reform. A New Foundation: Rebuilding Financial Supervision and Regulation*, White Paper.