Misconduct in Credence Good Markets*

Jennifer Brown
Northwestern University
NBER

Dylan Minor
Northwestern University

December 2012

Abstract

We study how monitoring, expert skill and consumer awareness affect the level of misconduct in markets with asymmetric information and price-taking experts. Theoretical predictions show that experts subject to more intense monitoring may be *less* ethical in equilibrium. Similarly, more experienced experts are predicted to exhibit greater levels of misconduct. We test these predictions in the insurance sales industry and find that monitored experts are 21 to 98% more likely to take advantage of customers, relative to unmonitored experts. We also find empirical evidence that more experienced experts are significantly more likely to mislead their customers.

Keywords: Misconduct, expert services, monitoring, asymmetric information, credence goods, insurance, ethics.

*We thank Tim Fedderson, Craig Garthwaite, Tom Hubbard, Mike Mazzeo, Nicola Persico and seminar participants at the Kellogg School of Management, Florida State University, 2012 NBER Law & Economics Summer Institute, Center for Research in Economics and Strategy Conference, and LMU Munich Workshop on Natural Experiments and Controlled Field Studies for helpful comments. We are grateful to Eric Zhang and Tongtong Shi for excellent research assistance. We also thank individuals at the Texas Department of Insurance for their help in acquiring the data. Brown: jen-brown@kellogg.northwestern.edu Minor: d-minor@kellogg.northwestern.edu

> *"... will people push the envelope and pitch lucrative and complicated products to clients even if they are not the simplest investments or the ones most directly aligned with the client's goals? Absolutely. Every day, in fact."*
>
> - Greg Smith, former executive at Goldman Sachs
> *New York Times* Op-Ed (March 14, 2012)

Expert services firms are often found in markets with substantial asymmetric information problems—providers of technical advice are common in the automotive, medical, engineering, and financial services industries. Experts benefit from customers trusting and buying their advice; however, experts may also face incentives that lead them to provide less than perfect recommendations. For example, a mechanic can provide a more extensive fix than warranted and a dentist can replace a filling that has not failed. In addition to over-treating a problem, experts can also suggest the wrong solution. For example, investment or insurance advisors can recommend products that offer customers less benefit, but provide themselves with greater revenue than the customers' ideal products.

With credence goods, it is difficult for a customer to determine whether the product or service is the best match for his or her needs. In extreme cases, the customer may never discover if the product was the most appropriate one—for example, the final benefit of life insurance is realized upon death.[1] When it is difficult for a customer to discern the correct product or service, an expert who both advises and receives revenue based on his advice faces conflicting incentives. High quality advice may improves the customer's payoff; yet, when taken by the customer, inappropriate advice may lead to higher expert revenue.

Many of the existing models of expert services allow advisors to adjust both quality and prices. In contrast, we explore credence good markets with price-taking experts. Examples of price-taking experts include individual physicians and dentists who may have limited scope to adjust prices for a particular patient; taxi cab drivers who face regulated rates; and, our empirical setting, insurance sales agents who face fixed commissions and prices.

In this paper, we ask: How do monitoring, experience and skill affect experts' propensity to engage in professional misconduct? We focus on monitoring because it is a salient point of differentiation in many expert services markets—experts may work as independent advisors and be subject to little oversight; or experts may work as representatives of large, hierarchical organizations. While monitoring may appear beneficial for customers who rely on experts' advice, just the opposite happens in our setting: Experts at firms with hierarchies and structure for monitoring are the ones most likely to take advantage of customers.

[1] While purchasers of "experience goods" gain utility from the actual consumption of the product or service, purchasers of "credence goods" gain utility based on their *beliefs* about the product or service.

The intuition is as follows: Since experts are price takers, their dimension of competition is their level of misconduct. For a given level of malfeasance, customers working with experts in firms with monitoring fare better in expectation relative to customers using unmonitored experts. Monitored experts cannot set their own prices to extract surplus from the larger expected consumer benefits of greater monitoring; instead, they extract surplus through greater misconduct. Similarly, more experienced experts are more skillful at providing the most appropriate solution for customers' needs. Hence, they can extract more rents through increased misconduct compared to less experienced experts.

We test the predictions of our theoretical model using data from insurance markets. Here, we have a clear credence good setting, particularly in life insurance and annuity sales. Additionally, experts themselves acknowledge the ethical quandary of their field. In Cooper and Frank (2005), a survey of insurance agents finds that agents consistently identify three primary ethical issues: failure to identify the customer's needs and recommend products that meet those needs; false or misleading representation of products or services; and conflicts between customer benefits and opportunities for personal financial gain.

For our empirical tests, we construct a rich dataset describing individual insurance agents operating in Texas. We match licensing data with company affiliations and detailed sales practice complaint records from the state regulator. From the company affiliation data, we identify two types of experts: monitored agents from large, branded companies, and unmonitored agents working as independents. We find that the odds of monitored experts from large, branded companies taking advantage of their customers are 21 to 98% greater than the odds for unmonitored independent experts. In a supplemental analysis, we use national sale practice complaints data to confirm our results. Finally, we find that more experienced agents are significantly more likely to mislead their customers.

Reputation has been offered as a solution to asymmetric information problems in markets. Reputation is built through repeated interactions across or within customers over time (for examples, see Kreps (1990) and Tadelis (1999)). However, the nature of credence good markets means that misconduct is seldom observed; the signals required for reputation building are not sufficiently informative (Mailath and Samuelson 2001). As a result, it is often not possible to build a reputation of good behavior. Yet, we still observe strong branding of firms in many credence good settings—for example, insurance companies, wirehouses, and hospital networks are often heavily advertised. Branding and reputation solve informational asymmetry in many markets; however, in our empirical setting, the correlation between strong branding and monitoring leads to a prediction that experts from large, branded firms are actually more likely to engage in misconduct.

Darby and Karni (1973) provide the foundation for the literature on credence goods.

Pitchik and Schotter (1987) isolate the problem of the expert honestly suggesting a mode of treatment and provide comparative statics results comparing price and quality controls and the level of honesty. Pessendorfer and Wolinsky (2003) study the first stage of a similar problem: the need to provide incentives for the expert to expend enough effort to identify and provide a correct solution. Sulzle and Wambach (2005) explore how changing physician and patient incentives through higher coinsurance levels may (or may not) induce patients to increase physician search and encourage physicians to reduce fraud. Alger and Salanie (2006) also consider the role of the client and find that a patient's ability to reject an expert's recommendation creates a market failure. Emons (1997) shows that market equilibria with honest expert behavior exist when customers can infer sellers' incentives for fraud from market data.

Customer heterogeneity may also drive the credence good problem. Fong (2005) shows that cheating arises when firms target high-valuation and high-cost customers. Feddersen and Gilligan (2001) find that third parties, namely activists, can ameliorate the credence good problem. Taylor (1995) examines multi-period contracts and warranties as another solution. Inderst and Ottaviani (2009, 2011, 2012) study firms trying to induce agents to provide advice to imperfectly informed customers. They find that mis-selling depends on firm asymmetries, customer awareness, and agents' utility from giving suitable recommendations. Broadly, in their models, agents provide honest advice when firms are symmetric or there are sufficiently many aware customers in the market. Dulleck and Kerschbamer (2006) present a model that unifies the extant literature and rationalizes many of the previous theoretical findings.

Hubbard (1998) explores empirically the incentives faced by experts in automotive repair services. He finds that private firms are more likely than state inspectors to help vehicles pass emissions tests. Moreover, he finds that independent experts are more likely to provide favorable inspection reports, relative to branded "chain" shops with non-owner managers. Hubbard (2002) suggests that the possibility of many future transactions provides incentives for experts to offer more favorable advice, particularly where experts are residual claimants. Free-riding may also dampen individual experts' incentives, as firms with more inspectors tend to help vehicles pass less frequently. Levitt and Syverson (2008) find that real estate agents invest more effort and secure a higher price for the sale of their own property, relative to their customers' homes. Similar to the mechanism proposed by Hubbard (2002), Levitt and Syverson argue that the absence of frequent and repeated interactions limits customers' abilities to verify their agents' service quality. They also find that the difference between agent-owned and non-agent-owned sale prices is increasing in the degree of asymmetric information about property values. In a very different context, Gruber and Owings (1996)

find that physicians perform more cesarean-section deliveries in response to negative income shocks. Mullainathan, Noeth, and Schoar (2012) conduct a field audit study in a U.S. market and find that financial advisors often recommend self-serving products. Anagol, Cole, and Sarkar (2012) conduct an audit study of insurance sales agents in India and find similar results.

The paper proceeds as follows. In the next section, we present our theory model and results. Section 2 provides institutional background on the insurance industry. Our data are described in Section 3 and our empirical results are presented in Section 4. Our final section discusses the implications of our findings.

1 A Model of Credence Goods Sales

In this paper, we present a model inspired by the unifying model in Dulleck and Kerschbamer (2006), hereafter DK. However, our model differs in important ways.

In DK, different outcomes are driven by experts offering services at different prices (e.g. mechanics choose quality and prices for auto repairs). In fact, virtually all of the aforementioned theory papers studied price-setting firms or advisors. In contrast, we consider a market in which experts are price takers. In our empirical setting, insurance agents are constrained to offer products with fixed premiums and commissions.[2] Industries with regulated prices also exhibit this feature.

We assume that a customer is made worse off by an inappropriate match, regardless of his or her level of need. Existing models have assumed that experts have only limited opportunities for misconduct—for example, experts can provide a major repair for a minor problem, but can only provide a major repair for a major problem. These models have assumed away an expert's flexibility to take advantage of customers that have major needs (see DK and cites therein). However, in practice, experts may provide inappropriate treatment to a variety of customer types. For example, a doctor can order excessive tests for any level of medical need. Similarly, an insurance agent can always oversell life insurance coverage to generate greater commission.

Rather than assuming that all experts work in a common institutional setting, we allow for two types of experts: monitored and unmonitored. This extension allows us to explore how different organizational structures support different levels of misconduct.

We assume there is always some chance of mistreatment being discovered, resulting in a penalty against the expert. In general, previous work has assumed that customers can only

[2] Rebating—where an agent kicks back some of the commission to a client to adjust the effective price of a product—is illegal in most jurisdictions.

identify and punish experts for an under-provision of services—only a minor treatment given for a major problem can be discovered *ex-post* by the customer.

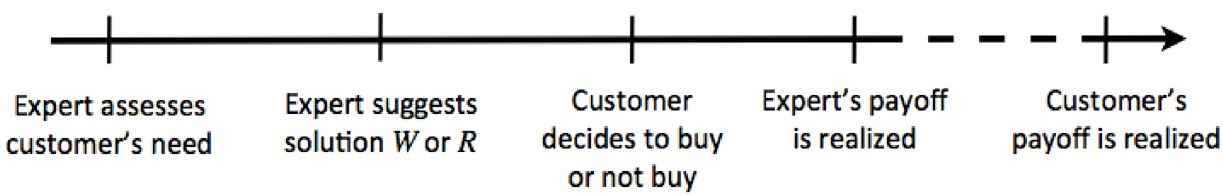
Finally, we extend the credence good model to allow for heterogeneous expert skill and the introduction of connoisseur consumers.[3]

1.1 Basic Model

Consider an interaction between an expert and a customer that can result in two outcomes: the expert can recommend either an appropriate or inappropriate product. For convenience, we will use the index "R" and "W" as mnemonics for the "right" (appropriate) and "wrong" (inappropriate) products, respectively. We assume that the expert knows which product is appropriate for the customer, but the customer does not. After the expert makes his product recommendation, the customer must chose to buy or not to buy.

Suppose that π^R and π^W are the payoffs to an expert for selling the appropriate and inappropriate products, respectively. It follows that π^t is a reduced form representation of the net payoff (i.e., gross revenue minus business expenses) of selling product $t \in \{R, W\}$, before any possible penalty for mis-selling to a customer (i.e., recommending W).

As depicted below, the timeline for the expert-customer interaction is sequential.



Since the customer cannot condition his purchase decision on any information about the quality of the expert's recommendation, the game can be solved simultaneously.

Let s be the probability that the expert recommends product W and $(1-s)$ be the probability that he recommends R. Now, assume that there is some expected cost $k(s) > 0$ for recommending W, where $k(0) = k'(0) = 0$ and $\frac{\partial}{\partial s} k(s) > 0$. Thus, $k(s)$ reflects the expected cost of mistreating customers which, in turn, reflects both the probability of detection and the magnitude of the punishment. Psychological costs associated with "doing the wrong thing" may also enter into $k(s)$.

In our model, if an expert offers the inappropriate product and the customer does not buy it, then the expert still faces a payoff of $-k(s)$. This captures the notion that experts who attempt to deceive their customers will face some probability of detection regardless of whether the customers actually buy the (inappropriate) products. For example, in the insurance industry, a customer typically receives a 10- to 30-day "free look" after paying for

[3] For a model of financial advisors in an experience good setting, see Bolton, Freixas, and Shapiro (2007).

an annuity or life insurance product. During this period, a customer could discover that he was sold W, report the agent to the regulator, and cancel the policy. This feature of our model relaxes the typical credence good model assumption that customers remain forever ignorant of their expert's misconduct.[4]

Let b be the probability that the customer buys the expert's recommended product and $(1-b)$ be the probability that the customer rejects the expert's recommendation. Suppose that the customer earns a net payoff of V^R from buying R and V^W from buying W, where $V^W < 0 < V^R$. If the customer decides not to buy any product, then her payoff is 0, her normalized outside option. Note that we assume that a customer is worse off buying the wrong product than he would have been simply not buying at all. Absent this assumption, the customer would rather be mistreated with certainty than reject the expert's advice, even knowing such advice is bad.

The payoffs can be described in a 2 x 2 matrix, where the first coordinate is the expert's payoff and the second is the customer's payoff:

	Buy	Don't Buy
Right Product	(π^R, V^R)	$(0, 0)$
Wrong Product	$(\pi^W - k(s), V^W)$	$(-k(s), 0)$

In the following section, we identify the mixed strategy equilibrium in which the customer is indifferent between buying and not buying. Following Harsanyi (1973), these mixed strategies can be reframed as representing a heterogeneous population of customers, each with a pure strategy. Since $k(s)$ is endogenous, the expert will choose a pure strategy.

1.2 Monitoring

We enrich the model to consider experts from firms with different levels of monitoring. We assume that the experts face similar payoffs across different levels of monitoring and, for a given level of misconduct, customers' payoffs are also equal.[5] This formulation of the model captures a common feature of expert industries: some experts operate in larger, branded firms with monitoring, while other experts operate as small, independent advisors with little (if any) monitoring. For example, in financial services, several monitored experts typically work in a branch office that is overseen by a manager. Independent experts may work in one-agent offices without supervision.

[4]In those models, an agent who is unsuccessful in selling W receives a payoff of 0—the same payoff he or she would earn from unsuccessfully marketing R.

[5]It is straightforward to show that our predictions hold if the monitored, branded firms offer customers' higher payoffs.

Assume that a monitor observes experts' recommendations with probability $q \in [0,1]$, where $q = 0$ represents no oversight and $q = 1$ means that every expert recommendation is reviewed. If the monitor observes an expert recommending W, then he stops the transaction—the consumer is indemnified for her loss V^W (i.e. she receives her outside option 0), and the expert faces penalty $-k(s)$ and does not keep any positive payoff π^W.[6] If the monitor observes an expert recommending R, then he does not intervene. Therefore, the expert's payoff for suggesting R is $b\pi^R$, but his payoff for recommending W is $(1-q)\left(b\pi^W - k(s)\right) + q(-k(s))$ or $(1-q)b\pi^W - k(s)$.

Since monitoring changes customers' payoffs, the level of misconduct s will also change. In particular, though the customer's payoff from the appropriate product is still V^R, she now receives $(1-q)V^W$ when she purchases the inappropriate product, where $V^W < (1-q)V^W < 0$. Monitoring saves her from some of the bad recommendations.

At low levels of monitoring, it cannot be an equilibrium for the customer to always buy; if this were so, the expert would always suggest W. But then the customer would be better off never purchasing. Never purchasing cannot be an equilibrium because then the expert will only suggest R and the customer would then choose to always buy. Hence, the customer's strategy must be a mixed one, at least for low levels of monitoring. Solving for the rate of s such that the customer is indifferent between buying and not buying yields

$$s^* = \frac{V^R}{V^R - (1-q)V^W} \tag{1}$$

Note that s^* is increasing in the level of monitoring ($\frac{\partial s^*}{\partial q} > 0$ since $V^W < 0$). A customer facing a monitored expert knows that there is some chance that the expert will offer the wrong product; however, there is also some probability that the monitor will detect this misconduct and refund the customer's payment. Overall, holding expert misconduct fixed, the customer has a higher expected payoff from a transaction with a more monitored expert.

In equilibrium,the expert must not want to deviate from her required course of action s^*. Thus, she solves the following problem, taking b as given:

$$\max_s \; s\left[(1-q)\left(b\pi^W - k(s)\right) + q\left(-k(s)\right)\right] + (1-s)\, b\pi^R$$

The first order condition yields

$$(1-q)b\pi^W - k(s) - sk'(s) - b\pi^R \equiv 0$$

[6] Alternatively, we could assume that the penalty $k(s)$ is greater when misconduct is discovered by the monitor. This does not change the qualitative results, so we omit this extension for ease of exposition.

Note that we already know the s^* that sustains an equilibrium. Hence, we can solve the first order condition in terms of b^*, such that the expert strictly prefers to provide s^* for a given level of monitoring. Note that $k(\cdot)$ is convex; the second order condition confirms that a maximum is obtained. Thus, we find

$$b^* = \frac{k(s^*) + s^* k'(s^*)}{(1-q)\pi^W - \pi^R}$$

Increased monitoring results in a greater buy rate b^*. Hence, experts with greater misconduct enjoy greater buy rates from customers; these experts extract more surplus from the value created by greater monitoring. We assume that $\frac{k(s^*)+s^*k'(s^*)}{\pi^W - \pi^R} < 1$. This assumption on the primitives ensures that customers do not always buy from the expert in the absence of monitoring—this would assume away the credence good problem.

After some level of monitoring $q \in (0, 1)$, the expected cost of recommending W is so great that the expert only recommends R. This occurs when

$$(1-q)b\pi^W - k(s) \leq b\pi^R$$

which happens necessarily when $q \in [\frac{\pi^W - \pi^R}{\pi^W}, 1]$ since $k(s) \geq 0$.

Define $\overline{q} = \frac{\pi^W - \pi^R}{\pi^W} < 1$. At some $q < \overline{q}$ with $k(s) > 0$, the buy rate will already equal 1; in particular, this occurs when

$$b^* = \frac{k(s^*) + s^* k'(s^*)}{(1-q)\pi^W - \pi^R} = 1$$

This means that the level of monitoring $\underline{q}$ where the customer always buys is

$$\underline{q} = \frac{\pi^W - \pi^R - k(s^*) - s^* k'(s^*)}{\pi^W} > 0 \tag{2}$$

It is straightforward to show that, although $k(\cdot)$ is indirectly a function of q, there is a unique value of q that solves (2).

Thus, when $q \in [\underline{q}, \overline{q}]$, the customer always buys. For this region of monitoring, the expert now faces the problem

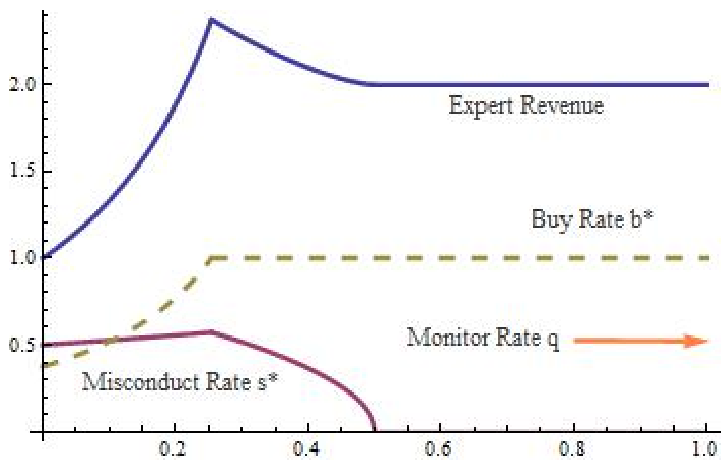
$$\max_s s\left((1-q)\pi^W - k(s)\right) + (1-s)\pi^R$$

and the expert chooses s according to the expression

$$(1-q)\pi^W - \pi^R = k(s) + sk'(s) \tag{3}$$

The left- and right- hand side of expression (3) represent the marginal benefit and marginal cost of misconduct, respectively. In contrast to all lower values of monitoring, the level of misconduct is now decreasing in the level of monitoring. For high levels of monitoring or detection (i.e., when $q > \underline{q}$), we find that increased monitoring reduces misconduct.

To illustrate these two regions of monitoring intensity, consider a simple example with $V^R = 2$, $V^W = -2$, $\pi^R = 2$, $\pi^W = 4$, and $k(s) = s^2$.



In the figure, misconduct increases in monitoring until customers always buy. With higher levels of monitoring, customers still always buy, but misconduct is weakly decreasing. Experts' revenues track their misconduct—revenues increase with low levels of monitoring and decline at higher levels.

We summarize our findings in the following proposition.

Proposition 1 *Define* $\underline{q} = \frac{\pi^W - \pi^R - k(s^*) - s^* k'(s^*)}{\pi^W}$. *With lower monitoring intensity, when* $q < \underline{q}$, *increasing the rate of monitoring* q *increases the level of misconduct* s^*, *expert revenue, and customer buy rate* b^*. *With higher monitoring intensity, when* $q \geq \underline{q}$, *increasing the rate of monitoring* q *weakly decreases the level of misconduct* s^* *and revenue, and customers always buy.*

One might wonder why all firms do not monitor their experts at high levels. First, intense monitoring may be too costly. Second, it may not be possible for the firm to monitor all activities, particularly when experts have the ability to hide some of their actions.

While the effect of monitoring depends on the particular empirical context, monitoring is far from perfect in credence good markets and, in practice, most likely in the lower region $q < \underline{q}$. More formally, consider an effective level of monitoring $q \equiv \widetilde{q}\,\Pr(\text{detected}|\widetilde{q})$, where $\widetilde{q}$ is the frequency an expert's advice is reviewed by a monitor and $\Pr(\text{detected}|\widetilde{q})$ is the

likelihood of detecting misconduct given a review. Even with $\widetilde{q} = 1$, since $\Pr(\text{detected}|\widetilde{q})$ is expected to be small for credence good markets, q is low.

Monitoring is similar in spirit to whistle-blowing. The critical difference is that, with whistle-blowing, we assume that the customer does not have an improved payoff in the event of misconduct and detection. That is, whistle-blowing involves the detection of bad behavior, but not the indemnification of abused customers. Since customers' payoffs are unaffected by whistle-blowing, experts do not increase their level of unethical sales behavior.

Note that we could also simply have written the downside customer payoff of V^W (and the upside expert payoff of π^W) as some general increasing (decreasing) function of the degree of monitoring q. In this case, the comparative statics of s^* and b^* for $q < \underline{q}$ still follow immediately. However, we would now need to put structure on $V^W(q)$ and $\pi^W(q)$ to ensure the existence of some level of monitoring that restores the market.

1.3 Connoisseur Consumers

In this section, we consider the impact of connoisseur consumers on the market equilibrium. Connoisseurs are defined as consumers who are perfectly informed about the appropriateness of the recommended product and, therefore, only and always buy from an expert who recommends R. We assume that experts cannot distinguish a connoisseur from a regular customer—otherwise, the expert simply always suggests R to such consumers and regular consumers are unaffected. Adding connoisseurs is equivalent to introducing some probability that a consumer knows the appropriate product for herself.

With a mass α of connoisseurs in the market, the expert's payoff for suggesting R increases while her payoff for suggesting W decreases. The expert's problem is now

$$\max_{s} \left[s\left((1-q)(1-\alpha)b\pi^W - k(s)\right) + (1-s)\left((1-\alpha)\, b\pi^R + \alpha\pi^R\right)\right]$$

The first order condition yields

$$b^* = \frac{1}{(1-\alpha)} \frac{k(s^*) + s^*k'(s^*) + \alpha\pi^R}{(1-q)\pi^W - \pi^R} > \frac{k(s^*) + sk'(s^*)}{(1-q)\pi^W - \pi^R}$$

Therefore, the overall market buy rate is

$$\begin{aligned}(1-\alpha)b^* + \alpha(1-s^*) &= \frac{k(s^*) + s^*k'(s^*) + \alpha\pi^R}{(1-q)\pi^W - \pi^R} + \frac{-\alpha\,(1-q)\,V^W}{V^R - (1-q)\,V^W} \\ &> \frac{k(s^*) + s^*k'(s^*)}{(1-q)\pi^W - \pi^R}\end{aligned}$$

This expression suggests that as monitoring increases, the market will reach a buy rate of 1 sooner than in a market without connoisseurs. Of course, holding q fixed, increasing the measure of connoisseurs also eventually leads to a buy rate of 1. As in a market without connoisseurs, once monitoring is sufficiently high, non-connoisseur customers always buy and the expert decreases her misconduct until $s = 0$.

We summarize these findings in our second proposition.

Proposition 2 *As the mass of connoisseurs α increases, the equilibrium buy rate of non-connoisseur consumers $b^* \to 1$. Expert misconduct is weakly decreasing in α.*

This extension links pure experience and credence good models: when $\alpha = 1$, customers can perfectly assess product quality after purchase and return a low quality product to the seller; when $\alpha = 0$, the customer never learns the true product quality.

1.4 Observable Differences in Expert Skill

We also consider a version of the model where, on occasion, experts *inadvertently* recommend the inappropriate product. Thus, we assume that an expert makes harmful mistakes.[7] Of course, the expert is also able to *choose* to recommend the inappropriate product, since that may increase his revenue at the customer's expense. In this extension, we consider the effect of experts' skill differences, conditional on a given level of monitoring.

Let h be the commonly known probability that an expert makes an error. Now an expert faces the problem

$$\max_s \left[s\left((1-q)b\pi^W - k(s)\right) + (1-h)\,(1-s)\,b\pi^R + h\,(1-s)\left((1-q)b\pi^W - k(s)\right)\right]$$

with the following first order condition:

$$(1-h)\left[(1-q)b\pi^W - k(s) - sk'(s) - b\pi^R\right] - hk'(s) \equiv 0$$

Thus, s^* will be smaller than when $h = 0$, since the marginal cost increased. In addition, s^* will differ by skill level because buy rates are a function of expert skill. A customer must be indifferent between buying from an expert with an error rate of h and earning her outside option of 0:

$$(s + h\,(1-s))\,(1-q)\,V^W + (1-s)\,(1-h)V^R = 0$$

[7] We assume that experts can make only harmful mistakes; they cannot intend to recommend W and mistakenly recommend R

$$\Longrightarrow s^* = \frac{V^R}{V^R - (1-q)\,V^W} + \frac{h}{(1-h)}\frac{(1-q)\,V^W}{(V^R - (1-q)\,V^W)} < \frac{V^R}{V^R - (1-q)\,V^W}$$

This suggests that the less skilled an expert, the less likely he is to engage in misconduct. All else equal, if an expert's experience or training is negatively correlated with the likelihood of making a mistake, then more experienced experts should have a greater rate of misconduct.

We summarize these findings below.

Proposition 3 *More error-prone experts are less likely to engage in misconduct.*

Corollary 4 *If the error rate h is negatively correlated with experience, more experienced experts engage in more misconduct.*

In summary, the model yields four main results:

1. Under low levels of effective monitoring, more heavily monitored experts are more likely to take advantage of customers.
2. The probability of misconduct is increasing in an expert's level of experience.
3. When the population of expert customers is sufficiently large, expert misconduct declines. Below this threshold, increases in the number of expert customers leaves the level of expert misconduct unchanged.
4. Customers are more likely to buy from more monitored experts.

2 The Insurance Industry

2.1 Insurance as Credence Goods

Insurance sales is a classic credence good market with price-taking experts. Products are complicated and multidimensional, and it is very difficult for even sophisticated consumers to identify the appropriate product for their needs. This is particularly true for life insurance and annuity products (LA) where insurers impose multiple "riders" and introduce modifications to policies that may be opaque to customers.[8] Consequently, a customer may be

[8]For example, life insurance policies can be term, universal, whole, variable and variable universal. In addition, a myriad of "riders" exist, including terminal illness and disability waivers, long-term care provisions, and accidental death benefits. The National Association of Insurance Commissions publishes a buyers' guide that describes some of the product complexities (http://www.naic.org/documents/consumer_guide_life.pdf).

sold an inappropriate product, but may never become aware of the seller's misconduct or mistake. With life insurance, the customer will never experience how well the policy serves his expected needs. Moreover, the insured customer and his beneficiaries may never learn whether there existed a superior product in the market at the time of purchase. Property and casualty insurance (PC) policies (e.g., auto or homeowners insurance) tend to be more understandable—the payouts and the conditions for payouts are often more transparent than other insurance products.

Insurance agents cannot adjust the prices faced by individual customers—indeed, this practice called "rebating" is illegal in most jurisdictions.[9] An insurance agent can enhance his commissions by recommending the wrong product to a customer. This increased revenue can come from simply "overselling" the level of insurance or from selling a product with a higher commission rate (i.e., percent of the customer's premium paid to the agent).

Commissions vary significantly across and within product types. For example, commissions from annuities typically range between 2 and 10% of the invested amount.[10] Typically, commission amounts are not disclosed to customers, allowing an agent to recommend an inferior product for a larger commission. In general, the tradeoff between the benefits to the policyholder and the revenue for the seller is substantial—for example, a so-called "bonus" annuity pays the customer an additional interest rate in the first year; however, the bonus rate and the commission rate are negatively correlated.

2.2 Monitoring and Organizational Forms

Insurance agents work primarily under two different organizational structures: agents work for large, branded companies that monitor their agents or as independent experts with little oversight.

Monitored company agents are typically affiliated with a single insurance company and may market only approved products from that company.[11] In practice, these product lists are quite large and there is little concern that company agents are too constrained. Companies using this organizational form may offer employment benefits packages and provide introductory training to inexperienced agents. New agents may also receive guaranteed salaries that phase out as they build up "books" of business, typically over 12 to 24 months. Company agents also have access to office space and administrative staff. Hierarchy within the firms ensures some level of monitoring—for example, branch managers may oversee and

[9] Rebating is illegal in our data environment (Texas Insurance Code CHAPTER 1806, Section 53).

[10] Our commission rate estimates and discussion of monitoring are based on personal communication with professional insurance agents.

[11] These agents may also be authorized to market selected products from other companies through agreements between their primary company and other firms.

approve large or complicated transactions. Company agents may earn 50 to 70% of the gross commissions of their sales, depending on the type of insurance product. State Farm, Farmers Insurance, Allstate, Northwestern Mutual and New York Life are examples of firms using the company agent model (A.M. Best 2011); in general, these firms have well-known, easily-recognized brand names.[12] We include a list of insurance companies using company agents in the Appendix.

In contrast, independent agents are not affiliated with a single insurance company. Typically, independent agents are responsible for all of their expenses; however, they generally earn 100% of the gross commissions on their sales. While independent agents are not restricted to selling insurance from any particular company, they usually cannot market products from insurance companies that use company agents—for example, an independent agent cannot market any State Farm products. Independent agents are often "one agent shops" and their transactions are not overseen by managers or supervisors. After accounting for business expenses, both company and independent agents earn roughly the same net commissions (Carson et al. 2007).

3 Data

Our Texas insurance agent dataset was compiled from multiple public sources and consists of licensing, appointment, complaint, and market share information. Broadly, the data cover the population of agents operating in the state and characterize both firm affiliations and reported incidents of misconduct in Texas's insurance industry.

3.1 Agents and Organizational Form

The licensing data were acquired from the Texas Department of Insurance (TDI) and cover all agents who were licensed to sell insurance in the state of Texas as of 2010. Overall, the data describe 235,604 agents: 60,812 agents are licensed to sell PC insurance only; 135,441 agents are licensed to sell LA only; and 39,351 agents hold licenses for both PC and LA. The licensing data include unique agent identifiers and the date on which each agent was first licensed in the state.

To identify the organizational form under which individual agents operate, we match company and appointments data from two sources. Company-level data were acquired from

[12] In 2010, State Farm, AXA, Allstate and Metropolitan Life appeared in Brandz's report on the top 8 most valuable global brands in the insurance industry (http://c1547732.cdn.cloudfiles.rackspacecloud.com/BrandZ_Top100_2010.pdf).

A.M. Best (2011) and allow us to identify companies by marketing type–these data distinguish firms that use monitored company agents from ones that sell through unmonitored independent agents. We then obtained appointments data from the TDI for firms employing a company agent model. Appointments data list all agents designated to sell a firm's products. Using agents' license numbers, we match license holders to firms and, thus, characterize individual agents' affiliations. Through this process, we identify 59,511 individuals who work as company agents (25.3% of licensees in the state).

We also acquired marketshare data from the TDI, describing the in-state total premiums written for all firms operating in Texas.

3.2 Complaints

The TDI maintains a public directory of complaints against insurance companies, agents and agencies. We accessed data describing 501,553 unique complaints filed between 1996 and 2010. The directory reports the date and nature of the complaint, the line of coverage (PC or LA), the license number of the subjects of the complaint, and whether the complaint was deemed "justified" or "unjustified" by the TDI.[13] Complaints vary considerably, from claims disputes to accusations about unfair cancellations.

Many complaints, even those leveled at agents, relate to actions under the control of insurance companies (e.g., denial of claims and premium-related complaints). To focus on misconduct at the agent level, we narrow our analysis to a subset of complaints relating to individual agents' sales practices. We also consider only complaints about PC and LA sales.[14] Table 1 summarizes these agent-level complaints by line of coverage and whether the complaints were justified or not.[15] In total, we identify 23,088 accusations of sales misconduct leveled against 13,356 individuals. Approximately 56% of these complaints were found to be justified. We match the complaints data to the population of agents licensed as of 2010 and find that 8,240 of these agents were the subject of at least one complaint.

Table 2 presents summary statistics for complaints, reported separately for company and independent agents. Complaints against insurance agents are rare events. Incident rates for both justified and unjustified complaints in PC sales are approximately three times higher than rates for LA, consistent with the notion that LA products have more pronounced

[13] In the Appendix, we discuss the role of potential reporting bias due to differences in perceived payoffs across expert types. Given that reporting costs are low in practice, we do not expect any substantial bias in our estimates.

[14] We exclude complaints relating to medicare supplements and employment insurance sales.

[15] The TDI dataset indicates that 122 agent-level marketing complaints were referred to other agencies for investigation; the broad descriptions of these individual complaints include "Agent mishandling," "Excessive physical force," and "Misrepresentation." Because we do not know the outcomes of these investigations, we drop these complaints from the analysis.

credence good attributes. Note also that, as our theory suggests, complaint rates appear substantially lower for independent agents relative to company agents. Of course, these summary statistics do not reflect other differences, including agent experience and market share across organizational forms—we account for these factors in our next section of results.

Table 3 reports aggregate premium and marketshare statistics for Texas by organizational form. While firms using company agents hold the majority of the marketshare in PC, the opposite is true for LA.[16] Even accounting for the number of agents under each structure, the data suggest that firms using company agents account for more PC sales.

4 Results

The credence good model that we analyzed in Section 1 yields four main predictions. In the following section, we present empirical evidence for each prediction. In the first two sub-sections, we present strong evidence about the difference in misconduct rates between company and independent agents and show that misconduct increases with agent experience. In the final two sub-sections, we consider predictions about connoisseur consumers and discuss differences in buy-rates across agent types.

4.1 Prediction 1: Monitoring and misconduct

Proposition 1 predicts that, at lower levels of effective monitoring, monitored agents are more likely to take advantage of customers, relative to unmonitored agents.

First, we ask: All else equal, are monitored company agents more likely to have been the subject of a complaint (justified or unjustified), relative to unmonitored independent agents?[17] As suggested by Table 2, complaints against insurance agents occur very infrequently in the data—in Texas, fewer than 4% of agents have been the subject of a complaint and less than half of those complaints were considered justified by investigators. Since typical econometric techniques, including logistic regressions, may underestimate the probability of rare events (King and Zeng 2001a), we estimate a logit model with a correction for rare events bias.

[16] In their seminal work on property rights theory, Grossman and Hart (1986) apply their model to the insurance industry. They predict that company firms will hold the majority of marketshare in LA and the minority of marketshare in PC. Their predictions align with the insurance industry structure in the early 1980s—65% independent firms in PC and 12% independent firms in LA. These marketshares are the opposite of what we find in Texas using more recent data.

[17] This question captures most misconduct—conditional on receiving any PC complaint, only 29% of agents receive additional PC complaints; similarly, only 16% of LA agents receive multiple complaints.

King and Zeng (2001a) describe the intuition of the correction: while the large number of zeros in the data allow the density $(X|Y=0)$ to be estimated well, the scarcity of observations for the rare event means that $(X|Y=1)$ is estimated relatively poorly with tails that are systematically too short. That is, the max $(X|Y=0)$ can be estimated well, but the min $(X|Y=1)$ will always be above the true minimum. Figure 1, adapted from King and Zeng (2001a), illustrates the intuition for the case of a single regressor, X. The vertical bars represent the actual observations when $Y=1$ and the solid line is the true density from which those observations were drawn; there are sufficient observations to draw a smooth density when $Y=0$, shown as a dotted line. The estimate X^* such that $X>X^*$ generates $Y=1$ and $X<X^*$ generates $Y=0$ will be greater than the true threshold value. Phrased informally, the sparse observations in the lower tail of the $(Y=1)$ density are traded-off with the dense observations in the upper tail of the $(Y=0)$ density to minimize the error between the estimated and observed values. Zero observations are overweighted relative to ones, yielding a higher estimate of X^*. Thus, coefficients will be systematically attenuated and the predicted $\Pr(Y=1)$ will be too small. For further details about the estimator, see King and Zeng (2001a,b).

We estimate the following equation:

$$\Pr(Complaint_i = 1) = \frac{1}{1+e^{-Q_i}} \tag{4}$$

where $Complaint_i$ equals 1 when agent i has been the subject of at least one TDI complaint and where

$$Q_i = \alpha CompanyAgent_i + \beta X_i$$

where $CompanyAgent_i$ equals 1 when agent i is a monitored company agent and matrix X_i contains the agent-specific controls described below. Coefficient and variance estimates are then corrected using the method of King and Zeng (2001a).

Although the main thrust of our analysis is concerned with differences between monitored and unmonitored agents (coefficient α), our predictions also speak to the role of agent experience. Recall that PC and LA products vary in terms of the ease with which customers can understand the match between their needs and the policy. To capture potential differences across product with differing credence good qualities, we distinguish between agents and complaints for PC and LA. We include the following controls in X_i, summarized in Table 4:

Years since first licensed: As a proxy for agent experience, we calculate the years since an agent was first licensed to sell insurance in Texas. If agents were licensed in other states prior to licensing by the TDI, we will underestimate their professional experience; if agents allowed their licenses to lapse in some interim periods, we will overestimate their

experience.[18]

Out-of-state agent: All agents who market insurance to consumers in Texas must be licensed by the TDI. We use the address on agents' licenses to determine residency and include a dummy variable to indicate when an agent resides outside of Texas.

Professional designation: Insurance agents may seek certification from several professional organizations. In general, these organizations require members to complete course work and exams, and participate in continuing education. We matched agents to member lists for 11 designations.[19] In our empirical analysis, we include a dummy variable indicating whether the agent holds any professional designation. While only a small percentage of agents hold these credentials, more LA agents have completed certification programs relative to PC agents.

License type: We include a dummy variable to indicate whether an agent is licensed to sell only one type of insurance (i.e. PC *or* LA). Most agents are licensed to sell only one type of insurance and slightly more sell LA only.

Table 5 reports estimation results from equation (4) with the rare events correction. In both PC and LA regressions, company agents are more likely to have received a complaint, justified or not. We transform our estimated coefficients into odds ratio form in Table 5. Results suggest that the odds of a company agent receiving any PC complaint is 39% higher than the odds of an independent agent receiving a complaint.[20] Examining LA, the odds of a monitored company agent receiving any complaint is 98% higher than for an unmonitored independent agent. PC and LA company agents are substantially more likely to be the subject of a justified complaint, relative to their independent peers.

One might ask: Do firms using company agents systematically hire less honest agents? This seems unlikely given that company agent firms have established screening processes (e.g. applications, background checks, and interviews). In contrast, independent agents establish their own practices and are not subject to this initial screening. Moreover, dishonest company agents who are fired are unlikely to gain employment at another company agent firm, but can readily move into independent sales. Thus, the pool of independent agents may include former company agents who were terminated due to misconduct.

Results also suggest that the difference in misconduct across organizational forms is affected by the extent of products' credence good qualities. Namely, when comparing monitored and unmonitored agents, LA products—which require more trust from the consumer—are associated with even more misconduct.

[18] The date of licensing was not available for approximately 1.5% of agents (3,455 individuals) and we exclude these agents from the analysis.

[19] The designations are: CFP, ChFC, CLU, CAP, CASL, CLF, FSS, LUTCF, MSFS, MSM, and REBC.

[20] Recall that odds are defined as $\frac{\Pr(\text{complaint})}{\Pr(\text{no complaint})}$.

One might be concerned that customers of branded companies are more likely to file a complaint due to the perceived "deep pockets" of these large firms. We explore this possibility in the Appendix and conclude that this is cannot fully rationalize the observed differences between the complaints against monitored and unmonitored agents.

4.2 Prediction 2: The Role of Expert Skill

In section 1.4, we describe our corollary 4 that more experienced agents are more likely to take advantage of customers.

Across the specifications in Table 5, an additional year of agent experience increases the odds of receiving a complaint by roughly 7%. Of course, agents with more experience have had more opportunities to receive a complaint. However, in this section, we present results suggesting that longevity alone cannot explain the estimated effect of experience.

In Table 6, we present results of a Tobit specification with complaints per licensed year as the dependent variable to account for agent experience.[21] Across all columns, the coefficients on agent experience are similar and statistically significant. In terms of magnitude, an additional year of experience results in an additional 0.01 complaints per year. In Table 2, we reported mean complaints per year of approximately 0.01—our Tobit results suggest that for an average agent, another year of experience may more than double the agent's complaint rate.

Our estimates are a lower bound on the true coefficient for experience for three reasons. First, the longer an agent has been in business, the greater the proportion of "bad apples" in his cohort that has been weeded out through disciplinary actions, leaving agents who are more ethical on average.[22] Since complaints against these "bad apples" are no longer included in the data, we expect our estimates of the effect of experience to be biased towards zero. Second, since our complaint data span 15 years, we cannot observe early-career complaints against agents with more than 15 years of experience.

Finally, client attrition may also attenuate estimates of the effect of agent experience. Consider our dependent variable $\frac{complaints}{Years}$, where $Years$ is years of experience. Assume for now that there is no client attrition and an agent acquires 10 clients per year. In ten years, a new agent has acquired 100 clients. Suppose that the chance of receiving a complaint is 1% per client per year. This means that an agent with 10 years of experience should (in expectation) receive one complaint. In an agent's 20th year, he has 200 clients and should

[21] While the Tobit results are consistent in magnitude and significance to our rare event logit analysis, the data fail strict tests of normality and homoskedasticity.

[22] Our model can be readily extended to include exogenously unethical agents—"bad apples"—who always suggest W. Our central results remain the same.

expect two complaints. Thus, without attrition, complaints per year does not depend on experience. Now consider the role of client attrition. Over the past 10 years, an agent with 20 years of experience has acquired the same number of clients as an agent with only 10 years of experience. However, due to attrition, the number of clients that he retained from his *first* 10 years is now less than the number of clients from the more recent decade. Thus, assuming that the chance of a complaint is still 1% per client per year, we would expect the ratio of complaints per year of the agent with 20 years of experience to be less than the ratio of the agent with 10 years of experience. Thus, we underestimate the true effect of experience on complaints.

One might worry that the most ethical company agents become independent operators after building up experience in the industry. If true, this could drive the difference in complaint rates between monitored and unmonitored agents. However, on average, company agents have been licensed significantly longer than independent agents ($p < 0.01$). Instead, one might wonder if bad agents are being detected and fired by the firms using company agents. Although our data do not allow us to observe this directly, this sorting would work against our predicted effect. That is, we would expect to observe higher complaints rates for independent agents if this organizational form included former "bad" company agents.

Although this is not an explicit component of our theory model, it is worth noting the sign and significance of our coefficient estimate for out-of-state agents. In Tables 5 and 6, these agents appear to be less likely to face complaints of misconduct for both PC and LA. This aligns with the intuition that out-of-state agents, from whom it might be difficult to recover compensation in the event of a misdeed, must be more ethical in order to attract clients. Another simple explanation is that these out-of-state agents are being prosecuted by their domiciled state's regulatory agency. Unfortunately, we observe only regulatory actions by the TDI. Finally, our empirical estimates provide little evidence that agents with professional designations are any less likely to have been the subject of complaints. Because these agents represent only 1% of the population of agents, we are unable to determine empirically whether these designations indicate skill or are simply attempted signals.

4.3 Prediction 3: Connoisseur Customers

The third prediction of the model, described by Proposition 2, is that an increase in the population of knowledgeable customers will weakly reduce agents' misconduct. To test this prediction in our data, we include a variable that is an estimate of the percentage of the population within a 25-miles radius of the agent that is employed in the finance industry. Here, we are assuming that employment in this industry is correlated with more knowledge

about insurance products.

Using a distance algorithm, we calculated the distance between the geographic centroid of all Texas ZIP codes and matched ZIP codes to 2010 County Business Pattern data from the U.S. Census Bureau. After identifying all ZIP codes within 25-miles of an agent's business address, we aggregated the employment statistics.[23] We do not include potential client employment statistics for non-resident agents because they do not have a Texas business location; as a result, non-resident agents are excluded from the analysis in Tables 7 and 8.

The results in Table 7 suggest that the employment type of local populations has little, if any, influence on agent misconduct. The coefficients on employment in finance are negative for LA and positive for PC, but are very small in magnitude. Tobit estimates of the effect of employment in finance are similar in Table 8, but also fail to achieve statistical significance. Note that the inclusion of the employment measure and the resulting exclusion of non-resident agents has little impact on the other coefficients of interest.[24]

Recall that the theory predicts that only sufficiently high levels of consumer education will reduce misconduct—our empirical results suggest that the population of consumers in Texas may not have reached this threshold of financial consumer literacy. Note also that, holding fixed the degree of malfeasance, if more educated people are more likely to report a complaint, then complaint rates should be greater for experts working in more finance-oriented areas. This works against finding evidence showing that complaint rates *fall* when customers are more knowledgeable.

4.4 Prediction 4: Monitoring and Marketshare

The final prediction of the theory model is that monitored agents will face higher buy rates than unmonitored agents. Since buy rate data are not available on an individual customer-agent level, we infer buy rates from marketshares.

Let n_m be the number of agents under organizational form $m \in \{C, I\}$, where C and I denote company and independent agents, respectively. Suppose that $r \in \mathbb{N}$ is an agent's *potential* customer flow rate per year. Denote customers' buy rate for organization type m by b_m. For example, if $r = 15$ and $b_m = 0.4$, then an agent faces 15 potential new customers each the year, resulting in 6 new clients per year. Let $p \in (0, 1)$ denote the persistence rate of clients, defined as the percentage of customers who remain clients into the next year (i.e. $(1 - p)$ is the client attrition rate).

[23] We multiplied the mid-point of the employment size class with the number of establishments in that class.

[24] We also consider consumers' education levels using the percentage of the nearby population with a college education. Results are similar.

In the long run, we can express the number of clients for a given agent as

$$\text{total \# of clients} = \sum_{j=0}^{t} p^j b_m r$$

Now, assuming that an average client has an annual total premium payment of π, the total premiums per long-lived agent (i.e. as $t \rightarrow \infty$) is

$$\frac{1}{n_m}\frac{b_m r}{1-p}\pi$$

We make three simplifying assumptions about the nature of the market: 1) both organizational forms have the same customer flow rate, r; 2) both organizational forms have the same customer persistence rate, p; and 3) the size of premium paid by an average customer is the same across organizational forms. These assumptions are particularly strong for relatively young agents. To accommodate this challenge, we compare average agents with at least three years of experience.[25] In our data, conditional on having at least three years of experience, company agents have approximately 14 years of experience, while independent agents average 11 years. For exposition, we assume a persistence rate of 0.9.

We compare the theoretical total premiums by organizational form

$$\begin{array}{lc} \text{Monitored Company Expert} & \sum_{j=0}^{14} 0.9^j b_C r \pi n_C \\ \text{Unmonitored Independent Expert} & \sum_{j=0}^{11} 0.9^j b_I r \pi n_I \end{array}$$

To inform our empirical test, we return to Table 3 reporting aggregate premium levels for both PC and LA. Rearranging the expressions above, we find that

$$\frac{b_C}{b_I} = \frac{\sum_{j=0}^{11} 0.9^j}{\sum_{j=0}^{14} 0.9^j}\frac{n_I}{n_C}\frac{\text{Total Company Expert Premiums}}{\text{Total Independent Expert Premiums}} \approx 0.91$$

Recall that if $\frac{b_C}{b_I} = 1$, then customers buy from independent and company agents at the same rate. While our main theory model predicts that $\frac{b_C}{b_I} > 1$, we find a buy rate ratio that is slightly lower than 1. Of course, we make many assumptions in constructing this empirical

[25] New agents experience a steep learning curve and often are not fully operational in their early years. For example, agents may begin as a trainee for two to three years or work as an assistant to a more experienced agent. This accounts for the difference between our estimates here and the values in Table 4.

comparison, and we cannot test whether our approximation is in fact different from 1.

Our main model assumes that k is endogenous—k is a function of expert's misconduct level s. However, if k is instead exogenous, (e.g. a fixed cost regardless of s) and q is small, then it is straightforward to show that although company experts still commit greater misconduct than independent experts, they face the roughly similar buy rates b^* and $\frac{b_C}{b_I} \approx 1$.

4.5 Level of complaints

We can also consider the impact of organizational form, experience, and customer education on the level of complaints with an OLS regression that is conditional on an agent having received one or more complaints. These results are reported in Table 9.

Estimates suggest that, conditional on receiving at least one complaint, independent agents are more likely to have been the subject of multiple complaints—that is, while fewer independent agents have complaints, they are more likely to be repeat offenders. One plausible explanation is that there exists a distribution of propensity for agent malfeasance. Assume that the level of complaints increase with this propensity. Since independent agents are less likely to receive complaints, those who actually do must have a greater propensity for malfeasance on average. Hence, conditional on having any complaint filed, we expect these independent agents to have more complaints.

While there is little evidence that the presence of a professional designation is associated with greater incidence of expert malfeasance, we do find that the *level* of complaints is negatively related to having a designation. As previously discussed, it is not clear precisely what these designations represent—for example, they might reflect skill, signalling or other unobservable attributes. Since fewer than 1% of agents have any professional designation and 80% of agents with a complaint have received only one, we interpret this finding very cautiously.

4.6 National Complaints Data

We also obtained a national-level dataset of sales practice complaints against firms for 2008 to 2011, collected by the National Association of Insurance Commissioners (NAIC). Agent-level data are not available nationally and, in general, states do not disclose individual agent- and firm-level sales practice complaints.[26]

We aggregated the complaints by firm and matched the NAIC data to firm characteristics obtained from A.M. Best. Using the A.M. Best marketing type classification, we identified

[26] Unfortunately, citing privacy issues, the NAIC could not accommodate our request for state-level complaints by firm.

firms using either the company or independent agent models; we excluded firms using Internet and direct sales. Firms that appeared in the A.M. Best data but not in the NAIC complaint data were coded as having faced zero complaints. In total, the final dataset includes 1930 PC and LA firms. Note that all complaints in the NAIC data were deemed justified by state regulators.

Table 10 presents results from the national data. In the first column, the dependent variable is an indicator for whether the firm has faced any complaints during the time period; in the second column, the dependent variable is the total number of complaints. Both regressions include a control for the net premiums written by the firm as well as dummy variables for all states in which a given firm markets insurance. These state-level controls capture both differences in the regulatory environment and differences in states' reporting to the NAIC.[27]

Using a logit regression with a rare-events correction, we find evidence that firms using monitored company agents are more likely to face a justified sales complaint, relative to a firm using independent agents. The odds ratio suggests that PC firms using monitoring are 58% more likely to have received a justified complaint. For LA, the overall effect of monitoring is not statistically different from zero ($p = 0.15$).

We find similar results examining the total number of complaints in Table 10. Firms using monitored company agents face more justified complaints. Again, we see that the effect is being primarily driven by PC firms. For LA, the overall effect of monitoring is not statistically different from zero ($p = 0.74$).

Overall, an analysis of these national data support our previous results, suggesting that, on average, monitored agents are more likely to take advantage of customers relative to independent agents.

5 Conclusion

In this paper, we explore how monitoring affects the level of misconduct in credence good markets with price-taking experts. Guided by theory, we find empirical evidence supporting the prediction that these markets operate differently than in standard asymmetric information problem settings.

In particular, rather than experts with strong reputations behaving more ethically, branded experts are actually less ethical in equilibrium. Similarly, experts who survive over time and become more skilled exhibit the greatest levels of misconduct. The intuition is as follows: in our setting, experts are price takers and thus extract surplus based on the value of their

[27]State reporting to the NAIC is voluntary. States' definitions of sale practice complaints may also vary.

firm's monitoring and their own skills through increased malfeasance. We find substantial empirical evidence that these predictions hold in the insurance industry.

Our work provides some preliminary suggestions for managing the credence good market problem. For low levels of monitoring, increases in monitoring may actually increase the level of misconduct; however, very intense monitoring will restore the market. Of course, intense monitoring is particularly challenging in credence good markets since detection is difficult. Our theoretical and empirical findings also suggest that regulators should focus their efforts on more experienced experts—not only do these experts have more customers, they are also more likely to take advantage of their clients. Of course, in practice, high levels of costly monitoring may not be feasible.

We also present theoretical results that show that increases in the population of expert consumers—those who can better discern misconduct—have the two-fold positive effect of increasing expert revenues and, with sufficient numbers, restoring the market. This suggests that informational campaigns to educate customers could prove promising. Our findings suggest that regulators should actually emphasize customer education over expert monitoring. Intuitively, while monitoring only provides a "stick" in the event of bad advice, the presence of informed consumers disciplines dishonest expert behavior by limiting the gains from misconduct while rewarding honest advice with higher purchase rates. A natural or field experiment, where consumers are randomly endowed with more information specifically about a credence good, would be enlightening.

References

[1] Alger, Ingela and Francois Salanie. 2006. "A Theory of Fraud and Overtreatment in Experts Markets." *Journal of Economics & Management Strategy*, 15(5): 853-881.

[2] A.M. Best Company. 2011. "Best's Key Rating Guide: Life/Health United State & Canada." Oldwick, New Jersey.

[3] Anagol, Santosh, Shawn Cole, and Shayak Sarkar. 2012. "Understanding the Incentives of Commissions Motivated Agents: Theory and Evidence from the Indian Life Insurance Market." Working paper.

[4] Bolton, Patrick, Xavier Freixas, and Joel Shapiro. 2007. "Conflicts of interest, information provision, and competition in the financial services industry." *Journal of Financial Economics*, 85: 297-330.

[5] Carson, James, Randy Dumm, and Robert Hoyt. 2007. "Incentive Compensation and the Use of Contingent Commissions: The Case of Smaller Distribution Channel Members." *The Journal of Insurance Regulation*, 25(4): 53-67

[6] Cooper, Robert, and Gary Frank. 2005. "The Highly Troubled Ethical Environment of the Life Insurance Industry: Has It Changed Significantly from the Last Decade and If so, Why?" *Journal of Business Ethics*, 58(1): 149-157

[7] Darby, Michael and Edi Karni. 1973. "Free Competition and the Optimal Amount of Fraud." *Journal of Law and Economics*, 16(1): 67-88.

[8] Dulleck, Uwe and Rudolf Kerschbamer. 2006. "On Doctors, Mechanics, and Computer Specialists: The Economics of Credence Goods." *Journal of Economic Literature*, 47: 5-42.

[9] Emons, Winand. 1997. "Credence Goods and Fraudulent Experts." *The RAND Journal of Economics*, 28(1) Spring: 107-119.

[10] Feddersen, Timothy and Thomas Gilligan. 2001 "Saints and Markets: Activists and the Supply of Credence Goods." *Journal of Economics & Management Strategy*, 10(1) Spring: 149-171.

[11] Fong, Yuk-fai. 2005. "When Do Experts Cheat and Whom Do They Target?" *The RAND Journal of Economics*, 36(1): 113-130.

[12] Grossman, Sanford, and Oliver Hart. 1986. "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration." *Journal of Political Economy*, 94(4) August: 691-719.

[13] Gruber, Jonathan and Maria Owings. 1996. "Physician Financial Incentives and Cesarean Section Delivery." *RAND Journal of Economics*, 27(1) Spring: 99-123.

[14] Harsanyi, J.C. 1973. "Games with randomly disturbed payoffs: a new rationale for mixed-strategy equilibrium points." *International Journal of Game Theory,* 2: 1-23.

[15] Hubbard, Thomas N. 1998. "An Empirical Examination of Moral Hazard in the Vehicle Inspection Market." *RAND Journal of Economics*, 29(2) Summer: 406-426.

[16] Hubbard, Thomas N. 2002. "How Do Consumers Motivate Experts? Reputational Incentives in an Auto Repair Market." *Journal of Law and Economics*, 45 October: 437-468.

[17] Inderst, Roman and Marco Ottaviani. 2009. "Misselling through Agents." *American Economic Review*, 99(3): 883-908.

[18] Inderst, Roman and Marco Ottaviani. 2011. "How (not) to pay for advice: A framework." Working paper.

[19] Inderst, Roman and Marco Ottaviani. 2012. "Competition through Commissions and Kickbacks." *American Economic Review*, 102(2) April: 780-809.

[20] King, Gary and Langche Zeng. 2001a. "Logistic Regression in Rare Events Data." *Political Analysis*, 9(2): 137-163.

[21] King, Gary and Langche Zeng. 2001b. "Explaining Rare Events in International Relations." *International Organization*, 55(3) August: 693-715.

[22] Kreps, David M. 1990. "Corporate Culture and Economic Theory." in James E. Alt and Kenneth A. Shepsle, Eds. *Perspectives on positive political economy*, New York: Cambridge University Press: 90-143.

[23] Levitt, Steven D. and Chad Syverson. 2008 "Market Distortions When Agents Are Better Informed: The Value of Information in Real Estate Transactions." *Review of Economics and Statistics*, 90(4): 599-611.

[24] Mailath, George J., and Larry Samuelson. 2001. "Who Wants a Good Reputation?" Review of Economic Studies, 68 (2) April: 415-441.

[25] Mullainathan, Sendhil, Markus Noeth, and Antoinette Schoar. 2012. "The Market for Financial Advice: An Audit Study." NBER Working Paper No. 17929.

[26] Pessendorfer, Wolfgang and Asher Wolinsky. 2003."Second Opinions and Price Competition: Inefficiency in the Market for Expert Advice." *The Review of Economic Studies*, 70(2): 417-437.

[27] Pitchik, Carolyn and Andrew Schotter. 1987. "Honesty in a Model of Strategic Information Transmission." *American Economic Review*, 77(5): 1032-1036.

[28] Sulzle, Kai and Achim Wambach. 2005. "Insurance in a Market for Credence Goods." *The Journal of Risk and Insurance*, 72(1): 159-176.

[29] Tadelis, Steven. 1999. "What's in Name? Reputation as a Tradeable Asset." *American Economic Review,* 89(3): 548-563.

[30] Taylor, Curtis R. 1995. "The Economics of Breakdowns, Checkups, and Cures" *Journal of Political Economy*, 103(1) February: 53-74.

6 Appendix

6.1 Different Payoffs and Reporting Rates for Customers

Overall, we find strong empirical evidence of our prediction that monitored company experts are more likely to take advantage of customers than unmonitored independent experts. However, one might be concerned that company agents' access to resources in the event of allegations of misconduct (i.e. "deeper pockets") might induce more customer complaints, relative to independent agents.

Assume that, in expectation, a customer is harmed more by the misconduct of an independent expert, relative to a company expert. That is, the expected value of reporting an abuse *conditional on conviction* is greater for the customer of a company expert. If the cost of filing is very low, then almost every discovered abuse should be reported and we would not see any difference in the ratio of justified to total complaints across organizational forms.[28] However, if there exists some material cost of filing a complaint, then customers of company experts will report suspected misconduct more often. If company and independent experts are equally ethical, then company experts will face more reported complaints.

To illustrate, let g_i be the probability that agent i is guilty of misconduct and let g_i be distributed uniformly between 0 and 1. Suppose that the expected payoffs to a customer after any conviction of a company or independent agent is \$1,000 or \$500, respectively. Let customers' reporting costs be \$100. A customer will not report an expert unless her expected net payoff from doing so is (weakly) positive. Therefore, the company agent's customer reports all cases where $g_i \geq 0.1$ and the independent agent's customer reports all cases where $g_i \geq 0.2$.

Define g^* as the threshold at which the customer chooses to report suspected misconduct. Now, the expected conviction rate given a report of the suspected impropriety is

$$\begin{aligned}\Pr(conviction|reported) &= \frac{\Pr(conviction \cap reported)}{\Pr(reported)} = \\ \frac{\int_{g^*}^{1} \Pr\left(guilty\right) f dg}{\Pr(reported)} &= \frac{\int_{g^*}^{1} g dg}{1-g^*} = \frac{1+g^*}{2}\end{aligned}$$

where f is the density of g.

For our example above, conditional on being reported, $g^* = 0.2$ leads to a conditional

[28] Empirically, the reporting cost is expected to be low. Customers can go online to the TDI website and fill out a form in a matter of minutes. Insurance policies also must list contact information for filing a complaint.

conviction rate of 60%, while $g^* = 0.1$ yields a conditional conviction rate of 55%. Thus, the company expert will have an unconditional conviction rate of $90\% \times 55\% = 49.5\%$ and the independent expert will be convicted $80\% \times 60\% = 48\%$ of the time.

More generally, we can write

$$\Pr(conviction) = \frac{1 - (g^*)^2}{2}$$

If reporting costs are low, g^* will be small for both independent and company customers. Hence, we would expect little distinguishable differences in reporting and conviction rates between expert types. By focusing on convictions (i.e., justified complaints) rather than all complaints, any potential difference is further minimized.

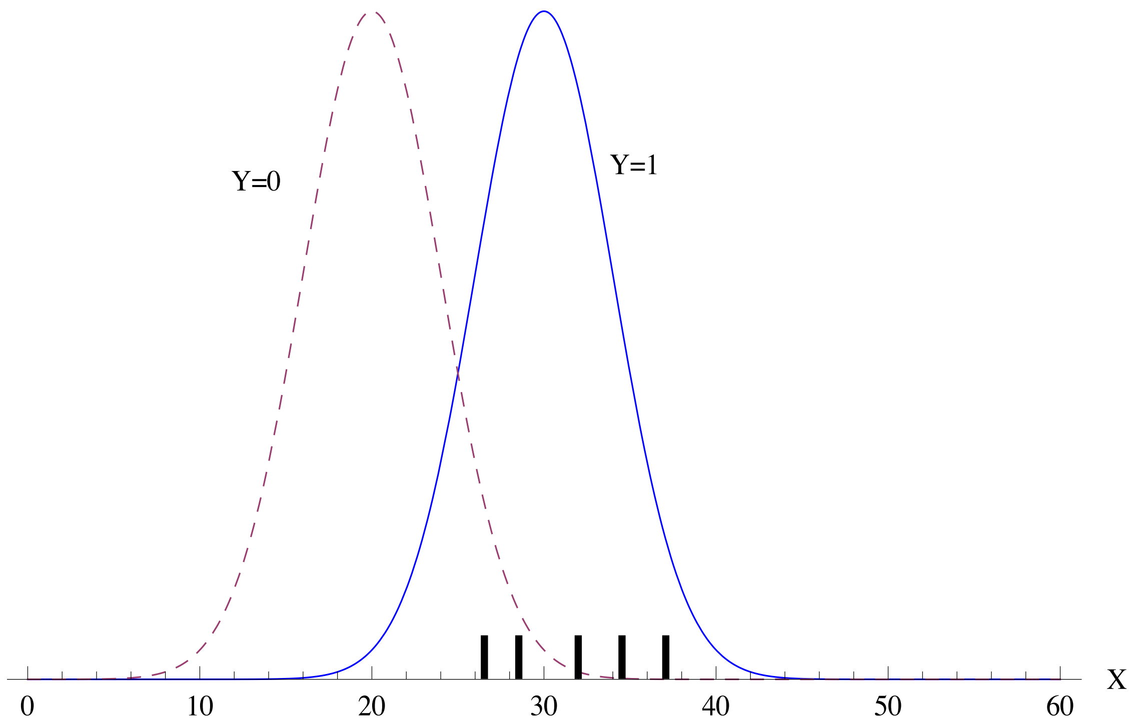
Empirically, there is also a countervailing force reducing reporting rates for company experts' customers. Recall that part of the monitor's role is to resolve disputes before they reach the regulator. In contrast, customers of independent experts have little recourse before contacting the regulator. As a result, in the data, we might expect observed complaint rates for company experts to represent a lower bound.

6.2 Insurance Companies Using Company Agents in Texas

Allstate Life Insurance Company
American General Life And Accident Insurance Company
American National Insurance Company
Axa Equitable Life Insurance Company
Baltimore Life Insurance Company
Beneficial Life Insurance Company
Farmers Insurance Exchange
First Acceptance Insurance Company
Guideone Mutual Insurance Company
Kansas City Life Insurance Company
Liberty Mutual Insurance Company
Metropolitan Life Insurance Company
Modern Woodmen Of America
Monumental Life Insurance Company
MONY Life Insurance Company Of America
Mutual Of Omaha Insurance Company
National Life Insurance Company
Nationwide Mutual Insurance Company
New York Life Insurance Company
Northwestern Mutual Life Insurance Company
Penn Mutual Life Insurance Company
Pennsylvania Life Insurance Company
Physicians Life Insurance Company
Provident American Life & Health Insurance Company
State Farm Life Insurance Company
Thrivent Financial For Lutherans
Western And Southern Life Insurance Company

A list of the 945 insurance companies licensed in Texas that use independent agents is available upon request.

Figure 1: Illustration of Rare Events Bias



Adapted from King and Zeng (2001a,b).

Table 1: Number of Agent-Level Complaints by Line of Coverage - Reasons and Outcomes

Nature of Complaint	**Property and Casualty Insurance**			**Life Insurance and Annuities**		
	Justified	**Unjustified**	**Total**	**Justified**	**Unjustified**	**Total**
Agent Mishandling	4746	4541	9287	836	2421	3257
Inappropriate Attitude	1	1	2	0	0	0
Churning	0	0	0	66	39	105
Commissions	89	192	281	73	190	263
Conversion	2981	127	3108	629	39	668
Failure to Provide Discount	4	7	11	0	0	0
Improper Inducements	33	16	49	37	10	47
Marketing Ethics	0	1	1	2	1	3
Misleading Advertising	74	51	125	190	57	247
Misrepresentation	340	131	471	965	1782	2747
Pressure to Take Higher Deductible Sales	1	3	4	0	0	0
Tie-In Sales	2	2	4	3	0	3
Twisting	0	0	0	20	28	48
Unauthorized Acts	993	286	1279	943	135	1078
Total Complaints	**9264**	**5358**	**14622**	**3764**	**4702**	**8466**

Table 2: Complaint Summary Statistics

		Property and Casualty	Life Insurance and Annuities
Monitored Company Agents	# of licensed agents	20032	56314
(n = 59,511)	# of agents with justified complaints	1124	812
	% of agents with a justified complaint	5.6%	1.4%
	# of agents with unjustified complaints	1451	1549
	% of agents with a unjustified complaint	7.2%	2.8%
	mean # of justified complaints per agent	0.028	0.018
	std dev	0.270	0.184
	mean # of unjustified complaints per age	0.032	0.034
	std dev	0.239	0.248
	mean # of justified complaints per year	0.005	0.001
	std dev	0.035	0.014
	mean # of total complaints per year	0.012	0.004
	std dev	0.054	0.026
Unmonitored Independent Agents	# of licensed agents	80131	118478
(n=176,093)	# of agents with justified complaints	1501	627
	% of agents with a justified complaint	1.9%	0.5%
	# of agents with unjustified complaints	1542	1081
	% of agents with a unjustified complaint	1.9%	0.9%
	mean # of justified complaints per agent	0.015	0.005
	std dev	0.282	0.097
	mean # of unjustified complaints per age	0.012	0.008
	std dev	0.150	0.116
	mean # of justified complaints per year	0.003	0.001
	std dev	0.033	0.015
	mean # of total complaints per year	0.005	0.002
	std dev	0.045	0.024

Note: 16,835 company agents are licensed to sell both PC and LA; 22,516 independent agents are licensed to sell both PC and LA.

Table 3 - Total Premiums and Marketshares by Organizational Form

	Agent Type	Total Premiums Written (in millions $)	Marketshare in %	# Agents	Premium per Agent (in thousands $)
Life Insurance and Annuities	Monitored Company	5661.39	11.20	56314	100.53
	Unmonitored Independent	44880.97	88.80	118478	378.81
	Total	50542.36		174792	289.16
Property and Casualty	Monitored Company	12082.20	62.97	20032	603.14
	Unmonitored Independent	7105.36	37.03	80131	88.67
	Total	19187.56		100163	191.56

Table 4 - Agent Summary Statistics by Insurance Type

	Property and Casualty		**Life Insurance and Annuities**	
	Mean	**Std. Dev**	**Mean**	**Std. Dev**
Monitored Company Agent Indicator	0.196	0.397	0.320	0.466
Agent Years Licensed	8.529	8.021	8.894	8.529
Texas Non-Resident Indicator	0.385	0.487	0.417	0.493
Professional Designation Indicator	0.003	0.058	0.011	0.106
One License Type Only Indicator	0.616	0.486	0.780	0.414
	n=98,435		n=171,476	

Table 5 - Logit results for Any and Any Justified Complaints By Agent with Rare Events Correction

Dependent variable: 1 if Agent has received any or any justified complaints, 0 otherwise

	Property and Casualty				***Life Insurance and Annuities***			
	Any Complaint		**Any Justified Complaint**		**Any Complaint**		**Any Justified Complaint**	
	Coefficient	Odds Ratio	Coefficient	Odds Ratio	Coefficient	Odds Ratio	Coefficient	Odds Ratio
Monitored Company Agent	0.331***	1.392	0.190***	1.209	0.681***	1.976	0.611***	1.842
	(0.035)		(0.046)		(0.036)		(0.056)	
Agent Years Licensed	0.064***	1.066	0.064***	1.066	0.073***	1.076	0.069***	1.071
	(0.001)		(0.002)		(0.001)		(0.002)	
Texas Non-Resident	-2.801***	0.061	-2.768***	0.063	-2.004***	0.135	-1.981***	0.138
	(0.122)		(0.162)		(0.076)		(0.119)	
Professional Designation	-0.306	0.736	-0.898**	0.407	-0.056	0.946	-0.433	0.649
	-0.214		-0.369		-0.159		-0.298	
One License Type Only	-0.681***	0.506	-0.623***	0.536	0.731***	2.077	0.793***	2.210
	(0.045)		(0.059)		(0.042)		(0.068)	
Constant	-3.339***		-3.945***		-5.303***		-6.202***	
	(0.037)		(0.049)		(0.051)		(0.080)	
N	98435		98435		171476		171476	

Note: Values in parentheses are robust standard errors.
** $p < 0.05$, *** $p < 0.01$

Table 6 - Tobit results for Total and Justified Complaints per Year by Agent

	Property and Casualty		***Life Insurance and Annuities***	
	Complaints Per Year	**Justified Complaint Per Year**	**Complaints Per Year**	**Justified Complaint Per Year**
Monitored Company Agent	0.0508***	0.0253***	0.0829***	0.0704***
	(0.0063)	(0.0078)	(0.0049)	(0.0073)
Agent Years Licensed	0.0098***	0.0098***	0.0092***	0.0084***
	(0.0003)	(0.0004)	(0.0003)	(0.0004)
Texas Non Resident	-0.3660***	-0.3471***	-0.2256***	-0.2182***
	(0.0143)	(0.0184)	(0.0084)	(0.0130)
Professional Designation	-0.0533	-0.1315**	-0.0163	-0.0629*
	(0.0370)	(0.0543)	(0.0220)	(0.0378)
One License Type Only	-0.0942***	-0.0821***	0.0968***	0.0994***
	(0.0070)	(0.0087)	(0.0058)	(0.0088)
Constant	-0.6321***	-0.7385***	-0.8067***	-0.9584***
	(0.0110)	(0.0159)	(0.0141)	(0.0260)
N	98435	98435	171476	171476

Note: Values in parentheses are standard errors. Complaints per year is calculated based on agents' years since first licensed in Texas.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7 - Logit results for Any and Any Justified Complaints By Agent with Rare Events Correction (in-state agents only)

Dependent variable: 1 if Agent has received any or any justified complaints, 0 otherwise

	Property and Casualty				***Life Insurance and Annuities***			
	Any Complaint		**Any Justified Complaint**		**Any Complaint**		**Any Justified Complaint**	
	Coefficient	Odds Ratio	Coefficient	Odds Ratio	Coefficient	Odds Ratio	Coefficient	Odds Ratio
Monitored Company Agent	0.349***	1.418	0.203***	1.225	0.694***	2.002	0.622***	1.863
	(0.037)		(0.049)		(0.039)		(0.060)	
Agent Years Licensed	0.065***	1.067	0.065***	1.067	0.070***	1.073	0.068***	1.070
	(0.002)		(0.002)		(0.002)		(0.002)	
Professional Designation	-0.188	0.829	-0.777**	0.460	-0.069	0.933	-0.39	0.677
	(0.217)		(0.372)		(0.180)		(0.328)	
One License Type Only	-0.661***	0.516	-0.594***	0.552	0.715***	2.044	0.787***	2.197
	(0.047)		(0.062)		(0.044)		(0.071)	
Fraction Finance Workers within 25mi.	0.001**	1.001	0.002***	1.002	-0.001***	0.999	-0.002**	0.998
	0.000		(0.001)		0.000		(0.001)	
Constant	-3.409***		-4.069***		-5.216***		-6.134***	
	(0.041)				(0.055)		(0.086)	
N	56906		56906		95343		95343	

Note: Values in parentheses are robust standard errors.

** $p < 0.05$, *** $p < 0.01$

Table 8 - Tobit results for Total and Justified Complaints per Year by Agent (in-state agent only)

	Property and Casualty		***Life Insurance and Annuities***	
	Complaints Per Year	**Justified Complaint Per Year**	**Complaints Per Year**	**Justified Complaint Per Year**
Monitored Company Agent	0.0543***	0.0281***	0.0826***	0.0710***
	(0.0067)	(0.0084)	(0.0050)	(0.0076)
Agent Years Licensed	0.0101***	0.0101***	0.0082***	0.0079***
	(0.0003)	(0.0004)	(0.0003)	(0.0004)
Professional Designation	-0.0339	-0.1141**	-0.0154	-0.05
	(0.0384)	(0.0565)	(0.0243)	(0.0407)
One License Type Only	-0.0918***	-0.0777***	0.0890***	0.0951***
	(0.0076)	(0.0095)	(0.0056)	(0.0089)
Fraction Finance Workers within 25mi.	-0.0889**	-0.0269	0.0288	0.0497
	(0.0405)	(0.0446)	(0.0340)	(0.0488)
Constant	-0.6374***	-0.7592***	-0.7496***	-0.9216***
	(0.0120)	(0.0176)	(0.0143)	(0.0268)
N	56886	56886	95317	95317

Note:Values in parentheses are standard errors. Complaints per year is calculated based on agents' years since first licensed in Texas.

** $p < 0.05$, *** $p < 0.01$

Table 9 - Level of Complaints Conditional on Complaints>0

	Property and Casualty		*Life Insurance and Annuities*	
	Complaints Per Year	**Justified Complaint Per Year**	**Complaints Per Year**	**Justified Complaint Per Year**
Monitored Company Agent	-0.0059	-0.0131**	-0.0106***	-0.0159***
	(0.0043)	(0.0053)	(0.0037)	(0.0051)
Agent Years Licensed	-0.0069***	-0.0065***	-0.0071***	-0.0076***
	(0.0003)	(0.0004)	(0.0003)	(0.0004)
Texas Non Resident	-0.0322*	-0.0410**	0.0332*	0.0104
	(0.0189)	(0.0182)	(0.0190)	(0.0180)
Professional Designation	-0.0249**	-0.0300***	-0.0170***	-0.0261***
	(0.0105)	(0.0105)	(0.0065)	(0.0092)
One License Type Only	0.0395***	0.0390***	0.0069**	0.0015
	(0.0093)	(0.0119)	(0.0034)	(0.0057)
Constant	0.2543***	0.2411***	0.2397***	0.2516***
	(0.0069)	(0.0086)	(0.0079)	(0.0124)
R-squared	0.16	0.16	0.25	0.33
N	4349	2411	3499	1372

Note: Values in parentheses are standard errors. Complaints per year is calculated based on agents' years since first licensed in Texas.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 10 - Regression Results for National Complaints Data

Dependent Variable:	***Any Justified Complaint (2008-2011)***		***Total Number of Justified Complaints (2008-2011)***
	Logit with rare events correction		**Tobit**
	Coefficient	Odds Ratio	Coefficient
Monitored Company Agents	0.4608*	1.585	12.1178***
	(0.236)		(3.662)
LA dummy	0.1884	1.207	4.7786*
	(0.162)		(2.843)
Monitored Agents x LA dummy	-0.8429**	0.430	-13.5346**
	(0.342)		(5.572)
Net Premiums Written	X		X
State Fixed Effects	X		X
N	1930		1930

Note: Values in parentheses are standard errors; standard errors are robust for the logit.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$