Persuasion and empathy in salesperson-customer interactions

Julio J. Rotemberg∗
February 3, 2014

Abstract

I consider a setting where firms can ask their salespeople to spend time and effort persuading customers. The equilibrium often features heterogeneity, with some salespeople being asked to persuade customers while others are not. Persuasive messages can either increase or decrease welfare, with price and quantity data generally being insufficient to determine whether welfare rises or falls. If one supposes that salespeople with different levels of empathy sort themselves across different selling jobs, the cross-sectional correlation between a salesperson’s sales and his or her empathy, can be helpful in determining whether persuasion is good or bad for consumers. JEL: M31, D64, D83

∗Harvard Business School, Soldiers Field, Boston, MA 02163, jrotemberg@hbs.edu. I wish to thank seminar participants at Boston University for helpful comments.
Whether working in a retail establishment or making calls on the road, many salespeople seek to persuade people to buy the products they have on offer.\footnote{For an overview of many persuasive methods used by salespeople, see Cialdini (2006).} Expenditure on the sales force thus bears some similarity with expenditure on advertising. There are enough differences between persuasive sales pitches and advertising, however, that the literature’s neglect of the former seems unwarranted. This paper focuses on the consequences of two of these differences, which can be thought of as variability and necessity. Because sales interactions are private, it is possible for customers to have very different experiences with salespeople. In practice, this appears to generate considerable variability in customer satisfaction with their sales encounters. Customers describe some as “delightful” while others are “terrible” (Arnold \textit{et al.} 2005).

Second, sales pitches often remain unavoidable, though the internet has provided an increasingly popular escape from them.\footnote{Interestingly, the unavoidability of some internet advertisements has generated complaints as well.} In many instances, potential customers find it necessary to obtain crucial information (including price information) by interacting with a salesperson who is then in a position where he can force the potential buyer to listen to a persuasive message. The combination of variability and necessity gives sales pitches the opportunity to lower welfare in a manner that is not open to advertisements. As Becker and Murphy (1993) emphasize, most advertising messages are avoidable so that consumers do not become exposed to them unless their individual welfare is thereby increased. The result is that, within their framework, advertising reduces overall welfare only if it leads to substantial price increases. As they argue, the avoidability of advertising is consistent with the delivery of most advertising messages in media such as TV, radio and magazines where people are induced to hear these messages by the lure of enjoyable non-advertising content.\footnote{Since these media outlets care about holding on to their audiences, they also prevent advertisers from using messages that these audiences would find particularly bothersome. Outdoor advertisement, while less avoidable, constitutes only a small fraction of total advertising expenditure (about 3.5% in the U.S. in 2005 according to ZenithOptimedia’s June 2012 \textit{Advertising Expenditure Forecasts}). It too is subject to control and, perhaps in part for this reasons, does not seem to generate nearly as much distress as “terrible” shopping experiences.}

Given that sales pitches are not avoidable, consumers are potentially subject to “terrible”
ones that reduce their welfare. Indeed, both Westbrook (1981) and Goff et al. (1997) suggest that the quality of sales encounters has long-lasting effects on consumers. They even seem to affect the satisfaction that consumers derive from the products that they have purchased. This raises the question of whether it is possible to ascertain the presence of negative welfare effects from sales pitches in any manner that does not involve direct reports from consumers.

This paper suggests that the variability of sales encounters can be helpful in this regard because it means that jobs in which salespeople use persuasion coexist with jobs in which they do not. Potential salespeople thus have some choice regarding the amount of persuasion they wish to engage in. If this persuasion leads customers to be hurt, people who empathize more with customers ought to be less willing to accept or keep jobs that involve persuasion. As I demonstrate, this can lead to a negative correlation between a salesperson’s level of empathy and his or her sales volume. In the absence of any persuasive effort by any salesperson, such a negative correlation would be surprising since the only reason some salespeople sell more would be that their price is lower, and one would expect more empathetic salespeople to be particularly eager to have jobs in which they charge low prices.

Consistent with the idea that the persuasion techniques used by salespeople sometimes increase welfare while other times they reduce it, the empirical literature studying cross-sectional correlations between salesperson empathy and salesperson effectiveness has produced mixed results. The early study of Tobolski and Kerr (1952) found a positive relationship between a measure of dispositional empathy and the effectiveness of new car salesmen at two dealers. By contrast, negative correlations with various measures of effectiveness are reported by Lamont and Lundstrom (1977) for a sample of building supply salesmen. Lastly, Dawson et al. (1992) reports a very weak relationship between empathy and sales.4

4More recent regression studies seeking to explain the performance of salespeople have tended to focus on “customer orientation,” (CO) measured as in Saxe and Weitz (1982) rather than on empathy per se. The two variables are strongly correlated with one another (see Widmier (2002) and Stock and Hoyer (2005)) and closely related conceptually. The CO scale gives considerable weight to agreement with the statements “I try to achieve my goals by satisfying customers,” and “A good salesperson has to have the customer’s best interest in mind.” Both of these fit well with the way I model empathy, which is by supposing that the salesperson’s objective includes the utility of his customers. The meta-analysis of Franke and Park (2006) shows that the correlation between salespeople’s self-reported customer orientation and the level of performance as evaluated by their managers is quite weak.
One way of thinking about the extensive literature focused on consumer search is that it considers random encounters between consumers and salespeople, where the only role of salespeople is to give a price quote. This paper maintains the framework of random encounters but lets salespeople carry out costly effort to change the amount that consumers are willing to pay for the product. The model I consider is a simplified version of Burdett and Judd (1983) in that an exogenous fraction of consumers encounter a single salesperson while the rest encounters two.\(^5\) I demonstrate that, for intermediate values of the ratio of the cost of persuasion to the increased willingness to pay that results from persuasion, the equilibrium features some sales encounters with persuasion and others without. The model is thus able to account for the variability of the quality of sales encounters even within narrow industries.

As emphasized by Bagwell (2007), the welfare consequences of persuasive messages depend on whether one thinks of them as affecting people’s true utility (as in Becker and Murphy 1993) or whether one thinks that they only lead people to act as if their utility were different from what it actually is. I focus mostly on the former case for two related reasons. First, consumers do seem to be negatively affected by certain sales encounters and recall these vividly (Arnold et al. 2005). Second, some of the sales techniques that are discussed both by observers and by authors giving advice to potential salespeople seem quite capable of affecting the utility of purchasers. Nonetheless, I also derive some results for the extreme case in which salespeople affect the willingness to pay without having any direct effect on consumer welfare. In this case, persuasion remains bad for consumers because it raises prices. As a result, this case bears some similarity to the case where persuasive effort also reduces welfare directly.

One way in which a salesperson can genuinely increase a customer’s utility from buying a product is by matching the customer with a product variant that is particularly suitable for her (Goff et al. (1997)). However, Westbrook (1981) shows that salespeople do not affect

---

\(^5\)This is simpler than following Burdett and Judd (1983), who focus on the fraction that makes people *ex ante* indifferent between meeting one salesperson and meeting two.
customer satisfaction only via their helpfulness and this may explain the popularity of sales techniques that are directed at buyers' emotions. One recurrent theme in books directed at would-be salespeople, is that they should seek to be liked by their potential customers (prospects) (Girard 1989, p. 17, Baber 1997, p. 40). A technique that has been recommended for accomplishing this objective (Baber 1997, p. 39), and which has been used in practice (Bone 2006, p. 61) is to imitate prospects' attitudes and mannerisms, and some academic research suggests that this technique should indeed be effective (Byrne 1961). If this and related techniques work, the customer's induced altruism can both raise the value that she attaches to a purchase and increase her cost from not purchasing.  

Consumers rationally expect salespeople to be more pleased when they complete a sale than when they do not. As a result, the act of purchasing can give consumers some vicarious pleasure. It can also lead consumers to suffer some vicarious losses when they fail to purchase. Arguably, these vicarious losses can be increased if the salesperson convinces the consumer that he already expected the sale to be completed, because this can lead the consumer to feel more guilt or vicarious disappointment if he does not buy.  

This fits with NCR’s 1887 advice to its salespeople that they should “not ask for an order [but] take for granted that he will buy. Say to him ‘Now, Mr. Blank, what color shall I make it?’ or ‘How soon do you want delivery?’ ” Girard (1989, p. 44) and MDRT (1999, p. 26) give a number of contemporary examples of this approach. As MDRT (1999) states, the idea behind this technique for “closing” sales is to make it difficult for the prospect to avoid purchasing the good.  

The company described in Bone (2006) used an additional technique to raise the cost of not buying. Salespeople at this company were expected to obtain agreement from their cus-  

---

6This is broadly consistent with Calvó-Armengol and Jackson (2010) who model peers as able to exert “positive pressure,” where they lower a peer’s cost of taking an action and “negative pressure,” where they increase the cost of not taking the action in question.

7A rare, but nonetheless effective method to obtain a sale by evoking sympathy is reported in Bone (2006, p. 90). He depicts a salesperson who, when customers declared an intention not to buy, routinely “burst into tears” while saying that she was “in trouble with the boss.” This sales method led at least some customers to buy.

8See Friedman (2004, p. 126)
tomers about desirable aspects of the product before they offered a price quote. If customers did not agree to buy, salespeople would ask for their reasons. Salespeople would then subtly suggest that the customer was being “disingenuous” or “irrational” if he tried to wiggle out by backtracking on some of his earlier statements.

Since these techniques are unpleasant for would-be buyers, they would not appeal to salespeople whose utility depends positively on the utility of their customers. This should lead salespeople with relatively high levels of altruism (or empathy) for their customers to self-select out of jobs that require these techniques. The logic is the same as that in Prendergast (2007), who considers the job choices of bureaucrats that vary in their concern for their clients. A formal difference between the current model and that of Prendergast (2007) is that I neglect moral hazard. This is done both for simplicity and because the work of Bone (2006) suggests that there are effective mechanisms for monitoring whether salespeople use the techniques discussed above.

The paper proceeds as follows. Section 1 introduces the model of costly persuasive effort. As in most of the paper, this effort is assumed to affect the willingness to pay of all customers by the same amount. The section shows that, for intermediate values of the cost of this effort some salespeople actively “sell” their products while others do not. Crucially for the results that follow, those that do sell their product more frequently. Section 1 also shows that, as in Becker and Murphy (1993) prices and quantities sold depend only on how sales efforts affect the difference between the utility of buying and the utility of not buying, they do not depend on whether selling activities increase the attractiveness of the product or reduce the utility of not buying the good. Section 2 shows that the welfare effects of these two approaches are quite different and compares them to those that arise when true utility is not directly affected by sales pitches.

Section 3 discusses employee preferences to prepare the ground for the sorting analysis of Section 4. Selling pitches directed at the emotions of prospects need not succeed with all customers, and Section 5 considers the case where they do not. It shows that customers who are immune to persuasion are likely to be hurt by the presence of customers who are
susceptible, because the latter encourage some salespeople to charge high prices while trying to persuade some customers to pay them.\footnote{If one views people that are susceptible to persuasion as “naive” while those that are immune are “sophisticated,” the focus of this analysis bears some similarity to Eliaz and Spiegler (2006), Gabaix and Laibson (2006) and Armstrong and Chen (2009).} Section 6 contains some concluding remarks.

1 A deterministic model of persuasive selling

A large number $N$ of firms sell a good that they can procure at marginal cost $m$. Each one does so through a single salesperson.\footnote{The assumption that each firm has a single salesperson is made for simplicity of exposition. The model can equally well be interpreted as one where each firm has many salespeople, each of which is assigned a pre-specified job.} Each salesperson, in turn, is visited by $\mu$ prospects whose initial valuation for a single unit of the good equals $v$, and who have no use for additional units. Prospects can be thought of as being engaged in a one-time purchase, and have no ex ante information about individual salespeople. Therefore, the $\mu$ potential customers of each salesperson are drawn independently from the population of consumers. A fraction $\theta$ of all consumers visits just one salesperson, while the rest visit two.

Salespeople can expend extra effort and thereby increase their customers’ willingness to pay. As discussed in the introduction, they can do so by either increasing the utility of purchasing or by reducing the utility from failing to purchase. This is captured by supposing that, after a prospect interacts with salesperson $i$, he assigns a value of $v + g_i$ to obtaining the good from $i$ while he suffers a loss $\ell_i$ if he does not buy it from $i$. This implies that a consumer prefers buying from $i$ than not buying at all if

$$v + g_i - p_i \geq -\ell_i \quad \text{or} \quad v + k_i \geq p_i \quad \text{where } k_i \equiv g_i + \ell_i. \quad (1)$$

If the prospect visits only $i$, this condition induces him to purchase the good. If he visits two sellers $i$ and $j$, he buys from $i$ if, in addition,

$$v + g_i - p_i - \ell_j > v + g_j - p_j - \ell_i \quad \text{or} \quad v + k_i - p_i > v + k_j - p_j. \quad (2)$$

If (1) is true and (2) holds as an equality, the consumer is equally willing to buy from either salesperson and I suppose he chooses $i$ with probability $1/2$. If the inequality in (2) is
reversed and (1) holds for $j$, the consumer buys from $j$. It follows that, from the perspective of a consumer’s purchase decision, $g_i$ and $\ell_i$ matter only through their sum $k_i$. To simplify further, suppose that there exists only one strictly positive level of effort, and that this leads $k_i$ to equals the constant $k$. If, instead, the salesperson makes no effort, $k_i = 0$.

The model is further simplified by supposing that a firm $i$ wishing its salesperson to set $k_i = k$ in a particular customer contact, must pay him an extra $w^e$ to secure the necessary effort. This could be justified by supposing that employee effort can be monitored and $w^e$ is the employee’s actual cost of effort. Each firm sets a single effort level and price, and these apply to every interaction of its salesperson.\footnote{The assumption that salespeople use a single approach with all its customers fits with the findings of Van Dolen et al. (2002). Since a range of prices gives the same profits in equilibrium, this assumption is inessential at this point. This is no longer true below when I let salespeople differ in their attitudes towards customers.} If firm $i$ wishes its salesperson to set $k_i = k$, it has to compensate him an extra $w^e$ for each customer contact. For the moment, the level of this extra compensation, as well as the base wage $w$ are exogenously fixed.

While the number of firms is finite, they are treated as sufficiently numerous that the distribution of prices across firms can be well approximated by a continuous density. An equilibrium is then a distribution of prices such that no firm has an incentive to deviate. This requires that profits be the same for all prices that are charged by a positive density of firms while profits would be lower if firms charged other prices. Assuming that $\mu \theta (v - m) > w$, as I do throughout, the equilibrium satisfies

**Proposition 1.** a)\textit{(Burdett and Judd 1983)} If $w^e/k > 1$, all firms set $k_i = 0$ and the equilibrium price has support $\{\theta v + (1 - \theta) m, v\}$ with cumulative density function

$$F_n(p) = 1 - \frac{\theta (v - p)}{(1 - \theta)(p - m)}. \quad (3)$$

b) If $w^e/k < \theta$, all firms set $k_i = k$ and each firm earns expected profits of $\mu [\theta (v + k - m) - w^e] - w$. The equilibrium price distribution across firms has support $\{\theta (v + k) + (1 - \theta) m, (v + k)\}$ with cdf

$$F_a(p) = 1 - \frac{\theta (v + k - p)}{(1 - \theta)(p - m)}. \quad (4)$$
c) If $\theta \leq w^e/k \leq 1$, a fraction $\gamma = (1 - w^e/k)/(1 - \theta)$ of firms require their employees to set $k_i = k$ for each customer contact while the rest require $k_i = 0$. All firms have expected profits of $\mu \theta (v - m) - w$. The firms with $k_i = 0$ have a distribution of prices with support $\{p^*, v\}$ where

$$p^* = m + (v - m)\theta k/w^e,$$

and a cdf of prices given by

$$F_0(p) = 1 - \frac{\theta(v - p)}{(w^e/k - \theta)(p - m)}. \quad (5)$$

The firms with $k_i = k$ have a distribution of prices with support $\{p^-, p^* + k\}$ where

$$p^- = w^e + \theta v + (1 - \theta)m.$$

The cdf of their prices given by

$$F_k(p) = \frac{p - p^-}{(1 - w^e/k)(p - m)}. \quad (6)$$

When the cost of raising a consumer’s willingness to pay by $k$ exceeds $k$ itself, no firm asks its employees to make the requisite effort. The resulting equilibrium exhibits the same price dispersion as Burdett and Judd (1983) and is equivalent to one in which sales pitches are successfully banned. If, instead, $w^e/k < \theta$, the cost of raising consumers’ willingness to pay by $k$ is so low that every firm adopts the persuasion method in question.

The intermediate case where $\theta < w^e/k < 1$ is the most interesting one, because it leads to heterogeneity both in prices and in the tactics used by salespeople. Let $q_i$ denote the probability that the salesperson working for firm $i$ makes a sale to a customer, so that this salesperson’s expected sales equal $\mu q_i$. For this intermediate case, we have the following corollary of proposition 1.

**Proposition 2.** If, for two firms $i$ and $j$, $k_i = k$ and $k_j = 0$, $q_i \geq w^e/k \geq q_j$. The second inequality is strict unless the price of firm $j$ equals $p^*$ while the first inequality is strict unless the price of firm $i$ equals $p^* + k$. 

8
This shows that when the cost/benefit ratio of persuasion is in the range in which only some salespeople persuade, persuasion is carried out exclusively by salespeople whose prices lead them to sell with high probability. An intuitive way of thinking about this is that the persuasive effort is wasted when it does not lead to a sale so that it should not be carried out when the probability of selling is low. To see this somewhat more precisely, imagine those salespeople whose overall offer is so unattractive that they sell only to those prospects that have no alternate offer, so that they sell with probability $\theta$. These salespeople should charge a price that is equal to the consumer’s maximum willingness to pay, which is either $v+k$ or $v$ depending on whether they incur effort or not. Thus, incurring effort is only worthwhile for such salespeople if $w^e < \theta k$. At the other extreme, consider the salespeople who offer the best deal that is available in equilibrium so that they sell to all their prospects. If such salespeople did not incur the requisite effort and $w^e < k$, a deviating firm could increase its price by slightly less than $k$, ask its employee to make the effort and thereby offer a more attractive deal that would also raise its own profits. Thus, $w^e/k < 1$ is sufficient for the firms that sell often to persuade their customers while $w^e/k < \theta$ is necessary for all firms to do so.

When $\theta < w^e/k < 1$, the highest price charged by a firm with $k = k$ exceeds (by $k$) the lowest price charged by firms with $k = 0$. If the difference were larger, the firm with $k = 0$ would be making a more attractive offer so that its sales would be larger, contradicting Proposition 2. If the difference were smaller, a firm with $k = k$ could profitably deviate and charge a slightly higher price than any of those currently charged. It would make sales just as often as the firm currently charging the highest price (since its offer would remain more attractive than any of those made by firms with $k = 0$) and earn more.

If the lowest price charged by firms with $k = k$ (which equals $[m + w^e + \theta(v - m)]$) exceeds $v$, the two distributions induced by the two levels of effort do not overlap. This requires that

$$w^e > (1 - \theta)(v - m),$$

which is satisfied if either few customers have access to two offers or if the cost of raising the
valuation by $k$ is substantial relative to the gap between consumer value and marginal cost. As a result, outcomes without overlap in the two price distributions may be relatively rare. Still, this case is conceptually interesting because it shows that this extension of Burdett and Judd (1983) makes it possible for the overall price distribution to have a hole, which is not possible in the standard case that excludes endogenous sales effort.

In spite of this strong effect of the choice of $k_i$ on the likelihood of selling, the availability of a technology to raise $k$ by spending $w^e$ has no effect on the overall distribution of sales. In particular,

**Proposition 3.** Regardless of $w^e$ and $k$, the distribution of the probability of selling $q$ across salespeople is uniform between $\theta$ and one.

The reason expected sales are uniform is that they depend linearly on the cumulative density function for prices, which is by necessity uniformly distributed.

## 2 The effects of employee effort on consumer welfare

This brief section discusses the implications for welfare of giving firms access to a persuasion technology that raises willingness to pay by $k$. I mostly focus on two special cases. In the first, only the utility of purchasing $g_i$ can be increased to equal $k$ rather than zero while $\ell_i = 0$. In the second, only the disutility of not purchasing $\ell_i$ can be raised while $g_i = 0$. As a shorthand, I refer to the former as the case where $g_i = k_i$ and the latter as the case where $\ell_i = k_i$. Though I emphasize it less, I also consider the case where the actual utility from consumption remains equal to $v$ after purchases and where the effect of salespeople on the willingness to pay can be ignored in welfare calculations. When prospects meet only one salesperson, this has the same effect as setting $\ell_i = k_i$ since the cost of not buying is not incurred in this case. Thus, the only difference between these cases is that people who meet two salespeople have an additional loss of $k$ when $\ell_i = k_i$

**Proposition 4.** Lowering $w^e$ below $k$ so that some firms set $k_i = k$ raises the expected utility of consumers if $g_i = k_i$ while it lowers it if either $\ell_i = k_i$ or if their “true” utility is
independent of the persuasive effort.

As already demonstrated, prices increase by the same amount in all three cases. The difference is that there is a direct positive effect on utility when \( g_i = k_i \) and this is sufficient to more than offset the effect of the price increase on utility. In the other cases, the rise in prices lowers utility and utility falls further if \( \ell_i = k_i \).

As noted in the introduction, a key difference with Becker and Murphy (1993) is that these sales messages are unavoidable if the consumer seeks to purchase the good. This still leaves the possibility that, knowing the some salespeople will seek to persuade them, consumers opt not to search for the good at all. If the persuasive effort has no effect on true utility or if \( \ell_i = k_i \), this will indeed occur if \( k \) is sufficiently large. What is more interesting is that consumers will continue shopping in spite of the welfare decline induced by persuasion if \( k \) is low enough. To see this, note first that consumers expect to receive some positive surplus from visiting salespeople when all firms set \( k_i = 0 \). The reason is that almost all firms charge less than the reservation price \( v \). By continuity, this ensures that customers can still expect to obtain positive surplus from visiting one (or two) salespeople when \( \ell_i = k_i = k \) as long as \( k \) is low enough. This raises the question of whether one can use information about salespeople to detect whether their persuasion is lowering welfare. I focus, in particular, on a characteristic of salespeople that has received considerable attention in the marketing literature, namely the extent to which they care about their customers.

3 The Preferences of Salespeople

There are two types of salespeople, \( \sigma \) and \( \lambda \), who differ in their empathy for their customers. Salespeople of type \( \sigma \) are selfish and care only about their effort and the financial compensation paid to them by their employer. In each period that they work in this industry, their utility is

\[
\omega^\sigma = w + \mu e(w^e - c) \tag{8}
\]

where \( w \) is their base wage, \( c \) is their cost of raising \( k_i \) from 0 to \( k \) for a single customer and
\(e\) is an indicator variable which takes the value of one if they are asked to make this effort. Otherwise \(e = 0\).

The salespeople of type \(\lambda\) are somewhat empathetic towards their customers so that their utility \(\omega^\lambda\) is given by

\[
\omega^\lambda = w + \mu e (w^e - c) + \tilde\lambda (u - \bar u)
\]

(9)

where \(\tilde\lambda\) is a parameter governing this type’s empathy (or altruism), \(u\) measures the expected utility obtained by each of the individual salesperson’s customers and \(\bar u\) is a baseline level of utility. The baseline level of utility \(\bar u\) is unimportant for the analysis below. For concreteness, it can be thought of as being the utility this salesperson expects these customers to obtain if he absents himself from his job. Regardless of the benchmark utility \(\bar u\), (9) captures the idea that salespeople with an empathetic disposition act as if they valued pleasing their customers.\(^{12}\)

Both when \(w^e > k\) and when \(w^e < \theta k\), salespeople only lose sales when their prospects meet another salesperson who charges a lower price. In the former case, the expected utility of a prospect who meets a salesperson with a price of \(p\) therefore equals

\[
u_c(p) = \theta(v + g - p) + (1 - \theta) \left[ \int_{\theta v}^p (v + g - \ell - y)dF_n(y) + \int_p^v (v + g - \ell - p)dF_n(y) \right]
\]

\[
= v + g - (1 - \theta)\ell - \left[ \theta + (1 - \theta)(1 - F_n(p)) \right]p - (1 - \theta) \int_{\theta v}^p ydF_n(y).
\]

(10)

where either \(g\) or \(\ell\) equal \(k\) depending on whether \(k_i = g\) or \(\ell_i = k\) and where both equal zero when the effect of persuasion on welfare can be ignored. In the first equation of (10), the first term involves the utility of a prospect who only meets the salesperson who charges \(p\) while the second one is the utility of a prospect who meets another salesperson and buys from that salesperson if her price is lower than \(p\). The expression in (10) is declining in \(p\) because customers are happier when they buy at a lower price, and this happiness gives

\(^{12}\)As discussed in footnote 2, the more recent empirical literature on sales performance has focused on “customer orientation” rather than empathy. While the Saxe and Weitz (1982) customer orientation scale involves answering 24 questions, several of them ask directly whether the salesperson cares about satisfying customers.
vicarious benefits to altruistic salespeople. Since salespeople that charge lower prices have a higher probability of selling, this force should lead more altruistic salespeople to seek jobs with higher sales volumes. Because of technical complications I describe below, I do not analyze this force in depth and focus instead on the correlation between sales and empathy induced by the sorting of salespeople into jobs with different levels of persuasive effort.

Salespeople who set \( e = 0 \) and charge \( p \) lose sales both when their customers meet another salesperson with a lower price and when they meet one that has set \( e \) equal to one. In the latter case, they lose the sale as long as the other salesperson charges less than \( p + k \). Thus, the expected utility of a customer that meets a salesperson with \( e = 0 \) and a price of \( p \) is

\[
u_0(p) = \theta(v - p) + (1 - \theta) \left\{ (1 - \gamma) \left( \int_p^{p_p} (v - y) dF_0(y) + \int_{p_p}^{v} (v - p) dF_0(y) \right) + \gamma \left[ \int_{p_p}^{p_p + k} (v + g - y) dF_k(y) \right] \right\}
\]

\[
= v + (1 - \theta) \gamma g - \left[ \theta + (1 - \theta)(1 - \gamma)(1 - F_0(p)) \right] p \\
- (1 - \theta) \left[ (1 - \gamma) \int_{p_p}^{v} y dF_0(y) + \gamma \int_{p_p}^{p_p + k} y dF_k(y) \right].
\]  

(11)

where \( g = k \) if this is the effect of setting \( k_i = k \).

By the same token, an employee who sets \( e = 1 \) sells to all those prospects whose other salesperson sets \( e = 0 \). Therefore, a salesperson who charges \( p \) with \( e = 1 \) expects his prospects to have a utility of

\[
u_1(p) = \theta(v + g - p) + (1 - \theta) \left\{ (1 - \gamma) \int_{p_p}^{v} (v + g - p) dF_0(y) + \gamma \left[ \int_{p_p}^{p_p + k} (v + g - p + \ell) dF_k(y) \right] \right\}
\]

\[
= v + g - (1 - \theta) \ell - \left[ 1 - (1 - \theta) \gamma F_k(p) \right] p - (1 - \theta) \gamma \int_{p_p}^{p} y dF_k(y).
\]  

(12)

where, again, either \( g \) or \( \ell \) can equal \( k \). The effect of \( e \) on the seller’s vicarious benefits depend on whether it \( g_i, \ell_i \) or neither is equal to \( k_i \). In the case where salesperson persuasion truly
raises the gain from buying by \( g \), \( g = k \) and \( \ell = 0 \) in equations (11) and (12). Therefore

\[
  u_0(p^*) = \theta(v - p^*) + (1 - \theta) \left\{ (1 - \gamma) \int_{p^*}^{v} \left( v - \frac{\theta v k}{c} \right) dF_0(y) + \gamma \int_{p^* - k}^{v} (v + y) dF_k(y) \right\} = u_1(p^* + k). \tag{13}
\]

Thus, when \( g_i = k_i \), the lowest price with \( e = 0 \) gives the same utility to employees of type \( \lambda \) as the highest price with \( e = 1 \). It follows that

**Proposition 5.** Suppose that \( g_i = k_i \) and let \( \omega^\lambda_i \) denote the utility of salesperson \( i \), where this individual is of type \( \lambda \). Then, if there exist \( i \) and \( j \) such that \( \omega^\lambda_i > \omega^\lambda_j \), it must be the case that \( q_i > q_j \).

This proposition follows directly from the earlier analysis. For a given level of \( k_i \), sales and the utility of altruistic salespeople are positively related because \( u_i(p) \) is decreasing in price while Proposition 1 implies that lower prices lead to more sales. Moreover, Proposition 2 implies that firms with \( k_i = k \) provide higher output than firms with \( k_i = 0 \), with the two providing the same output only when the latter charge \( p^* \) while the former charge \( p^* + k \). Equation (13) implies that, with \( g_i = k_i \), consumer utility is also the same in these two circumstances since the higher price of the latter is matched by a higher perceived value of the good. Thus, utility only rises when output is increased and the proposition follows.

Now consider the opposite case, where \( \ell_i = k_i \) so that \( \ell = k \) in (11) and (12). The effect of \( e = 1 \) is now not only to increase the price that consumers pay if they encounter another salesperson with \( e = 0 \) but also to cause a loss of \( k \) if consumers encounter another salesperson with \( e = 1 \). Consumers welfare is thus strictly higher with the combination of a price \( p^* \) with \( k_i = 0 \) rather than a combination of price \( p^* + k \) with \( k_i = k \). In particular, (11) and (12) now imply that

\[
  u_1(p^* + k) = u_0(p^*) - k.
\]

There are two additional relationships between utility levels that prove to be important below for the analysis of the case where \( \ell_i = k_i \). These are the relationships between the
highest value of \( u_1 \) (namely that which accrues from a price of \( p^- \)) and both the highest and lowest utility levels obtained by consumers who meets a salesperson that sets \( e = 0 \). If \( u_1(p^-) > u_0(p^*) \), the highest level of expected utility is obtained by consumers who meet (and buy from) the lowest-priced firms who sets \( e = 1 \). By contrast, when \( u_0(v) > u_1(p^-) \) (which obviously implies that \( u_1(p^-) > u_0(p^*) \) is violated), salespeople of type \( \lambda \) that set \( e = 1 \) are all worse off than any employee of type \( \lambda \) that sets \( e = 0 \). I study these conditions under the assumption that the marginal cost \( m \) equals zero. The condition on parameters that lead \( u_1(p^-) \) to be smaller than \( u_0(v) \) is given in the following proposition.

**Proposition 6.** Suppose \( \ell_i = k_i \) and consider the case where \( \theta k \leq w^e \leq k \) so that some salespeople set \( e = 0 \) while others set \( e = 1 \). Suppose also that \( m = 0 \). Then,

\[
\begin{align*}
\quad u_0(p) &= v(1 - \theta) - \theta v \log(p/\theta v) - w^e \log(k/w^e), \\
u_1(p) &= (v - k)(1 - \theta) - w^e - (w^e + \theta v)[\log(p) - \log(w^e - \theta v)],
\end{align*}
\]

so that

\[
\begin{align*}
u_0(v) > u_1(p^-) &\iff 1 + (1 - \theta) \frac{k}{w^e} - \log \left( \frac{k}{w^e} \right) > -\frac{v}{w^e} \theta \log(\theta) \quad (16) \\
u_1(p^-) > u_0(p^*) &\iff (\theta v + w^e) \log(k/w^e) > k(1 - \theta) + w^e. \quad (17)
\end{align*}
\]

Condition (16) is very strong, since it ensures that every salesperson that sets \( e = 1 \) gives lower expected utility to his prospects than any salesperson that sets \( e = 0 \). It turns out that, nonetheless, there are reasonable parameter values for which (16) is satisfied. To obtain some insight as to why this is so, note first that the left hand side of (16) is positive since \( k/w^e > \log(k/w^e) \). Moreover, \( \theta \log(\theta) \) is negative because \( \theta < 1 \). For given \( \theta \) and \( k/w^e \), \( u_0(v) \) thus exceeds \( u_1(p^-) \) as long as \( w^e/v \) is large enough. Rises in \( w^e/v \) raise \( p^- \) relative to \( v \) since firms that set \( e = 1 \) have to recoup the cost \( w^e \). They thus raise the utility of consumers that receive an offer of \( v \) with \( e = 0 \) relative to that of consumers who receive an offer of \( p^- \) with \( e = 1 \). Notice that, since \( k/w^e \) is bounded between 1 and 1/\( \theta \), the relatively high value of \( w^e/v \) that is needed implies that \( k/v \) must be substantial as well.
The parameters $k/w^e$ and $\theta$ have more complex effects on condition (16). Raising $k/w^e$ above 1 increases the potential price a customer with an offer of $v$ might have to pay if his second offer comes from a high-priced salesperson with $e = 1$ and this raises $u_1(p^-)$ relative to $u_0(v)$. On the other hand, an increase in $k/w^e$ also raises the psychological cost that a customer with an offer of $p^-$ faces if he gets a second offer with $e = 1$ and this lowers $u_1(p^-)$ relative to $u_0(v)$. In the case of $\theta$, intermediate values lead to a maximum for $|\theta \log(\theta)|$, which favor $u_1(p^-)$ because they raise the expected price that a consumer who receives an offer of $v$ will have to pay if he receives a second offer from a salesperson with $e = 0$.

Too little is known to calibrate the parameters of (16) on the basis of empirical observations. Still, one would not expect the actual cost of effort $w^e$ to be substantial relative to the subjective value of the product $v$. It is thus worth knowing that a value of $w^e/v$ of about .23 is sufficient to make (16) hold for any $\theta$ in the case where $k = w^e$. The derivative of the left hand side of (16) with respect to $k/w^e$ is negative for $1 \leq k/w^e \leq 1/(1 - \theta)$ so that increases in $k/w^e$ above the value of 1 initially imply that higher values of $w^e/v$ are needed for $u_0(v)$ to exceed $u_1(p^-)$. However, smaller values of $w^e/v$ are again compatible with (16) if $k/w^e$ is sufficiently larger than $1/(1 - \theta)$, though this requires that $\theta$ be relatively small since the conditions of Proposition 6 require that $k/w^e$ be smaller than $1/\theta$. If, in particular, $k/w^e$ is equal to 10, so that $\theta$ is less than or equal to .1, (16) is true as long as $w^e/v$ is less than or equal to .12.

This analysis demonstrates that there are robust conditions under which (16) is true, and there are equally robust conditions under which it is violated. I demonstrate below that the validity of this condition has implications for the correlation of the empathy of salespeople with the volume of their sales. The same is true for the condition (17) that makes $u_1(p^-)$ larger than $u_0(p^*)$.

To satisfy this latter condition, $k$ must be sufficiently small. To understand the reason, note that consumers care only about price when $k$ is zero. More generally, the lowest price value of $k/w^e$ that makes (16) hold as long as $w^e/k$ exceeds $-\theta \log(\theta)/(2 - \theta)$ and this expression reaches a maximum of about .23 for $\theta$ equal to about .46. 

---

13When $k = w^e$, (16) holds as long as $w^e/k$ exceeds $-\theta \log(\theta)/(2 - \theta)$ and this expression reaches a maximum of about .23 for $\theta$ equal to about .46.
with \( e = 1 \) remains the lowest price overall when \( k \) is small and, in this case, the loss of \( k \) has a limited impact on consumer welfare. As a result, the lowest-priced firm with \( e = 1 \) remains the consumers’ favorite supplier and \( u_1(p^-) > u_0(p^*) \).

One possible outcome from having an employee absent himself from his job is that the \( \mu \theta \) customers who would have visited only him are unable to buy the good and get zero utility from this transaction. The \( \mu(1 - \theta) \) customers that have access to another offer, meanwhile, would obtain the expected utility from a single offer. The baseline utility \( \bar{u} \) that enters into (9) would then be given by

\[
\bar{u} = (1 - \theta) \left\{ (1 - \gamma) \int_{\theta v + c}^{\theta v + c + k} (v - y) dF_0(y) + \gamma \int_{\theta v + c}^{\theta v + c + k} (v + g - y) dF_k(y) \right\},
\]

where \( g = k \) when the effort raises \( g \) whereas \( g = 0 \) if it raises \( \ell \). Notice that, since the lowest value of \( u_0 \) is \( u_0(v) \), (18) and (11) imply that \( u_0(v) = \bar{u} \). Since the salesperson that charges \( v \) is providing consumers no surplus, it is reasonable to think of consumers as getting as much utility from such a salesperson as from one that is absent. While I included this definition of \( \bar{u} \) for concreteness, the behavior of the employees of type \( \lambda \) once they are in the industry is independent of this variable. This variable would play a role if, instead of treating their total supply as exogenous, the decision of individuals of type \( \lambda \) to enter this industry were modeled explicitly.

Before closing this section it is worth noting that the case where persuasion has no effect on true consumer welfare has some elements in common with the case where \( \ell_i = k_i \) and, in particular, it can make empathetic salespeople worse off when they set \( e = 1 \). This occurs, in particular, when (7) is satisfied because, as discussed when this inequality is derived, this ensures that all salespeople with \( e = 1 \) charge a strictly higher price than all salespeople with \( e = 0 \).

4 The correlation between empathy and sales

The purpose of this section is to demonstrate that the empathy of employees and their individual sales are positively correlated when employee effort raises customer’s utility but
can be negatively correlated when it reduces it. Since the employee’s type is binary, this correlation can be studied by treating the type as a dummy variable which equals 1 for employees of type \( \lambda \) and zero for employees of type \( \sigma \). Then, as is well known from the difference-in-differences literature, the regression coefficient of individual sales on the dummy variable representing the salesperson’s type equals the difference between the average sales of individuals of type \( \lambda \) minus the average sales of employees of type \( \sigma \). Let \( \eta^{\lambda}(q) \) represent the fraction of salespeople of type \( \lambda \) among all salespeople who sell \( q \). The average sales of people of this type are

\[
q_{\lambda} = \frac{1}{a} \int_{\theta}^{1} \eta^{\lambda}(q) q dq \frac{1}{1 - \theta},
\]

where I have made use of Bayes’ rule, the uniform distribution of \( q \) between \( \theta \) and 1 and the fact that the overall proportion of altruistic salespeople equals \( a \). The average sales of salespeople of type \( \lambda \) exceed those of type \( \sigma \) if \( q_{\lambda} \) exceeds the overall average of sales across firms \( (1 + \theta)/2 \).

The connection between empathy and sales thus depends on the way that salespeople of the two types sort themselves into jobs with different prices and effort levels. It might be thought that the simplest approach to this problem would be to consider the full-information setup of Rosen (1986) where workers know all the prices charged and the effort required by employers before choosing their jobs. Unfortunately, it is then impossible for the wage paid by firms to be independent of the price they charge unless only type \( \sigma \) workers are employed in equilibrium. The reason is that, if the wage is independent of price and some type \( \lambda \) workers are employed, these workers strictly prefer to be employed by firms that charge a low price. Such firms are thus able to raise their profits by decreasing their wage slightly.

If the wage that firms are required to pay depends on the price that they charge, the earlier analysis based on Burdett and Judd (1983) is invalid. Moreover, computing equilibrium prices in the case where a firm’s wage cost depends both on the price the firm charges and the price charged by others is likely to be complicated. I thus opt for a simpler solution, namely to suppose that the pricing decisions of firms are not known to workers when they decide which job to accept. It is worth stressing, however, that the conclusion that altruistic
workers are concentrated in the jobs whose prices and effort levels they find more attractive, is likely to remain valid if this complexity were included in the analysis.

I consider both a one-period model and a dynamic extension. In each period there are \( \mu \) new consumers per firm, and these consumers have the same properties as those assumed earlier (for example, a fraction \( \theta \) visits one salesperson while the rest visit two). At the beginning of each period there is also a queue of potential employees who decide, in sequence, whether to accept a job posted by a firm in the industry under study. Potential employees are assumed to accept one of the available jobs for the period if they are indifferent between doing so and not working in the industry. If they choose to work in other industries, potential employees of type \( \sigma \) obtain a reservation utility of \( \bar{w} \) per period.

Consider first a model with a single period. At the beginning of this period, potential firms publicly post their compensation terms, which consist of a base wage \( w_i \) and a compensation for effort \( w_e^i \). After firms post these offers, a series of potential salespeople are placed in a queue. These employees decide, in sequence, which job offer to accept, if any. After a job is accepted, it ceases to be available to potential employees occupying later positions in the queue.

There is a finite number \( A \) of potential employees of type \( \lambda \) at the head of this queue, and there is an unlimited number of potential employees of type \( \sigma \) after this. The assumption that potential employees of type \( \lambda \) are at the head of the queue is a simplification which is meant to capture that they are in sufficiently finite supply that they are inframarginal from the point of view of wage setting. I suppose, in particular that the fraction of empathetic employees \( a = A/N \) is smaller than \( \gamma(1 - \gamma) \) where \( \gamma \) is the fraction of firms that set \( k_i = k \). Notice that my assumption of a fixed finite supply obviates the need to explicitly discuss the reservation wage of employees of type \( \lambda \), although, implicitly, this wage must be low enough that these employees gain utility from accepting the compensation in this industry.

I now show that, if potential employees do not know the price that individual firms intend to charge, there is an equilibrium where all firms offer a base wage \( w \) equal to \( \bar{w} \) and set their effort compensation \( w^e \) equal to \( c \). Since firms can attract an unlimited supply of workers of
type $\sigma$ on these terms, any firm that offers a higher $w$ or $w^e$ is strictly worse off. Now consider a firm planning to set $e = 0$. If it deviates and offers a lower base wage, it cannot attract either employees of type $\lambda$, who have better options at the same effort level, or employees of type $\sigma$, who find the offer insufficient. The same logic applies to a firm planning to set $e = 1$ which offers either a lower base wage or a lower compensation for effort.

These wages apply in all three cases considered in Proposition 1. Since $w^e = c$, Proposition (1) implies that all firms ask their employees to make the same effort when $c < \theta k$ or $c > \theta$. Altruistic employees are then indifferent with respect to all the offers that are made in equilibrium. They pick employers at random, and there is no connection between employee altruism and the level of sales.

The more interesting case is where $\theta k < c < k$ so that only some firms ask their employees to set $e = 1$. Propositions 2 and 4 then imply immediately

**Proposition 7.** Suppose that there is a single period, that $\theta k < c < k$ and that firms credibly communicate to potential employees whether they will set $e = 0$ or $e = 1$. Then $\eta^\lambda(q) = a/\gamma$ if $q > c/k$ while it equals zero if $q < c/k$ when salesperson effort is good for consumers. When this effort makes consumers worse off and $E(u_1(p)) < E(u_0(p))$, $\eta^\lambda(q) = a/\gamma$ if $q < c/k$ while it equals zero for $q > c/k$.

When $e$ benefits consumers, there is a high proportion $\eta^\lambda$ of altruistic salespeople among those with relatively high sales so that the correlation of altruism with sales is positive. By contrast, the correlation is negative when $e$ hurts consumers. I now show that the result that negative correlation between empathy and sales is indicative of a situation where the selling effort reduces welfare extends to a setting that is somewhat more realistic along two dimensions. The first assumption of the previous analysis that can be questioned is that job seekers know the effort level required by firms, though they do not know the prices that they are expected to charge. This asymmetry is unattractive. Moreover, competition for workers forces all firms to offer the same compensation, so that firms are left with no reason to convey this information themselves. Bone’s (2006) description of how he was hired suggests that,
in fact, some sales-oriented firms make little effort in this regard. Instead of describing the job in detail, his employer simply expected those that did not “fit in” to move elsewhere. This brings up the second weakness of the analysis to this point, which is that its focus on a single period precludes labor mobility.

I therefore consider a multi-period model that starts in period 1 and where a worker of type \(i\) has a utility of

\[
E \sum_{t=1}^{\infty} \hat{\rho}^{t-1} \omega_i^t \quad i = \lambda, \sigma,
\]

where \(\hat{\rho}\) is a discount factor and \(\omega_i^t\) represents the period utility at \(t\). The period utility is given by (8) or (9) if the individual works in the industry. Individuals of type \(\sigma\) receive a per-period utility of \(\bar{w}\) outside the industry. The labor market in period 1 is identical to the one in the single-period model. It starts with firms posting compensation levels. I suppose that firms are not allowed to change these compensation offers in subsequent periods. After these offers are posted, a queue of potential employees chooses which, if any, firm to work for. The queue continues to be headed by \(A\) workers of type \(\lambda\), though workers are now assumed not to know the effort level required by potential employers when they are choosing whether to accept one of these offers.

Workers who accept a job at the beginning of period \(t\) are required to complete the job’s requirements for that period, and thereby learn the required level of effort. At the end of each period, a fraction \(\psi\) of workers leaves the industry for exogenous reasons. These workers are replaced by workers from the queue that forms at the beginning of the next period. To keep the total number of employees of type \(\lambda\) constant, the queue of potential new recruits that forms every period is headed by \(\psi AN\) workers of type \(\lambda\) with all subsequent potential employees being of type \(\sigma\).

The \((1 - \psi)N\) workers who are not forced to leave the industry at the end of the period can leave their employer at this point. If they do so, they take positions at the front of the queue of potential employees. In other words, if \(X\) employees quit their current job and are willing to remain in the industry, these \(X\) individuals get to pick first which of the \(\psi N + X\) open jobs they wish to join in the next period. When they make this choice, they again do

21
not know the effort level required by potential employers.

Before analyzing the mobility of workers, it is worth noting that an equilibrium continues to exist in which the base wage is equal to \( \bar{w} \) while \( w^e \) equals \( c \). Firms have no problem attracting employees at these wages, so raising wages only raises their costs. Similarly, a firm that lowers either of these forms of compensation is unable to attract workers in the initial period. At this equilibrium, workers of type \( \sigma \) are indifferent among all jobs and I thus suppose they stay at their existing job until they are forced to leave the industry.

Workers of type \( \lambda \), on the other hand, obtain more vicarious utility at those firms that have a higher \( u \), where this represents the expected utility of customers that meet these firms. Let \( z_t(\tau, u) \) denote the joint density at \( t \) (across jobs or firms) of \( u \) and the type \( \tau \) of employees. The cdf of the marginal distribution of expected customer utility \( u \) is denoted by \( G(u) \). This is constant over time because the distribution of prices and effort levels is constant and \( u \) depends only on these variables. While the overall probability that employees are of type \( \lambda \) is also constant, the joint distribution \( z_t \) varies over time as long as the assignment of employee types to jobs varies as well. My focus, however, is on its steady state \( z(\tau, u) \).

A worker of type \( \lambda \) expects to obtain a vicarious utility \( \mu \tilde{\lambda} u \) in every period in which his job gives customers an expected utility of \( u \). Maximizing the worker’s expected present discounted value of this vicarious utility is thus equivalent to maximizing the present value of the expected utilities \( u \) earned by the salesperson’s customers. Let \( V(u) \) represent, in steady state, this latter present value for a salesperson who provided an expected utility \( u \) to his customers in his previous job and who thus has the option of providing this expected utility once again by remaining in this job. By moving, this individual gets a random draw from the distribution of consumer utility levels offered by those firms that need to fill a job. Let \( H(u) \) denote the cdf of this steady state distribution. Then \( V(u) \) satisfies

\[
V(u) = \max \left( u + \rho V(u), \left[ \int (y + \rho V(y))dH(y) \right] \right)
\]

where \( \rho = \hat{\rho}(1 - \psi) \),

where this equation takes into account that the salesperson has a probability \( \psi \) of exiting the
industry in the next period. Since the term in square brackets in this equation is independent of \( u \), \( V \) is increasing in \( u \). There thus exists a critical value of \( u \), which I label by \( \hat{u} \) such that \( u + \rho V(u) \) is smaller than the term in square brackets for \( u < \hat{u} \) and is larger for \( u > \hat{u} \). For \( u = \hat{u} \) the two expressions are equal.

This implies that employees of type \( \lambda \) stay in jobs that provide customers an expected utility greater than or equal to \( \hat{u} \) while they leave all jobs that provide smaller utility. Thus,

\[
V(u) = \begin{cases} 
  u + \rho V(u) = u/(1 - \rho) & \text{if } u \geq \hat{u} \\
  \int (y + \rho V(y)) \, dH(y) = V(\hat{u}) & \text{if } u < \hat{u},
\end{cases}
\]

where the last equality follows from the fact that individuals are indifferent between staying and leaving at \( \hat{u} \). Combining these equations, we have

\[
V(\hat{u}) = \frac{\hat{u}}{1 - \rho} = \int_{-\infty}^{\hat{u}} y \, dH(y) + \rho H(\hat{u}) V(\hat{u}) + \int_{\hat{u}}^{\infty} \frac{y}{1 - \rho} \, dH(y),
\]

so that

\[
\hat{u}(1 - \rho H(\hat{u})) = \int_{-\infty}^{\hat{u}} (1 - \rho) y \, dH(y) + \int_{\hat{u}}^{\infty} y \, dH(y),
\]

or

\[
\rho = \frac{\hat{u} - \int_{-\infty}^{\infty} y \, dH(y)}{\hat{u} H(\hat{u}) - \int_{-\infty}^{\hat{u}} y \, dH(y)}. \tag{22}
\]

It follows from equation (22) that, when \( \rho \) equals zero, \( \hat{u} \) equals the mean of \( u \) (according to the distribution \( H \)) while it equals the maximum value in the support of \( H \) when \( \rho \) equals one. Moreover, it is easily verified that the right hand side of (22) is monotonically increasing in \( \hat{u} \). This implies that, for any \( \hat{u} \) above the mean of \( H \), one can use this equation to find a \( \rho \) such that this \( \hat{u} \) is the minimal level of utility that is required for an employees of type \( \lambda \) to stay at his current job. I take advantage of this in the analysis that follows by treating \( \hat{u} \) as given, with the understanding that one still needs to check that a \( \rho \) can be found to rationalize this \( \hat{u} \). For a given value of \( \hat{u} \), I now study \( H \) and steady state distribution of employees of type \( \lambda \) across jobs.

To do this, one must first clarify the properties of \( G(u) \), the probability that a job chosen at random offers a utility less than or equal to \( u \) to its individual customers. If all jobs had
the same $k_i$, this probability would equal the probability that the price exceeds that which

gives an expected utility of $u$. Thus, when $c < \theta k$ or $c > \theta$, we have

$$G(u) = 1 - F_i \left( u_i^{-1} (u) \right) \quad i = a, c$$

When $\theta k < c < k$, expected utility can be below $u$ as a result of encounters with

salespeople who set $e = 0$ or encounters with ones that set $e = 1$. As a result,

$$G(u) = (1 - \gamma) \left[ 1 - F_0 \left( u_0^{-1} (u) \right) \right] + \gamma \left[ 1 - F_k \left( u_k^{-1} (u) \right) \right].$$

As we saw, all salespeople of type $\lambda$ whose job gives customers an expected utility level

$u$ below $\hat{u}$ quit their jobs, while those whose jobs give higher utility stay. Job changers give

an expected utility level drawn from $H$. Therefore, if both $u^1$ and $u^2$ are no smaller than $\hat{u}$
or if both are strictly smaller

$$\frac{z(\lambda, u^1)}{z(\lambda, u^2)} = \frac{g(u^1)}{g(u^2)},$$

where $g(u)$ is the density of $u$ implied by $G$. It follows that all jobs with consumer utility

$u \geq \hat{u}$ have the same fraction $\eta$ of employees of type $\lambda$ while all those jobs with utility $u < \hat{u}$
have a fraction $\hat{\eta}$ of employees of type $\lambda$ where

$$\eta = \frac{z(\lambda, u)}{g(u)} \quad \forall u \geq \hat{u}$$

$$\hat{\eta} = \frac{z(\lambda, u)}{g(u)} \quad \forall u < \hat{u}$$

$$a = (1 - G(\hat{u}))\eta + G(\hat{u})\hat{\eta}. \quad (23)$$

The last equality follows from the fact that the overall fraction of employees of type $\lambda$ equals

$a$. Knowledge of $\eta$ is sufficient to obtain $H(u)$ from $G(u)$. To see this, note first that, if

$\bar{u} \leq \hat{u}$, the total number of posted jobs with an $u \leq \bar{u}$ is the sum of $\psi G(\bar{u})N$ (the number of
dissolved jobs with $u \leq \bar{u}$) and $(1 - \psi)\hat{\eta}G(\bar{u})N$ (the number of non-dissolved jobs with $u \leq \bar{u}$
held by employees of type $\lambda$). If, instead, $\bar{u} > \hat{u}$ the total number of posted jobs with $u \leq \bar{u}$
equals the dissolved jobs $\psi G(\bar{u})N$ plus the total number of jobs abandoned by employees of
type $\lambda$, $(1 - \psi)\hat{\eta}G(\hat{u})N$. Therefore

$$H(u) = \begin{cases} 
[\psi + \hat{\eta}(1 - \psi)]G(u)/[\psi + \hat{\eta}(1 - \psi)G(\hat{u})] & \text{for } u \leq \hat{u} \\
[\psi G(u) + \hat{\eta}(1 - \psi)G(\hat{u})]/[\psi + \hat{\eta}(1 - \psi)G(\hat{u})] & \text{for } u > \hat{u}.
\end{cases}$$

(24)

To characterize the equilibrium, one now needs the value of either $\eta$ or $\hat{\eta}$. The former is given in the following proposition.

**Proposition 8.** The steady state proportion $\eta$ satisfies

$$(1 - \psi)(1 - G(\hat{u}))\eta^2 - [a + \psi(1 - a) + (1 - \psi)(1 - G(\hat{u}))]\eta + a = 0. \quad (25)$$

For $0 < \psi < 1$, this equation has a root strictly between $a$ and one so that $\eta > \hat{\eta}$.

In the case where $\psi$ equals one so that every job terminates after one period, employees of type $\lambda$ spread themselves evenly across all available jobs so that $\eta = a$. For lower values of $\psi$, the departure of employees of type $\lambda$ from jobs with $u < \hat{u}$ leads them to be more concentrated in jobs that give them higher utility so that $\eta > a$. Employees of type $\sigma$ end up correspondingly concentrated in jobs whose $u$ is below $\hat{u}$. The mean value of $u$, which is a measure of the utility that employees of type $\lambda$ would obtain in a particular job, is thus larger for employees of type $\lambda$ than for employees of type $\sigma$.

In the case where $g_i = k_i$, Proposition 5 implies that higher values of $u$ are associated with higher expected sales. Thus, the over-representation of altruistic employees in jobs with high values of $u$ implies that their mean sales are higher as well, so that empathy and sales are positively correlated.

The next proposition shows that, both when $\ell_i = k_i$ and when persuasive effort has no direct effect on customer’s true welfare, the correlation between empathy and sales can be negative.

**Proposition 9.** Suppose $\ell_i = k_i$ and (16) is satisfied or that $e$ has no direct effect on true welfare while (7) holds. There then exist ranges of values for $\psi$ and $\rho$ such that employees of type $\lambda$ quit all jobs with $e = 1$ and stay at those jobs with $e = 0$. As a result, the proportion
\( \eta^\lambda(q) \) is equal to \( \eta \) for values of \( q \) below \( c/k \) while it equals \( \hat{\eta} < \eta \) for \( q > c/k \). The correlation between empathy and sales is therefore negative.

For these parameters, employees of type \( \lambda \) abandon jobs with \( e = 1 \) to look for new jobs. They thus end up being concentrated in jobs with \( e = 0 \). The reason there is more than a single combination of \( \psi \) and \( \rho \) that fulfills the requirements of Proposition 9 is that, when (16) holds with \( \ell_i = k_i \) or (7) holds with \( \ell_i = g_i = 0 \), altruistic salespeople obtain strictly less utility at the “best” job with \( e = 1 \) than at the “worst” job with \( e = 0 \). As a result, several discount rates lead the expected present discounted value of staying at any job with \( e = 1 \) to be lower than that of leaving while the opposite is true for the best job with \( e = 0 \).

The next proposition shows, however, that the correlation between empathy and sales can be positive even when persuasion by salespeople lowers welfare.

**Proposition 10.** Suppose \( \ell_i = k_i \) and (17) is satisfied. There then exist a critical price \( \hat{p} > p^- \) such that \( u_1(\hat{p}) = u_0(p^*) \) as well as values of \( \psi \) and \( \rho \) such that \( \hat{u} = u_0(p^*) \). As a result, the proportion \( \eta^\lambda(q) \) is equal to \( \eta \) for all levels of \( q \) greater than or equal to \( \hat{q} = 1 - \gamma(1 - F_1(\hat{p})) \) and equals \( \hat{\eta} < \eta \) for all levels of \( q \) smaller than \( \hat{q} \). The correlation between altruism and output is therefore positive.

Since (17) is satisfied, the cost \( k \) is relatively small so that consumers receive their highest possible utility when they encounter a salesperson with \( e = 1 \) that sets a price of \( p^- \). Empathetic salespeople are then willing to have their customers incur the cost \( \ell \) as long as they do so at firms that charge low enough prices. The result is that the sales of empathetic salespeople are relatively high.

The combination of Propositions 9 and 10 and the more general result for the case where \( g_i = k_i \) implies that, in the deterministic case I have considered so far, a negative correlation between empathy and sales is indicative of a situation where salespeople reduce welfare when they use persuasion to raise sales while a positive correlation is not enough to rule out the possibility that the sales effort is socially deleterious. I now turn to the case where the persuasiveness of salespeople is stochastic.
5 Heterogeneous consumer susceptibility to persuasion

When salespeople reduce welfare either because $\ell_i = k_i$ or because their persuasive effort has no effects on “true” welfare, their techniques are presumably geared to affect their customer’s emotions. It is reasonable to suppose, then, that not all potential customers are equally susceptible to these efforts. This section explicitly considers such a situation by supposing that only a fraction $\alpha$ of prospects raises their willingness to pay from $v$ to $v + k$ when they encounter a salesperson that sets $e = 1$. It shows that, even if salespeople do not know who is persuadable, equilibria may still feature some salespeople that incur effort. It also shows that the resulting increase in prices is bad for people who are immune to this effort.

I start by giving conditions on parameters under which there is no equilibrium in which all firms set $e = 0$ because some firms would benefit by deviating and setting $e = 1$. To do this, recall from Proposition 1 that expected sales for a firm that sets $e = 0$ and a price of $p$, $q_0(p)$ in an equilibrium where all firms set $e = 0$ satisfy

$$q_0(p) = \begin{cases} 
1 & \text{if } p < m + \theta(v - m) \\
\theta(v - m)/(p - m) & \text{if } p_0 \equiv m + \theta(v - m) \leq p \leq v \\
0 & \text{if } v < p.
\end{cases} \quad (26)$$

At this equilibrium, all firms earn expected profits per customer (net of the base wage) equal to $\theta(v - m)$. A firm that deviates and sets $e = 1$ while charging $p$ sells with probability $q_0(p - k)$ to susceptible customers while it sells with probability $q_0(p)$ to immune ones. Such deviations are thus profitable if a price exists such that the expected profits per customer (net of base wages)

$$[\alpha q_0(p - k) + (1 - \alpha)q_0(p)](p - m) - w^e$$

exceed $\theta(v - m)$. An equilibrium where all firms set $e = 0$ exists if and only if (27) is smaller than $\theta(v - m)$ for all $p$. To obtain more specific conditions one must distinguish between the cases where

$$(1 - \theta)(v - m) < k \quad (28)$$
holds, and the case where it is violated. In the latter case, \( k \) is small enough for a range of prices \( p \) to exist such that \( k + p_o^- \leq p \leq v \) where \( p_o^- \) is defined in (26). For these prices, (26) implies that both \( q_o(p - k)(p - k - m) \) and \( q_o(p)(p - m) \) are equal to \( \theta(v - m) \). Using this in (27), the profits from deviating to a \( p \) in this range equal
\[
\theta(v - m) + \alpha q_o(p - k)k - w^e.
\] (29)
The profits from this deviation are maximized by setting \( p \) equal to \( p_o^- + k \), which leads \( q_o(p - k) \) to equal one. Profits from deviating then equal \( [\theta(v - m) + \alpha k - w^e] \) so that, when (28) is violated, equilibria where all firms set \( e = 0 \) fail to exist if
\[
\alpha k - w^e > 0.
\]
Notice that lowering \( p \) below \( p_o^- \) reduces profits from deviating because the probability of selling to susceptible consumers does not rise above one while the price they pay falls. Deviating by setting a price above \( v \) is equally unattractive since it eliminates sales to the skeptical consumers while lowering the profits from the susceptible ones. Thus, when (28) does not hold, equilibria where all firms set \( e = 0 \) do exist if the above condition is violated as well.

Now turn to the case where (28) is satisfied. Two sorts of deviations are now worth considering. In the first, a deviating firm sets a price above \( v \) and turns away all immune customers while, in the second, it sets a price smaller than or equal to \( v \) and sells to some of them. The expected profits per customer from the first kind of deviation are given by the left hand side of (29) with the term in square brackets set to zero (because there are no sales to immune customers). These profits increase when \( q_o(p - k) \) rises, so that the most profitable price is \( p_o^- + k \). For this deviation to be profitable, it must be the case that
\[
\alpha k - (1 - \alpha)\theta(v - m) - w^e > 0.
\] (30)
In the second type of deviation, the deviating firm’s price \( p \) is no larger than \( v \) so that \( p - k \) is strictly smaller than \( p_o^- \). Equation (26) then implies that its profits per customer
are given by the left hand side of (29) with the term in curly brackets set equal to $\alpha(p - m)$ (because it sells to all its susceptible customers). This is maximized by setting the highest possible price, which here involves setting $p = v$. This deviation is thus profitable when

$$\alpha(v - m) + (1 - \alpha)\theta(v - m) - w^e > \theta(v - m) \quad \text{or} \quad \alpha(1 - \theta)(v - m) - w^e > 0. \quad (31)$$

When (28) holds, an equilibrium where all firms set $e = 0$ exists only if both (30) and (31) are violated.

I now turn to the study of equilibria in two contrasting situations where no equilibrium exists with all firms setting $e = 0$. In both of these situations, (28) holds. In the first, (30) holds as well, while (31) does not. This can be thought of as a situation where both $w^e$ and $k$ are “large.” Given the high cost of setting $e = 1$ and the substantial impact this has on susceptible customers, all the firms that set $e = 1$ charge a price above $v$ so that immune customers do not buy from them. In the second case I consider, (31) holds so $w^e$ is more modest. Together with the other conditions I impose, this leads to equilibria where firms that set $e = 1$ charge prices below $v$ and, indeed, charge prices that are also charged by firms that set $e = 0$. The result is that, at these equilibria, firms that set $e = 1$ have larger sales.

When, instead, firms that set $e = 1$ charge prices above $v$, their sales tend to be lower. I start by studying parameters that give rise to outcomes of this sort.

**Proposition 11.** Suppose that (31) is violated. Then, for any fraction of susceptible customers $\alpha > 0$ there exist equilibria where a fraction $\gamma > 0$ of firms set $e = 1$ and charge a price above $v$ while the rest set $e = 0$. These equilibria requires that $k$, $w^e$ and $\gamma$ satisfy

$$\alpha(1 - \gamma(1 - \theta))k - w^e = \left[\frac{1 - \alpha}{1 - (1 - \theta)\alpha \gamma}\right](\theta + (1 - \theta)(1 - \alpha)\gamma)(v - m) \quad (32)$$

which implies (28) and (30).

Let

$$p^* = m + \frac{\theta + (1 - \theta)(1 - \alpha)\gamma}{1 - (1 - \theta)\alpha \gamma}(v - m) \quad (33)$$

$$p^- = m + \frac{\theta + (1 - \theta)(1 - \alpha)\gamma}{\alpha}(v - m). \quad (34)$$
The equilibrium prices of firms that set $e = 0$ lie between $p^*$ and $v$ and have cdf

$$F_0(p) = 1 - \frac{[\theta + (1 - \theta)(1 - \alpha)\gamma](v - p)}{(1 - \theta)(1 - \gamma)(p - m)}. \quad (35)$$

The equilibrium prices of the firms that set $e = 1$ lie between $p^-_1$ and $p^* + k$ and have a cdf

$$F_1(p) = \frac{p - p^-_1}{(1 - \theta)\gamma(p - m)}. \quad (36)$$

When $\gamma = 0$, (35) reduces to (3) so that the equilibrium is the standard Burdett-Judd (1983) outcome displayed in part a) of Proposition 1. As $\gamma$ is increased, prices rise for two reasons. First, note that the violation of (31) implies that $p^-_1$ in (34) exceeds $v$. Thus, the fraction $\gamma$ of firms that sets $e = 1$ charges more than $v$, which was originally the highest price. This not only raises prices directly but also implies that firms that set $e = 0$ find themselves with a larger group of customers whose purchases are insensitive to price, namely the immune customers whose other salesperson set $e = 1$. As a result, profits of firms that set $e = 0$ must be higher, and this requires higher prices. Formally, $F_0(p)$ in (35) falls when $\gamma$ increases so that the distribution of prices conditional on finding a firm with $e = 0$ for a particular $\gamma$ stochastically dominates the corresponding distribution for a lower value of $\gamma$.

This increase in prices makes immune consumers worse off. If one interprets a simultaneous increase in $\alpha$ and $\gamma$ as an increase in the number of “unsophisticated” consumers coupled with an increase in the number of firms willing to devote themselves exclusively to serve them, then this increase is costly to “sophisticated” (i.e., immune) customers. This can be compared with Eliaz and Spiegler (2006), Gabaix and Laibson (2006), and Armstrong and Chen (2009), who also consider firms with customers with varying levels of sophistication.

As in Armstrong and Chen (2009) but unlike in the other two papers, the presence of unsophisticated customers is costly to sophisticated ones. The reason is similar in both models, namely that firms switch their product in the direction of those purchased by the susceptible agents. Here they do so by increasing their selling effort, which reduces the competition for immune consumers, while in Armstrong and Chen (2009) they do so by producing a low quality good that only naive customers are willing to buy.\(^{14}\)

\(^{14}\)While ignored in the current analysis, there is an even more direct way in which immune consumers
When \( \ell_i = k_i \), the effort \( e \) of salespeople also makes susceptible consumers worse off (since they pay more than the valuation \( v \) and sometimes lose an additional \( k \)). Thus, overall welfare is reduced by the availability of the persuasion technology in this case. Matters are more complex when \( g_i = k_i \) so that the utility of susceptible consumers is increased by the effort \( e \). Even if the result is a net gain in the welfare of susceptible consumers, average consumer welfare can still fall if the decline in the utility of immune consumers is large enough. I now construct a numerical example where this does indeed occur.

For all the simulations carried out in this paper, \( v = 1 \) and \( m = 0 \), where these choices are inessential. Suppose \( \theta \) is equal to .2 while \( \alpha \) takes values on a grid between 0.001 and .999. For each value of \( \alpha \), let \( \gamma = \alpha \) and set \( w^e \) so that it equals .01 plus the value that makes (31) hold as an equality. The first panel of Figure 1 displays the resulting values of \( w^e \) and \( k \), which is given by (32), as a function of \( \alpha \). The second panel displays, for the case where \( g_i = k_i \), the resulting values of expected utility. The expected utility of immune consumers is simply equal to the expectation of \( v - p \), where \( p \) is the price that they pay, while that of susceptible ones is the expected value of \( v + k - p \). The Figure also displays the average expected utility of consumers, which gives a weight of \( (1 - \alpha) \) to the former and a weight of \( \alpha \) to the latter.

This Figure shows that, for low values of \( \alpha \), average expected utility is declining in the proportion of susceptible consumers. The main reason for this is that, as just discussed, the utility of immune consumers declines as \( \alpha \) and \( \gamma \) increase. The effect of changes in \( \alpha \) on the utility of susceptible consumers depends to an important degree on the assumed changes in \( k \), so the figure does not clarify the effect of \( \alpha \), by itself, on this utility. Rather, the Figure is only intended to demonstrate that one can change parameters so that susceptible consumers are better off, as when \( \alpha \) increases beyond the value of about .13, while average utility falls because the losses of immune consumers outweigh the gains of susceptible ones.

This possibility that an effort that makes susceptible consumers better off is nonetheless pay for the resources involved in the selling effort. They do so by spending time listening to arguments by salespeople. As discussed extensively by Bone (2006), these often purposefully delay giving a price quote until after they have presented these arguments.
bad for consumers as a whole raises the question of whether it is now possible for the correlation between empathy and sales to be negative even though the effort \( e \) enhances the utility of certain consumers. The third panel of Figure 1 demonstrates that this is indeed possible. For \( \alpha \) above about .37, the expected utility of consumers that meet a salesperson with \( e = 1 \) and a price of \( p^- \) exceeds the expected utility of consumers that meet a salesperson with \( e = 0 \) and a price of \( p^* \).

As discussed earlier, this implies that the jobs that give the highest utility to altruistic salespeople are the ones with \( e = 1 \) and a price of \( p^- \). It thus becomes possible to find values of \( \psi \) and \( \rho \) so salespeople of type \( \lambda \) are mostly found in jobs with \( e = 1 \) and prices near \( p^- \). The Figure shows that, for values of \( \alpha \) between about .35 and .5, expected sales of these employees, \( q_1(p^-) \) are lower than average sales \( E(q) \). The parameters associated with \( \alpha \) between .37 and .5 thus lead to a negative correlation between altruism and sales even though persuasive effort is good for susceptible consumers. The parameters that accomplish this seem fairly special. Nonetheless, this demonstrates that empirical instances of negative correlations between empathy and sales need to be studied further before it is determined that the salespeople involved are causing harm to their consumers. This is particularly so because, as the Figure demonstrates, these negative correlations can be found for parameters where the persuasive efforts of salespeople also increase the average expected utility of consumers.

I now briefly discuss some implications of these parameters for the case where \( \ell_i = k_i \). These are depicted in the bottom panel of Figure 1. This panel shows that, for all these parameters, the expected utility of consumers who encounter a salesperson with \( e = 0 \) and a price of \( v \) exceeds the expected utility of consumers who meet a salesperson with \( e = 1 \) and a price of \( p^- \). Both of these salespeople give the same utility to immune consumers since these do not buy from the former and obtain no surplus from the latter. The difference is that susceptible consumers are at risk of losing \( k \) as soon as one of their salespeople sets \( e = 1 \).

For low values of \( \alpha \), salespeople who set \( e = 1 \) also have low sales. As \( \alpha \) increases, the expected sales of firms that set \( e = 1 \), \( E(q_1) \), rise above those \( E(q_0) \), the expected sales of those that set \( e = 0 \). In the numerical exercises this starts occurring when \( \alpha \) reaches .7.
From that point on, it is straightforward to choose values of $\psi$ and $\rho$ such that altruistic salespeople stay at jobs with $e = 0$ (which give utility of $u_0(p^*)$ or more) while they leave jobs with $e = 1$ (which give utility of $u_1(p^-)$ or less). There is then a negative correlation between empathy and sales. This demonstrates that, just as in the deterministic case, it is relatively straightforward to obtain parameters where the correlation between empathy and sales is negative if the effort of salespeople increases consumers’ cost of not buying from them. The reason is that such salespeople make consumers worse off while their sales tend to be high, at least if susceptible consumers are sufficiently numerous.

I now turn to a set of parameters that lead firms that set $e = 1$ to charge less than $v$. This leads such firms to have relatively high sales, since they also sell to immune consumers. Since these firms still impose costs on susceptible consumers when $\ell_i = k_i$, it becomes easy once again to find parameters such that the correlation between sales and altruism is negative.

**Proposition 12.** An equilibrium with a proportion $\gamma > 0$ of firms setting $e = 1$ exists if

\[ w^e = \alpha(1 - \gamma)(1 - \theta)(v - m), \]  
\[ \theta \geq \frac{(1 - \alpha)(1 - \gamma)}{1 + (1 - \alpha)(1 - \gamma)}, \]  
\[ (1 - \theta)^2 \alpha \gamma^2 - ((1 - \theta)\alpha + 1 + \theta)\gamma + 1 > 0, \]  
and

\[ v - m < k + p^* - m \leq \frac{\alpha(1 - \gamma)(1 - \theta) + \theta}{\alpha(1 - \gamma)(1 - \theta) + \alpha \theta} (v - m), \]  
where

\[ p^* - m = \frac{\theta(v - m)}{1 - \alpha \gamma (1 - \theta)}. \]

These conditions also ensure that there is no equilibrium where all firms set $e = 0$. Let $F_i(p)$ denotes the cumulative distribution function of the prices charged by firms with $e = i$. 

33
At the equilibrium where a fraction $\gamma$ of firms sets $e = 1$, these cdf’s satisfy

$$F_1(p) = \frac{1 + (1 - \alpha)(1 - \theta)\gamma}{\gamma(2 - \alpha)(1 - \theta)^2} - \frac{\alpha\theta(v - m)}{\alpha(2 - \alpha)(1 - \theta)(p - m)} \quad \text{for } p^- \leq p \leq v \quad (42)$$

$$F_0(p) = \frac{1 - \alpha\gamma(1 - \theta)}{(1 - \gamma)(1 - \theta)} - \frac{\theta(v - m)}{(1 - \gamma)(1 - \theta)(p - m)} \quad \text{for } p^* \leq p \leq p^- \quad (43)$$

$$= \frac{\theta}{(2 - \alpha)(1 - \gamma)(1 - \theta)^2} - \frac{\alpha\theta(v - m) - (1 - \alpha)w^e}{\alpha(2 - \alpha)(1 - \gamma)(1 - \theta)(p - m)} \quad \text{for } p^- \leq p \leq v \quad (44)$$

where

$$p^- - m = \frac{1 - \gamma(1 - \theta)}{1 + (1 - \alpha)(1 - \theta)\gamma}(v - m). \quad (45)$$

As shown below, there exist numerical values of the parameters that satisfy all the proposition’s conditions. Before showing these, it is worth understanding more qualitatively why the range of parameters where it applies is nontrivial. Note first that the term in square brackets in (40) is greater than one so that one can find a range of k’s that fulfill this condition for any values of the other parameters. Moreover, (38) is satisfied for $\theta$ larger than one half when $\alpha$ and $\gamma$ both zero and becomes less restrictive as either of these parameters becomes larger.

Finally, the quadratic equation on the left hand side of (39) has two positive roots, only one of which is smaller than one. Thus (39) is satisfied for values of $\gamma$ below a critical value that is between zero and one. Analysis of this equation reveals that this critical value is greater than or equal to 1/2 for any values of $\theta$ and $\alpha$. It equals 1/2 in the limit when $\theta$ goes to one, and is larger for smaller values of $\theta$.

Figure 2 presents some outcomes from a set of parameters satisfying the requirements of Proposition 12. The Figure is drawn under the assumption that $\gamma$ is equal to .3. It allows $\alpha$ to vary between .01 and .99. Whenever possible, the value of $\theta$ is set equal to .2. For values of $\alpha$ below .66, this would violate (38) and the value of $\theta$ is thus set just above the value that makes (38) hold. The Figure plots the resulting value of $\theta$ as well as $E(q_1) - E(q_0)$, the difference between the expected sales per customer for firms that set $e = 1$ and the expected sales per customer for firms that set $e = 0$. Not surprisingly, this is strictly increasing in the fraction of susceptible customers $\alpha$. A larger fraction of such customers ensures higher sales.
for the fixed fraction of salespeople that set $e = 1$.

Lastly, the Figure plots the difference $u_1(p^-) - u_0(v)$ between the average utility of customers who visit a salesperson with $e = 1$ and a price of $p^-$ and the average utility of customers who visit one salesperson with $e = 0$ and a price of $v$. This difference in utility, is drawn under the supposition that $\ell_i = k_i$ so that the effort of salespeople reduces utility. This difference in utility depends on $k$, with higher values of $k$ implying larger losses to susceptible customers who do not buy from a salesperson that set $e = 1$. The constraint (40) limits the allowable range of $k$, however. For $\alpha = .8$ and the other parameters of the Figure, for example, $k$ has to be between .75 and .82 (where $v$ has been normalized to equal 1 and $m$ has been set to equal zero). This differences has only a very mild effect on the results, so the Figure is drawn for $k$ set at the midpoint of the two extremes allowed by (40).

The effect of an increase in $\alpha$ is to reduce $u_1(p^-) - u_0(v)$. The reason is not that increases in $\alpha$ endogenously increase $k$. In fact both limits of (40) decline slightly when $\alpha$ rises. A factor that contributes to the effect of $\alpha$ is the induced decline in $\theta$, since this lowers $p^*$ and thus the average prices charged by firms that set $e = 0$. However, $u_1(p^-) - u_0(v)$ also declines with $\alpha$ in the part of the Figure where $\theta$ is constant. This occurs because an increase in $\alpha$ raises the number of susceptible customers who lose $\ell$ when they meet two salespeople who set $e = 1$ and this reduces $u_1(p^-)$.

The result is that, for $\alpha$ large enough, one can again construct equilibria with a negative correlation between sales and empathy. When $\alpha$ exceeds .5, $u_1(p^-) - u_0(v)$ is negative while there are sufficient susceptible consumers that $E(q_1) - E(q_0)$ is positive. The former implies that any job with $e = 1$ gives lower utility to altruistic salespeople than any job with $e = 0$. There thus exist values of $\psi$ and $\rho$ such that altruistic salespeople stay in the latter but quit the former. The fact that $E(q_1) > E(q_0)$ then implies that altruistic salespeople have lower sales on average.

I have not been able to determine whether it is possible, when $g_i = k_i$, to have a negative correlation between empathy and sales for the parameter configurations that satisfy Proposition 12. It is worth noting, however, that all the prices charged by firms with $e = 1$
are also charged by firms that set $e = 0$. For a given price, the sales of the former are clearly larger, since they sell to more susceptible customers. Moreover, if they must charge a particular price, salespeople of type $\lambda$ prefer to set $e = 1$ rather than $e = 0$ since this gives more utility when $g_i = k_i$. While they do not prove it, these observations suggest that the sales of employees of type $\lambda$ are likely to be greater than average.

6 Conclusions

Salespeople draw mixed emotions from their customers: sometimes customers like them, at other times they regard them with apprehension. This paper has sought to present a model in which both these situations are possible. One important message of the model is that the unavoidability of sales pitches makes them different from advertising messages and, in particular, allows sales pitches to have considerably larger negative welfare consequences. Unfortunately, as in Becker and Murphy (1993), price and quantity data are not sufficient to distinguish sales pitches that lower utility by making it more costly not to buy from sales messages that raise utility by increasing the attraction of buying. The reason is that price and quantity data depend only on the extent to which sales pitches change the willingness to pay as opposed to their effect on the utility of not buying.

A key result of the paper is that, nonetheless, there is another difference between sales pitches and advertising that permits one to gain some insight into whether sales interactions raise or reduce welfare. This is that sales pitches are delivered by individuals, and individuals differ in their willingness to deliver different sorts of sales messages. In particular, the paper shows that a negative correlation between a salesperson’s empathy towards customers and her level of sales can be indicative of a situation in which sales pitches lower welfare. One advantage of this diagnostic method is that correlations of this sort have been computed in the empirical marketing literature and that this literature has displayed both negative and positive correlations. This empirical literature has relied on questionnaires to elicit empathy measures but other approaches may be feasible as well. One might, in particular, use measurements of mirror neuron responses, as in Singer et al. (2004), who show that these
responses are stronger for subjects who score higher on the Balanced Emotional Empathy Scale of Mehrabian and Epstein (1972).

A question raised by the paper is how the analysis would change if salespeople were given more independence regarding the messages that they use. Monitoring the behavior of salespeople, as described in Stroh (1978) and Bone (2006), requires resources. It is thus not surprising that, while many organizations closely evaluate the way that salespeople behave with customers, others evaluate them mostly based on their sales outcomes.\(^\text{15}\) In the case where the actions of salespeople are unobservable to their supervisors, the sorting of employees across jobs would resemble even more closely the analysis of Prendergast (2007). Still, some of the results of the current analysis may well extend to such a setting. The fact that altruistic salespeople would try to avoid “hard sell” tactics should continue to lead their sales to be relatively low in industries where such tactics are particularly effective. By the same token, their sales should be relatively high in industries where persuasion involves increasing the utility of purchasing.

\(^{15}\) Oliver and Anderson (1994) show that both approaches to salesforce control are used by companies and try to measure the consequences of these approaches.
References


Proof of Proposition 1

Notice that firms can assure themselves of profits equal to $\mu \theta (v - m) - w$ by charging the reservation price $v$. Since this is positive, no firm $i$ charges more than $v$ if $k_i = 0$ and no firm charges more than $v + k$. Doing so would lead to zero sales and profits of $-w$.

As demonstrated by Burdett and Judd (1983), the fact that some but not all customers observe two prices implies that the distribution of prices is smooth and has no holes. To see this, suppose first that the distribution has an atom at $p_1$. Then, a firm charging slightly less than $p_1$ makes strictly more profits because it increases the probability of selling by a discrete amount (namely by half the probability that the other firm charges exactly $p_1$). Now suppose that there is a hole in the price distribution between $p_1$ and $p_2$ with $p_1 < p_2$. Then, a firm can make strictly more profits than the firm at $p_1$ by charging a price above $p_1$, since it sells just as frequently but at a higher price.

Consider case a) where $k < w^\epsilon$. If a firm were to set $k_i = k$ in this case and the probability of selling at $p$ were equal to $q$, it expected profit per customer would equal $q(p - m) - w^\epsilon$. By setting $k_i = 0$ and setting the price at $p - k$, the firm would continue to sell with the same probability $q$, but its profit per customer would be higher. This implies that it is not desirable to set $k_i = k$.

The expected profit per customer is thus $\theta (v - m)$ for all firms so that the probability of selling at $p$, $q_0(p)$ must satisfy $q_0(p)(p - m) = \theta (v - m)$. Since $q_0(p)$ equals $(\theta + (1 - \theta)(1 - F_n(p)))$ where $F_n(p)$ is the cdf of prices charged by firms, (3) follows.

Since the price distribution has no mass points, any firm charging either the reservation price $v$ with $k_i = 0$ or the reservation price $v + k$ with $k_i = k$, sells only to a fraction $\theta$ of its potential customers. Of these two reservation price strategies, charging $v + k$ with $k_i = k$ is more profitable if

$$\theta (v + k) - w^\epsilon > \theta v \quad \text{or} \quad \theta k > w^\epsilon.$$  \hfill (46)

When this is satisfied, the highest price charged in equilibrium is $v + k$ (with $k_i = k$). The reason is that, if the highest price were lower, expected profits per customer (ignoring the
base wage, which acts as a fixed cost at this stage) at the highest price would be lower than \( \theta(v + k) - w^e \), which is achievable by charging \( v + k \). Since firms can assure themselves of this level of profits per customer, none has lower ones. Moreover, no firm has higher profits because, otherwise, a firm would prefer undercutting rather than charging \( v + k \).

I now show that, given that (46) implies that profits per customer equal \( \theta(v + k) - w^e \), no firm sets \( k_i = 0 \). For a firm to benefit from setting \( k_i = 0 \), there would have to exist a price \( p \) so that the probability of selling at this price \( q \) satisfies

\[
q(p - m) \geq \theta(v + k - m) - w^e.
\]

If such a combination of \( p \) and \( q \) existed, \( q \) would have to exceed \( \theta \) because this inequality is false for \( q = \theta \) even when \( p = v \). Even then, such a firm would then have the option of setting \( k_i = k \) and charging \( p + k \). It would then sell with the same probability \( q \) and its expected profit per customer (ignoring the base wage) would be \( q(p + k - m) - w^e \). This exceeds \( q(p - m) \), contradicting the desirability of setting \( k_i = 0 \).

Since the expected profit per customer is \( \theta(v + k - m) - w^e \) for all firms, the probability of selling at \( p \) for firms with \( k_i = k \), \( q_k(p) \), must satisfy \( q_k(p)(p - m) = \theta(v + k - m) \). At the same time, the probability of selling \( q_k(p) \) equals \( (\theta + (1 - \theta)(1 - F_a(p))) \) where \( F_a(p) \) is the cdf of prices charged by firms, and this implies (4).

Turn now to case b), where \( \theta k \leq w^e \leq k \). Since (46) is violated, all firms have an expected profit per customer equal to \( \theta(v - m) \). Denote by \( q_0(p) \) the probability that a firm setting \( k_i = 0 \) sells if it charges \( p \), while \( q_k(p) \) is the probability that a firm setting \( e_i = 1 \) sells at \( p \). The consumer consumer decision rule (2) implies that \( q_0(p) = q_k(p + k) \).

Note that the equations

\[
q_0(p^*)(p^* - m) = q_0(p^*)(p^* + k - m) - w^e = \theta(v - m).
\] \hspace{1cm} (47)

have the unique solution \( q^0(p^*) = w^e/k \) and \( p^* = m + \theta k (v - m)/w^e \). This establishes that the price \( p^* \) is the only one which ensures that both the expected profit per customer of firms with \( k_i = 0 \) that charge \( p^* \) and that of firms with \( k_i = k \) that charge the price \( p^* + k \) equal
\( \theta(v - m) \). Since \( q_0(p) \) is monotone in \( p \), firms with \( k_i = 0 \) make higher profits at \( p \) than do firms with \( k_i = k \) at price \( p + k \) if \( p^* < p \) while they make lower profits for \( p^* < p \). Thus, all firms with \( k_i = k \) charge prices lower than or equal to \( p^* + k \) while all firms with \( k_i = 0 \) charge prices greater than or equal to \( p^* \). Thus, \( q_0(p) = \theta + (1 - \theta)(1 - \gamma)(1 - F_0(p)) \) for prices between \( p^* \) and \( v \). Plugging this in (47) and using the fact that \( F_0(p^*) = 0 \) demonstrates that \( \gamma \) is indeed equal to \( (1 - w^e/k)/(1 - \theta) \). Using this the resulting expression for \( q_0(p) \) in (47), one obtains (5).

Similarly, for prices between \( w^e + m + \theta(v - m) \) (the price such that a firm setting \( k_i = k \) obtains an expected profit per customer of \( \theta(v - m) \) if it sells with probability one) and \( p^* + k \), \( q_k(p) = \theta + (1 - \theta)(1 - \gamma + \gamma(1 - F_k(p))) \). Using this in (47), one obtains (6).

**Proof of Proposition 2**

Since some firms set \( k_i = 0 \) and others set \( k_i = k \), we must be in case b) of proposition 1. Customers are indifferent between buying from a firm with \( k_i = 0 \) that charges \( p^* \) and buying from a firm that sets \( k_i = k \) and charges \( p^* + k \), so the expected sales of these firms are the same. Proposition 1 implies that firms with \( k_i = 0 \) that set a different price charge more so that their offering is more desirable and they sell less. Similarly, firms with \( k_i = k \) that do not charge \( p^* + k \) charge less, so they expect higher sales. It follows that firms that set \( k_i = k \) and charge \( p^* + k \) have a probability of selling equal to \( \theta + (1 - \theta)(1 - \gamma) \) and this equals \( w^e/k \).

**Proof of Proposition 3**

For \( w^e < \theta k \) the probability of selling \( q \) equals \( (\theta + (1 - \theta)(1 - F_a(p))) \) while it equals \( (\theta + (1 - \theta)(1 - F_n(p))) \) when \( w^e > k \). Since both \( F_a(p) \) and \( F_n(p) \) are distributed uniformly between 0 and 1, \( q \) is distributed uniformly between \( \theta \) and 1.

Now let \( \theta k \leq w^e \leq k \) and consider firms that set \( k_i = 0 \). Since their expected profits equal \( \mu \theta(v - m) - w \), their sales have to satisfy \( q(p)(p - m) = \theta(v - m) \). The probability that a random firm has a \( q \) less than or equal to \( \bar{q} \) must thus be equal to the probability that

43
its price exceeds \( m + \theta(v - m)/\bar{q} \). Using (5), this conditional probability must thus equal

\[
\text{Prob}(q \leq \bar{q}|k = 0) = 1 - \frac{(w^e/k)(\theta(v - m)/\bar{q}) - \theta(v - m)}{(w^e/k - \theta)(\theta(v - m)/\bar{q})} = \frac{\bar{q} - \theta}{w^e/k - \theta}
\]

for \( \bar{q} \) between \( \theta \) and \( w^e/k \). Since firms that set \( k_i = k \) sell more, the overall probability that \( q \leq \bar{q} \) when \( \bar{q} < w^e/k \) is thus \((1 - \gamma)\) times this probability. With \( \gamma = (1 - w^e/k)/(1 - \theta) \), this overall probability equals \((\bar{q} - \theta)/(1 - \theta)\).

Consider now the firms that that set \( k_i = k \). Their expected profits per customer must also equal \( \mu \theta(v - m) - w \) so \( q(p)(p - m) - w^e = \theta(v - m) \). For these firms, the conditional probability that \( q \) is less than or equal to \( \bar{q} \) is thus equal to the probability that \( p \) exceeds \( m + (\theta(v - m) + w^e)/\bar{q} \) or, using (6),

\[
\text{Prob}(q \leq \bar{q}|k_i = k) = 1 - \frac{(\theta(v - m) + w^e)/\bar{q} - (\theta(v - m) + w^e)}{(1 - w^e/k)(\theta(v - m) + w^e)/\bar{q}} = \frac{q - w^e/k}{1 - w^e/k}.
\]

Since firms with \( k = 0 \) have a \( q \leq w^e/k \), the overall probability that \( q \) is below \( \bar{q} \) when \( \bar{q} > w^e/k \) is equal to the sum of \( \gamma \) times the conditional probability that \( q < \bar{q} \) conditional on \( q \geq w^e/k \) and \((1 - \gamma)\). Using the value of \( \gamma \) above, this overall probability equals \((\bar{q} - \theta)/(1 - \theta)\) once again. This is the cdf for the uniform distribution between \( \theta \) and 1.

**Proof of Proposition 4**

Let \( Q(u) \) be equal to the cdf of drawing an ex post utility smaller than \( u \) from a visit to a single firm. When \( w^e \) is prohibitively high, this ex post utility is \( v - p \) where \( p \) is the price paid. Thus, \( Q(u) \) is the probability that \( p \) exceeds \( v - u \) and this equals

\[
Q(u) = 1 - F_n(v - u) = \frac{\theta u}{(1 - \theta)(v - u - m)}, \quad (48)
\]

where \( F_n \) is given by (3).

Turn now to case b) where \( w^e \) is low enough that all salespeople incur this cost. Suppose first that \( g_i = k_i \). The cdf \( Q_k(u) \) is now the probability that \( p \) exceeds \( v + k - u \), or \( 1 - F_a(v + k - u) \) where \( F_a \) is given by (4). Using this equation, we have

\[
Q_k(u) = \frac{\theta u}{(1 - \theta)(v + k - u - m)}. \quad (49)
\]
This expression is smaller than that in (48) when \( k \) is positive. This means that the distribution of utilities from a visit to a single salesperson first order stochastic dominates the distribution when \( w^e \) is prohibitive, so consumers are better off. It follows that, when salespeople raise utility by \( k \), the maximum utility obtained from two different salespeople also dominates the maximum utility from visiting two salespeople when all firms set \( k_i = 0 \). Thus, all consumers can be expected to be better off when all firms set \( g_i = k \).

Now suppose that \( \ell_i = k_i \) or that one can ignore the effect of persuasion on welfare and focus first on the case where \( w^e \) is low enough that all salespeople raise the willingness to pay by \( k \). The cdf \( Q(u) \) is the probability that \( p \) is greater than \( v - u \) with \( F \) still being given by (4). This now equals

$$Q_{\ell}(u) = \frac{\theta(u + k)}{(1 - \theta)(v - u - m)}.$$  \hspace{1cm} (50)

This is larger than the expression in (48) so the case with infinite \( w^e \) leads to a \( u \) that first order stochastically dominates this one. This means that a customer that sees a single salesperson can expect to be worse off. It also implies that customers that visit two salespeople are worse off because individuals get two draws from a less favorable \( u \) distribution. If, in addition, \( \ell_i = k_i \), individuals who visit two salespeople lose \( k \) for sure.

I now study case c), where there is a probability \( \gamma \) that firms set \( k_i = k \). To compare the resulting consumer welfare with welfare when \( w^e \) is prohibitively high, it is useful to decompose the distribution of \( u \) from a single visit in the case of prohibitively high \( w^e \). In particular, let \( Q^-(u) \) be the cdf of \( u \) conditional on \( u \) being smaller than \( v - p^* \) (the level of \( u \) when the price is \( p^* \)) while \( Q^+(u) \) is the cdf of \( u \) conditional on \( u \) being larger than \( v - p^* \). Using Bayes’ rule

\[
Q^-(u) = \frac{Q(u)}{Q(v - p^*)} \quad \text{and} \quad Q^+(u) = \frac{Q(u) - Q(v - p^*)}{1 - Q(v - p^*)}.
\]

In the case where \( w^e \) is prohibitive, \( Q(u) = 1 - F_n(v - u) \) with \( F_n \) given by (3). Therefore,

\[
Q^-(u) = \frac{\theta u}{(w^e/k - \theta)(v - u - m)} \quad \text{and} \quad Q^+(u) = \frac{\theta(v - m) - (v - u - m)w^e/k}{(1 - w^e/k)(v - u - m)};
\]
while
\[ Q(v - p^*) = 1 - F_n(p^*) = \frac{w^e/k - \theta}{1 - \theta} = 1 - \gamma, \]  
(51)

where the third equality follows from Proposition 1. When \( \theta k \leq w^e \leq k \), there is a probability \( (1 - \gamma) \) that employees set \( e_i = 0 \) so the resulting ex post utility also equals \( v - p \). The cdf of \( u \) conditional on encountering a salesperson who sets \( e_i = 0 \), \( Q^0(u) \) is thus equal to \( 1 - F_0(v - u) \). Using (5), this implies that \( Q^0(u) = Q^-(u) \).

Therefore, the distribution of utility in case c) conditional on finding a salesperson who has set \( k_i = 0 \) is the same as the distribution of utility when \( w^e \) is prohibitive and the price is above \( p^* \). Moreover, (51) implies that the probability in case c) of finding a salesperson who has set \( k_i = k \) is the same as that of finding a price above \( p^* \) when \( w^e \) is prohibitive.

Thus, the question of whether the utility distribution in case c) dominates the distribution of utility conditional on finding a salesperson who sets \( k_i = k \) dominates the distribution of utility conditional on finding a price below \( p^* \) in the case where \( w^e \) is prohibitive.

When \( g_i = k_i \), a consumer’s utility when purchasing from a salesperson who has set \( k_i = k \) is \( v + k - p \). Therefore, the conditional probability of obtaining a level of utility below \( u \) from a single encounter with a salesperson, \( Q^1_g(u) \), equals \( 1 - F_k(v + k - u) \) or, using (6)
\[ Q^1_g(u) = 1 - \frac{v + k - u - m - \theta(v - m) - w^e}{(1 - w^e/k)(v + k - u - m)} = \frac{\theta(v - m) - (v - u - m)w^e/k}{(1 - w^e/k)(v + k - u - m)}. \]

For \( k > 0 \), this is smaller than \( Q^+(u) \). This implies that the distribution of utility when some salespeople set \( g_i = k \) first order stochastically dominates the utility when \( w^e \) is prohibitive.

When \( \ell_i = k_i \) or when persuasion has no effect on “true” utility, the utility from finding a single salesperson who has set \( k_i = k \) remains \( v - p \) so that the conditional probability of obtaining a level of utility below \( u \), \( Q^1_{\ell}(u) \), equals \( F_k(v - u) \) or, using (6)
\[ Q^1_{\ell}(u) = 1 - \frac{v - u - m - \theta(v - m) - w^e}{(1 - w^e/k)(v - u - m)} = \frac{\theta(v - m) + w^e - (v - u - m)w^e/k}{(1 - w^e/k)(v - u - m)}. \]

For \( w^e > 0 \), this is larger than \( Q^+(u) \). Thus the distribution of utility when some salespeople set \( \ell_i = k \) or when the direct effect of persuasion on utility can be ignored is stochastically
dominated by the distribution that obtains when \( w^e \) is prohibitive. Consumers are therefore made worse off by the higher prices that accrue when salespeople engage in persuasive effort. In addition, when salespeople incur the requisite effort to raise \( \ell \), consumers who visit two salespeople who both have \( \ell_i = k \) suffer the direct loss of \( k \).

**Proof of Proposition 6**

Notice first that, when \( m = 0 \), Proposition 1 implies that revenue \( [\theta + (1 - \theta)(1 - \gamma)(1 - F_0(p))]p \) must equal \( \theta v \) while revenue \( [1 - (1 - \theta)\gamma F_k(p)]p \) must equal \( w^e + \theta v \). It further follows from (5) that, when \( m = 0 \),

\[
dF_0(p) = \frac{\theta v}{(w^e/k - \theta)p^2}
\]

so that

\[
\int_{p^*}^{p} ydF_0(y) = \int_{v\theta k/w^e}^{p} \frac{\theta v}{(w^e/k - \theta)y} dy = \frac{\theta v}{(w^e/k - \theta)} \log \left( \frac{pw^e}{v\theta k} \right)
\]

Similarly (6) implies that

\[
dF_k(p) = \frac{w^e + \theta v}{(1 - w^e/k)p^2}
\]

so that

\[
\int_{p^-}^{p} ydF_k(y) = \int_{w^e + \theta v}^{p} \frac{(w^e + \theta v)}{(1 - w^e/k)y} dy = \frac{w^e + \theta v}{(1 - w^e/k)} \log \left( \frac{p}{w^e + \theta v} \right)
\]

Substituting these expressions in (11) and (12) and simplifying, (14) and (15) follow. By evaluating these equations for \( u_0(p) \) and \( u_1(p) \) at prices of \( v \), \( p^* \) and \( p^- \), one obtains the conditions (16) and (17).

**Proof of Proposition 8**

Suppose we are in a steady state. At any point in time \( t \), the total number of salespeople of type \( \lambda \) whose jobs give customers utility greater than or equal to \( \hat{u} \) equals \( \eta(1 - G(\hat{u}))N \). Of these, a fraction \( (1 - \psi) \) keep their period \( t \) jobs at \( t + 1 \). In addition \( \psi a + (1 - \psi)\eta G(\hat{u}) \) individuals of type \( \lambda \) search for jobs at the beginning of \( t + 1 \), where \( \psi aN \) are new recruits and the rest are individuals who are dissatisfied with their period \( t \) job. These searchers have a probability \( 1 - H(\hat{u}) \) of finding a job at \( t + 1 \) that provides utility greater than or
equal to \( \hat{u} \). Thus, the total number of people of type \( \lambda \) that have such jobs at \( t + 1 \) is

\[
\left\{ (1 - \psi)(1 - G(\hat{u}))\eta + \left[ \psi a + (1 - \psi)\hat{\eta}G(\hat{u}) \right] \frac{\psi(1 - G(\hat{u}))}{\psi + (1 - \psi)\hat{\eta}G(\hat{u})} \right\} N
\]

Equating this to \( \eta(1 - G(\hat{u}))N \) and simplifying gives equation (25) as long as \( \psi \neq 0 \). For \( 0 < \psi < 1 \), the left hand side of this equation equals plus infinity for \( \eta = \infty \) while it equals \( \psi(a - 1) < 0 \) for \( \eta = 1 \). This implies that the equation has one real zero smaller than one and another real zero that is greater than one. For values of \( \eta \) between these zeros, the left hand side of (25) must be negative. At \( \eta = a \), this expression equals

\[
a^2(1 - \psi)(1 - G(\hat{u})) - a[1 + (1 - \psi)(a - G(\hat{u}))] + a = a(1 - a)(1 - \psi)G(\hat{u})
\]

Since this is positive and \( a < 1 \), the smallest zero of (25) is larger than \( a \). Equation (23) then implies that \( \eta \) exceeds \( \hat{\eta} \).

**Proof of Proposition 9**

When \( \ell_i = k_i \) and (16) is satisfied as well as when effort has no effect on true welfare and (7) is satisfied, \( u_1(p^-) < u_0(v), G(u_1(p^-)) = G(u_0(v)) = \gamma \). Changes in \( \hat{u} \) within the range \( [u_1(p^-), u_0(v)] \) thus have no effect on \( \eta \), implying that \( H(\hat{u}) \) is also constant in this range. Equation (24) implies that \( \lim_{\psi \to 0} H(\hat{u}) = 1 \). One can therefore find both a sufficiently small \( \psi \) such that the mean of \( H \) is below \( \hat{u} \) for any \( \hat{u} \) between \( u_1(p^-) \) and \( u_0(v) \) and a \( \rho \) less than 1 that satisfies (22). Since \( \psi \) is small, (20) implies that the discount rate \( \hat{\rho} \) that accomplishes is less than one as well.

As \( \hat{u} \) moves within the range \( [u_1(p^-), u_0(v)] \), the set of values for \( \psi \) that ensure that \( \hat{u} \) exceeds the mean of \( H \) varies as well. For a given \( \psi \), varying \( \hat{u} \) within this range also changes the value of \( \rho \) that leads \( \hat{u} \) to be the cutoff level of utility. All values of \( \hat{u} \) in the range \( [u_1(p^-), u_0(v)] \) lead to the same behavior, namely the departure of employees of type \( \lambda \) from jobs with \( e = 1 \). It follows that there is a nontrivial set of values of \( \psi \) and \( \rho \) that rationalize these actions. Moreover, these actions ensure that a fraction \( \eta \) of the employees that set \( e = 0 \) (and thus have output levels below \( c/k \)) are of type \( \lambda \). Among salespeople with higher output, the fraction is only \( \hat{\eta} \), which is lower.
Proof of Proposition 10

Given that $u_1(p^* + k) < u_0(p^*) < u_1(p^*)$, there does exist a $\hat{p}$ greater than $p^-$ and smaller than $p^* + k$ such that $u_1(\hat{p}) = u_0(p^*)$. With this being equal to $\hat{u}$ salespeople of type $\lambda$ leave all jobs with $e = 0$ and stay only at those jobs with $e = 1$ that have prices below $\hat{p}$. These firms all have expected sales greater than or equal to $1 - \gamma F_1(\hat{p}) > w^e/k$. As in the proof of Proposition 9, one can always find a low enough $\psi$ so that this choice of $\hat{u}$ is above the mean of $H$. At a low $\psi$, the $\hat{\rho}$ that satisfies (20) and (22) is smaller than one, and this discount rate ensures that salespeople do indeed set the $\hat{u} = u_0(p^*)$.

Proof of Proposition 11

Since (32) implies that $\alpha k > w^e$ and (31) implies that $w^e > \alpha(1 - \theta)(v - m)$, $\alpha k > \alpha(1 - \theta)(v - m)$, which implies (28). Moreover, since the term in square brackets in (32) exceeds one, the equation implies that $\alpha k - w^e$ exceeds $\theta(v - m)$, which obviously exceeds $\theta(1 - \alpha)(v - m)$ so that (30) holds.

Now notice that $F_0(p)$ in (35) is indeed a proper cdf. It is increasing in $p$, and takes the value of zero for $p = p^*$ and the value of one for $p = v$. I first show that, if this is the distribution of prices for firms that set $e = 0$, they all make the same profits.

From (34), we have

$$\alpha(p^-_1 - m) = [\theta + (1 - \theta)(1 - \alpha)\gamma](v - m) + w^e > \alpha(v - m)$$

where the inequality uses the violation of (31), the fact that $\theta > \theta\alpha$ and the fact that $(1 - \theta)(1 - \alpha)\gamma > 0$. This implies that $p^-_1 > v$ so that, at the proposed equilibrium, no immune consumer buys from a firm that sets $e = 1$. Firms whose price is no greater than $v$ therefore face a fraction $\theta + (1 - \theta)(1 - \alpha)\gamma$ of consumers who are insensitive to price. Their expected sales per customer at a price $p$, $q_0(p)$, are

$$q_0(p) = \theta + (1 - \theta)[(1 - \gamma)(1 - F_0(p)) + \gamma(1 - \alpha F_1(p + k))] = 1 - (1 - \theta)[(1 - \gamma)F_0(p) + \gamma \alpha F_1(p + k)]$$

For prices between $p^*$ and $v$, $F_1(p + k) = 1$ and, using (35), profits per customer equal

$$q_0(p)(p - m) = [\theta + (1 - \theta)(1 - \alpha)\gamma](v - m)$$
so that firms that set \( e = 0 \) are indifferent among all these prices.

I now turn to firms that set \( e = 1 \). Note first that \( F_1(p) \) in (36) is increasing in \( p \) because \( p_1^- > m \). It obviously equals 0 for \( p = p_1^- \). Moreover (32) together with (33) and (34) implies that \( F_1(p^* + k) = 1 \). Thus, \( F_1(p) \) is a proper cdf between \( p_1^- \) and \( p^* + k \).

For any price \( p > v \), firms that set \( e = 1 \) have expected sales per customer, \( q_1(p) \) equal to

\[
q_1(p) = \alpha\{\theta + (1-\theta)[(1-\gamma)(1-F_0(p-k))+\gamma(1-F_1(p))]\} = \alpha\{1-(1-\theta)[(1-\gamma)F_0(p-k)+\gamma F_1(p)]\}.
\]

For prices below \( p^* + k \), \( F_0(p - k) = 0 \) so that, using (36), expected profits per customer equal

\[
q_1(p)(p - m) - \omega^e = \left[\theta + (1-\theta)(1-\alpha)\gamma\right](v - m).
\]

Therefore, firms that set \( e = 1 \) are indifferent among these prices and also have nothing to gain or lose by switching to \( e = 0 \) with a price between \( p^* \) and \( v \).

All that is left to prove is that charging other prices for either level of \( e \) is strictly less profitable. Consider first deviations involving \( e = 1 \). Charging prices strictly between \( v \) and \( p_1^- \) gives lower profits because it attracts no sales beyond those available at \( p_1^- \) while charging strictly less. Charging the price \( v \) gives profits per customer of

\[
(1-\alpha)[\theta + (1-\theta)(1-\alpha)\gamma](v - m) + \alpha(v - m) - \omega^e.
\]

Condition (31) implies that this is less than the equilibrium level of profits. Now consider prices above \( p^* + k \). A firm charging such a price makes no sales to susceptible individuals who encounter another firm with \( e = 1 \). Its probability of selling is

\[
q_1(p) = \alpha\{1-(1-\theta)[\gamma + (1-\gamma)F_0(p-k)]\} = \alpha q_0(p - k) - (1-\theta)(1-\alpha)\gamma,
\]

where the second equality is based on the fact that, for \( p \geq p^* \), \( q_0(p) \) equals \( 1-(1-\theta)(\gamma \alpha + (1-\gamma)F_0(p)) \). Therefore, profits at such a deviation equal

\[
q_1(p)(p - m) = \{\alpha q_0(p - k)(p - k - m)\} + \alpha q_0(p - k)k - (1-\theta)(1-\alpha)\gamma(p - m).
\]
The term in curly brackets is independent of \( p \) while both other terms decline when \( p \) rises. Thus profits decline as a firm with \( e = 1 \) raises its price above \( p^* + k \).

Now consider deviations involving \( e = 0 \). Raising price above \( v \) eliminates all sales and is thus unprofitable. A firm with \( e = 0 \) and a price below \( p^* \) sells to every immune customer it encounters and to all customers whose other salesperson sets \( e = 0 \). It also sells to susceptible customers whose offer from a firm with \( e = 1 \) has a price that exceeds \( p + k \). Thus, expected sales for such a firm are

\[
q_0(p) = 1 - (1 - \theta)\gamma \alpha F_1(p + k) = q_1(p + k) + 1 - \alpha,
\]

where the second equality stems from the fact that, for prices below \( p^* + k \), \( q_1(p) = \alpha[1 - (1 - \theta)\gamma \alpha F_1(p + k)] \). Its profits are therefore

\[
q_0(p)(p - m) = \{q_1(p + k)(p + k - m)\} - q_1(p + k)k + (1 - \alpha)(p - m).
\]

The term in curly brackets is independent of \( p \) while the others rise with \( p \). Therefore, profits decline as the price is reduced below \( p^* \).

**Proof of Proposition 12**

When \( \gamma > 0 \), equation (37) implies that (31) holds, so that no equilibrium where all firms set \( e = 0 \) would exist if (28) held. Moreover, using (41) to substitute for \( p^* \) in the first inequality of (40) and then using (37) to substitute for \( (v - m) \) leads to the conclusion that \( \alpha k > w^e \). An equilibrium where all firms set \( e = 0 \) would therefore also fail to exist if (28) were violated.

Equations (44) and (42) are the solutions to the following two equations

\[
q_0(p)(p - m) = \theta(v - m) \quad q_1(p)(p - m) = \theta(v - m) + w^e
\]

where

\[
q_0(p) = \theta + (1 - \theta)[(1 - \gamma)(1 - F_0(p)) + \gamma(1 - \alpha)(1 - F_1(p))]
\]

\[
q_1(p) = \theta + (1 - \theta)[(1 - \gamma)(\alpha + (1 - \alpha)(1 - F_0(p)) + \gamma(1 - F_1(p))].
\]
The first inequality in (40) requires that \( p^* + k > v \). As a result, susceptible customers do not buy from salespeople that set \( e = 0 \) if they also meet a salesperson with \( e = 1 \) that charges \( v \) or less. This implies that, if the \( F_i(p) \)'s are the cdf’s of equilibrium prices, the \( q_i(p) \) above measure the probability of selling for a firm that charges \( p \) and sets \( e = i \). This further implies that these \( F \)'s lead expected profits per customer to equal \( \theta(v - m) \) for all firms that charge between \( p^- \) and \( v \).

Equation (37) implies that \( F_0(v) = F_1(v) = 1 \) solves these equations. Moreover, the definition of \( p^- \) in (45) implies that \( F_1(p^-) = 0 \). Since (42) is increasing in \( p \), \( F_1(p) \) is thus a proper cdf for a distribution taking values between \( p^- \) and \( v \).

\( F_1(p) \) continues to equal zero for prices below \( p^- \). Moreover, (39) ensures that \( p^* < p^- \). Thus \( F_0(p) \) in (43) implies that \( q_0(p)(p - m) \) is equal \( \theta(v - m) \) for prices between \( p^- \) and \( p^* \). It remains to show that \( F_0(p) \) is a proper cdf. We saw that it equals one for \( p = v \). In addition, the definition of \( p^* \) ensures that \( F_0(p^*) = 0 \) in (43). Moreover, (43) rises with \( p \). For (44) not to fall with \( p \), we must have \( \alpha \theta(v - m) \geq (1 - \alpha)w^e \). Using (37) to substitute for \( w^e \), it is readily verified that (38) ensures that this is indeed the case.

We have thus shown that the proposed \( F_i(p) \) are cdf’s that lead firms to be indifferent among all the prices in their ranges. It remains to show that firms do not benefit by deviating to different prices. For a firm that sets \( e = 0 \), raising its price above \( v \) eliminates all sales. Prices between \( (v - k) \) and \( p^* \) yield the same expected sales than the price \( p^* \) and are therefore less profitable. For any price \( p \) below \( v - k \), the expected probability of selling \( q_o(p) \) equals \([1 - \alpha \gamma (1 - \theta)F_1(p + k)]\). Using the second of the two equations above to substitute for \( F_1(p + k) \), expected revenue per customer \( q_o(p)(p - m) \) at these prices equals

\[
(1 - \alpha)[1 + \alpha(1 - \theta)(1 - \gamma)F_0(p + k)](p - m) + \alpha\{q_1(p + k)(p + k - m)\} - \alpha q_1(p + k)k.
\]

Since the term in curly brackets is independent of \( p \), this expression is increasing in \( p \). This means that cutting price below \( v - k \) is unprofitable as well.

Now consider the firms that set \( e = 1 \). If they were to raise the price above \( v \), the earlier argument implies that their most profitable deviation is to set it equal to \( p^* + k \). To prevent
this deviation it must be the case that
\[ \alpha(\theta + (1 - \theta)(1 - \gamma))(p^* - m + k) < \theta(v - m) + w^e, \]
where the term multiplying \((p - m)\) is the probability of sale. Using (37) to substitute for \(w^e\), it can be verified that the second inequality in (40) ensures that this is true. If firms that set \(e = 1\) lower their prices below \(p^-\), the linear equations above imply that their revenues per customer fail to fall only if \(F_1\) becomes negative, which is impossible.
Figure 1: Utility and output levels in “high-\(w^e\)” stochastic model
Figure 2: Relative utilities and relative outputs in “low-$w$” stochastic model.