



## Marketing Science

Publication details, including instructions for authors and subscription information:  
<http://pubsonline.informs.org>

### In the Eye of the Beholder: An Analysis of the Relative Value of a Top Sales Rep Across Firms and Products

David Godes,

To cite this article:

David Godes, (2003) In the Eye of the Beholder: An Analysis of the Relative Value of a Top Sales Rep Across Firms and Products. Marketing Science 22(2):161-187. <https://doi.org/10.1287/mksc.22.2.161.16040>

Full terms and conditions of use: <https://pubsonline.informs.org/Publications/Librarians-Portal/PubsOnLine-Terms-and-Conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact [permissions@informs.org](mailto:permissions@informs.org).

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

© 2003 INFORMS

Please scroll down for article—it is on subsequent pages



With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes.

For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

# In the Eye of the Beholder: An Analysis of the Relative Value of a Top Sales Rep Across Firms and Products

David Godes

Graduate School of Business Administration, Harvard University, Morgan Hall 165,  
Soldiers Field, Boston, Massachusetts 02163  
dgodes@hbs.edu

---

We ask the question, “when should the most highly skilled salespeople sell the best products?” Our main result is that the highly skilled reps should sell better products when the task is very complex and worse products when the task is very simple. This is shown using a general analytical model of selling in which sales are a joint function of the salesperson’s skill and the complexity of the selling task. Complexity varies across products and industries. Intuitively, when the selling task is complex, few salespeople of *any* level of ability will be successful with a low-quality product. Therefore, the high-skill rep’s value is higher on the better product. Conversely, when the task is simple, salespeople of *any* ability can sell the better product fairly easily so the high-skill rep’s impact is more pronounced on the worse product. This general result offers insight into many key problems: Which salespeople should we hire? How should we organize our salespeople? How should we allocate training funds? We show that the insights hold for salespeople that either *create* value or simply *persuade* the customer about the product’s value. Finally, we contrast this set of questions with the question of *how many* salespeople the firm should hire. We find that the firm that has the biggest sales force does not always have the best.

(*Sales Management; Selling; Hiring Policies; Game Theory; Agency Theory*)

---

## 1. Introduction

No aspect of marketing receives more investment than the sales force: U.S. firms spend \$140 billion a year on selling (Kotler 2000). This translates into about 7% of gross revenues on the sales force (Marchetti 1997). In fact, over 11 million Americans are employed in sales-related jobs (Kotler 2000). The effective management of this function is thus essential to the success of many firms and business units. Managing the sales force, however, is a complex and multidimensional problem. Sales managers must decide how to hire, organize, train, and compensate their salespeople. While most of the recent research in

the area addresses the latter issue, compensation, this paper offers insight into each of these other important areas by analyzing the *value* of a highly skilled sales rep and how this value differs across firms and products. Specifically, do top-level reps have a bigger impact on better products or on worse? We show that the answer to this question has important implications for how the firm hires, organizes, and trains the sales force. We address the firm’s *hiring* problem by asking which firms should pay “what it takes” to attract the most highly skilled salespeople. Evidence for the importance of this question lies in the staggering difference in costs. On average, a top

sales rep costs about 50% more than a midlevel rep. In retail, the premium for a top rep is more than 200% (Marchetti 1997)! The decision by a firm to hire a highly skilled rep, then, depends on whether or not her incremental production will compensate the firm for *up to triple* the cost. Which firms or products should pay this premium? A first guess might be that firms with better products (those that are easier to sell) will experience the bigger gain and therefore should hire the high-skill reps. The market for their product may be bigger and, therefore, the firm would experience the impact of the rep's selling ability on a broader scale. Casual observation suggests that this may not always be the case. In its July 2000 issue, *Sales & Marketing Management* named its annual "Best Salesforces in America." Among the list of 25 were several companies that are well known as providers of superior products: Northwestern Mutual Insurance, Cisco Systems, and Dell Computer, for example. The fact that they also employ the "best" sales forces seems to offer evidence that selling skill is more valuable to firms with superior products. However, there are also companies on this prestigious list that do not seem to offer superior products. Viacom's CBS television network is on the list, while rivals NBC and ABC are not, suggesting that CBS's sales force is better than NBC's and ABC's. However, by just about any measure, it is clear that NBC and ABC both offered better products than CBS at the time.<sup>1</sup> Similarly, Charles Schwab & Co. was the only broker named to the list, suggesting that, at least according to *Sales & Marketing Management*, they have the industry's best sales force. Again, as compared with the likes of TD Waterhouse and Merrill Lynch, Schwab's offerings are unlikely to be rated as the best in the business.<sup>2</sup> Thus, it seems that the naïve model that "the best always hire the best" is too simplistic. In this paper, we present a model that offers insight into why and when relatively inferior firms or products should

optimally employ the best salespeople and when they should not. The paper also addresses the firm's *organization* of its sales force. Specifically, a multiproduct firm must decide on which products to place its most highly skilled sales reps. Should they always sell the firm's best products, or would they have a bigger impact on products that are more difficult to sell? Likewise, we are interested in how the firm should allocate its *training* budget. Should it always invest more in improving the skills of those selling its best products?

While these questions are central to our understanding of sales force management, there exists no general theory to address them. Montgomery et al. (1971) is closest in spirit to the problem we address here. They employ a decision calculus methodology (Little 1970) to assist managers in allocating salespeople's time across products. However, these authors are concerned with interproduct response differences and not intersalesperson differences. We are concerned with the interaction between both of these factors. Also related to our work is the stream of research beginning with Zoltners and Sinha (1980). They apply numerical optimization techniques to the question of *how many* salespeople to hire and how to organize them. Finally, Lodish (1971) employs similar methods to the organization of sales territories and prioritization of selling efforts across customers.

We take a two-stage approach to answering these questions. First, in §2, we present a micromodel of the selling process. The objective of this model is to understand the *impact* of selling ability: which firms or products will experience a higher increase in sales as a result of an increase in the skill of its salespeople? In equilibrium, of course, relative impact should imply relative value, and we therefore interpret the results of the model as informing us about which firms or products optimally place the highest *value* on premium selling ability. We specify the firm's sales as a joint function of (a) the salesperson's skill and (b) the "complexity" of the selling task, or how hard the product is to sell. The key result from the model is that selling skill has a higher impact on the *easier* of two products (which we will call the "better" product) when the industry is complex and on the

<sup>1</sup> For example, CBS's rating/share for adults 18–49 in the 11/99 sweeps was 3.8/10 compared with ABC's 6.1/16 and NBC's 5.4/14. In 1998, the results were similar: CBS delivered 4.0/11, ABC 4.9/13, and NBC 6.0/16 (Schlosser 1999).

<sup>2</sup> *SmartMoney* ranked TD Waterhouse and Merrill Lynch the two best brokers in the three categories they surveyed.

*harder* of two products (the “worse” product) when the industry is simple. Intuitively, when the context is complex, neither a highly skilled rep nor a lesser-skilled rep can sell the difficult product, so selling skill has a bigger impact on the easier product. On the other hand, when the context is very simple, *anyone* can sell an easy product, so there is a bigger impact of selling skill on the harder product.

In the sections that follow, we use this model as the basis for a series of inquiries into a set of specific and practical questions facing sales managers: hiring, training, and organizing. The micromodel is essential to our analysis of these questions because each of them is fundamentally about the valuation of selling ability. For example, some firms value selling ability more than others and therefore in equilibrium will *hire* better salespeople more often. Likewise, selling skill is more valuable on some products than others and therefore *ceteris paribus* the firm should ensure that the best salespeople are selling these products. Thus, they should allocate their training funds and organize their salespeople across products in accordance with this objective. In §3, we look at a single-product monopolist employing salespeople who are responsible for “adding value” to the product they are selling. We show that the general intuition developed in §2 helps us to answer a wide variety of important questions. Specifically, we find that in complex industries the firm should hire the top-level rep only if its product is good. It should do so in a simple industry only if its product is not. We also find that a two-product monopolist should put its best reps on its better product in a complex industry, but on its worse product in a simple industry. Similarly, the two-product monopolist should invest more in training the better product’s sales staff in a complex industry and the worse product’s staff in a simple industry. In §4, we generalize the hiring problem by considering a setting in which competing firms bid for salespeople. Here, we begin by showing that our monopoly findings hold in a competitive context in that the better firm optimally hires away (“poaches”) the inferior firm’s salesperson only when the industry is complex, but the opposite is true when the industry is simple. We then introduce the potential for a mismatch between the

supply of, and demand for, salespeople. In this setting, the firms will have sales forces of unequal size. We are therefore able to investigate a related question, which firm should hire the *most* salespeople, along with our ongoing analysis of which firm should hire the *best*. When the product is simple, the answers to these two questions are not necessarily the same. In some cases, the better firm hires the most salespeople but the worse firm hires the most highly skilled salesperson. The intuition for this is that in a simple selling context, any salesperson of any ability will bring significant sales as compared with an empty slot. This is particularly true of the better product. However, when comparing two nonempty slots, the same forces at work throughout the rest of the paper take over, and the worse firm has a higher *marginal* return from a high-skill rep. Thus, the absolute effect of hiring *a* salesperson may be higher in this case for the better firm, but the marginal effect of hiring a *more highly skilled* salesperson is higher for the worse firm. Finally, in §5, we change our definition of “selling” and assume instead that the salesperson may only *persuade* the customer of the product’s value. Besides yielding similar intuition to the other models, and therefore establishing the robustness of the results, the model also offers interesting insight into why rational customers may believe salespeople, even though they are biased. Section 6 concludes with suggestions for future research. To aid the reader, Appendix B contains a summary of the notation used throughout the paper.

## 2. A Model of Selling

Our model is closely related to Weitz’s (1981) “adaptive selling” model, which has five stages: The salesperson (1) forms an impression of the customer, (2) formulates a strategy to influence choice, (3) transmits information to implement this strategy, (4) evaluates the strategy, and (5) adjusts her approach as necessary. We depart from Weitz’s model at several points. First, while Weitz allows for an independent level of selling skill in each stage, we collapse Weitz’s Stages (1)–(3) and focus on the salesperson’s ability to choose an “effective action.” Second, our model does not include feedback and adjustment (Stages (4)

and (5)), though the model could easily be extended to do so. Finally, our model of selling is somewhat more general in that the rep is assumed to do more than transmit information. Her choice of an effective action may also include, for example, adding value to the product. We assume that there are  $K$  possible “actions” that the rep may select to influence the customer’s choice. Each action  $k$  is either “effective” or “ineffective” for a given customer:

$$a_k = \begin{cases} 0 & \text{Action } k \text{ is ineffective,} \\ 1 & \text{Action } k \text{ is effective.} \end{cases} \quad (2.1)$$

The meaning of effectiveness will depend on the task. If the task is value creation, effectiveness implies the selection of a valuable product enhancement. If the task is persuasion, effectiveness implies convincing the customer that the underlying product has a high value. The salesperson selects one of the  $K$  actions that will be “effective” for a given customer. This restriction that the salesperson may only select one action is without loss of generality because we can simply consider combinations of  $N$  subactions to be actions. We will denote the salesperson’s choice as  $k^*$ . A given selling context will be characterized by a parameter  $\delta$  which will capture the difficulty or “complexity” of the task:

$$\delta = \frac{1}{K} \sum_{k=1}^K (1 - a_k). \quad (2.2)$$

So,  $\delta$  is the proportion of actions that are ineffective. The salesperson thus has equal prior beliefs about each action:  $\Pr[a_k = 1] = 1 - \delta, \forall k$ .<sup>3</sup>

The salesperson chooses an action after forming an impression of the customer. We model this impression-formation process as the reception by the salesperson of a  $K$ -dimensional vector of binary signals  $\hat{a}$  where each scalar element  $\hat{a}_k$  is given by:

$$\hat{a}_k = \begin{cases} 1 & \text{Action } k \text{ appears effective,} \\ 0 & \text{Action } k \text{ appears ineffective.} \end{cases} \quad (2.3)$$

Equation (2.3) captures the idea that the salesperson gets a noisy signal about the possible effectiveness of

each action. We model selling skill  $\sigma$  as the accuracy of this signal:<sup>4</sup>

$$\sigma \equiv \Pr[\hat{a}_k = a_k] \quad \forall k, \quad \sigma > \frac{1}{2}. \quad (2.4)$$

The restriction that  $\sigma$  is greater than  $\frac{1}{2}$  avoids the awkward implication that the salesperson should do the opposite of what her signal suggests. We interpret  $\sigma$  as the probability that the salesperson reads the customer correctly. The salesperson inspects the vector  $\hat{a}$  and randomly chooses from among those actions on which she received a positive signal  $\hat{a}_k = 1$ . This randomization is optimal given her equal prior beliefs over actions. With nonzero probability, she may receive a positive signal on none of the potential actions. For example, her initial impression of a customer may not be sufficiently precise to allow her to select a pitch. We will assume that in this case she repeats the process until she receives at least one positive signal. Therefore, she might spend more time to form a second impression by asking more questions, etc. Combining our definitions of complexity  $\delta$  and selling skill  $\sigma$  with this assumed selling process, we can calculate the probability that an effective action is chosen.

**LEMMA 1.** *The probability that a salesperson of skill  $\sigma$  selects an effective action for a customer of complexity  $\delta$  is given by:*

$$p(\sigma, \delta) \equiv \Pr[a_{k^*} = 1 \mid \sigma, \delta] = \left[ 1 + \frac{1 - \sigma}{\sigma} \cdot \frac{\delta}{1 - \delta} \right]^{-1}. \quad (2.5)$$

Further,  $\partial p / \partial \delta < 0$ ,  $\partial p / \partial \sigma > 0$ ,  $\partial^2 p / \partial \delta^2 < 0$ ,  $\partial^2 p / \partial \sigma^2 < 0$  for  $\delta < (>) \frac{1}{2}$ , and  $\partial^2 p / \partial \sigma \partial \delta < (>) 0$  for  $\delta > (<) \sigma$ .

All proofs other than those noted are contained in Appendix A. The key insight in Lemma 1 is contained in the cross-partial result: When the selling task is complex enough, the impact of increasing selling skill is decreasing in complexity. So, assume that we have two customers that differ only in the sense that one is slightly more difficult to sell than the other. Further, we have two salespeople that differ in selling

<sup>3</sup> One way to think about this is that each customer is randomly “assigned” an  $a_k$  for each  $k$  such that  $\Pr[a_k = 1] = 1 - \delta$ .

<sup>4</sup> This implies that the probability of Type I error and Type II error are equal. It is easy to show that relaxing this restriction and allowing for two parameters instead of one does not qualitatively change the results.

ability by a slight amount. The more highly skilled of the two will always have a higher success probability, but we are interested in how the difference compares across customers. Lemma 1 says that when the customers are relatively complex (the  $\delta$ 's are both greater than the  $\sigma$ 's), then the difference in success probabilities experienced by the two salespeople will be higher on the simpler of the two customers. On the other hand, if the  $\delta$ 's are less than the  $\sigma$ 's—indicating a relatively simple selling context—then the difference in the success probabilities will be higher on the more complex of the two customers. Intuitively, as the selling task becomes either very complex or very simple, the difference across people in terms of performance narrows. In the limit, we will observe no difference across people on a product that is either (a) impossible to sell or (b) so easy to sell that anyone can sell it 100% of the time.

Because this will form the basis for the results to follow, it is worthwhile to discuss the robustness of the properties of the cross partial. One assumption that we have made is that the salesperson has equal priors over all actions. This understates the importance of experience because many salespeople will know that, say, 7 out of 10 times a specific action will be effective. In the Technical Appendix to this paper, available at <http://mktsci.pubs.informs.org>, we show that as long as the beliefs are such that the impression-formation stage still adds value (i.e., the salesperson may change her opinion about the best action for *some* signal realizations), and if the priors are not too different, then the cross-partial result still holds for unequal priors. We have also assumed that the impact of selling skill is equal across all actions.<sup>5</sup> It may be the case that actions  $A$  and  $B$  are associated with different impression-formation parameters,  $\sigma_A$  and  $\sigma_B$ : Some “aspects” of the customer are harder to read than others. The Technical Appendix contains a proof that, as long as the parameters are not too different, the cross-partial result holds. Finally, we assume that the salesperson keeps drawing signal vectors  $\hat{a}$  until she receives at least one  $\hat{a}_k = 1$ . It should be clear that, as

long as  $K$  is high enough,<sup>6</sup> this is not a very restrictive assumption since, as  $K$  approaches infinity, she receives a positive signal on the first draw with probability one. We formalize this intuition in the Technical Appendix.

Lemma 1 offers a parsimonious statement capturing the essence upon which we will be building in the sections to follow. It tells us about the relative impact of selling ability: Where will it be higher and when? However, one limitation of the lemma is that the relative impact is captured in terms of derivatives with respect to selling ability. In reality, of course, the difference in ability between salespeople is likely to be significant. A second limitation is that while the lemma tells us about when the relative impact is going to be higher for some customers than for others, it tells us little about the *magnitude* of this differential impact. That is, when will the impact of selling ability be *significantly* higher on, say, complex customers as compared with simpler customers? This will be important as we evaluate the cost and benefit of hiring the high-skill reps and how this relationship varies across firms. Lemma 2 addresses both of these limitations. It looks at two customers—one with complexity  $\delta^i$  and the other with complexity  $\delta^j > \delta^i$ —and two salespeople—one with skill  $\underline{\sigma}$  and another with skill  $\bar{\sigma} > \underline{\sigma}$ —and shows that the relative impact of selling skill on easy customers, as compared with difficult customers becomes very high when the task is very complex and very low when the task is very simple.

LEMMA 2. Let  $\bar{\sigma} > \underline{\sigma}$  and

$$R \equiv \frac{p(\bar{\sigma}, \delta^i) - p(\underline{\sigma}, \delta^i)}{p(\bar{\sigma}, \delta^j) - p(\underline{\sigma}, \delta^j)}.$$

(a) Let  $\delta^i = \delta^j - e$ ,  $e < \frac{1}{2}$ . For any  $\bar{\sigma}$  and  $\underline{\sigma}$ , there exists a  $\tilde{\delta}(\bar{\sigma}, \underline{\sigma})$  such that, for all  $\delta^j \geq \tilde{\delta}$ ,  $\partial R / \partial \delta^j > 0$ . Further,  $R \rightarrow_{\delta^j \rightarrow 1} \infty$ . (b) Let  $\delta^j = \delta^i + f$ ,  $f < \frac{1}{2}$ . For any  $\bar{\sigma}$  and  $\underline{\sigma}$ , there exists a  $\underline{\delta}(\bar{\sigma}, \underline{\sigma})$  such that for all  $\delta^j \leq \underline{\delta}$ ,  $\partial R / \partial \delta^j > 0$ . Further,  $R \rightarrow_{\delta^j \rightarrow 0} 0$ .

<sup>5</sup> We are indebted for this to an anonymous referee.

<sup>6</sup> In fact, we think of  $K$  as being “very” high. To see this, consider the set of possible actions that a salesperson has in her arsenal, particularly accounting for the dimensionality of this set (she can send letters, send e-mail, find reference accounts, take the customer to lunch... all of these in any combination).

**Table 1** Summary of Model Assumptions

	Monopoly with Value-Adding Salespeople	Duopoly with Value-Adding Salespeople	Monopoly with Persuasive Salespeople
Firms	<ul style="list-style-type: none"> <li>• Single firm of one of two types</li> <li>• “Better” type matches preferences of larger segment</li> </ul>	<ul style="list-style-type: none"> <li>• Two firms</li> <li>• “Better” firm matches preferences of larger segment</li> </ul>	<ul style="list-style-type: none"> <li>• Single firm of one of two types</li> <li>• “Better” type matches preferences of larger segment</li> </ul>
Customers	<ul style="list-style-type: none"> <li>• Two segments</li> <li>• Customer unsure of which segment they are in (i.e., uncertain about preferences)</li> </ul>	<ul style="list-style-type: none"> <li>• Two segments. Within each, some shop both firms others shop just one</li> <li>• Customer unsure of which segment they are in (i.e., uncertain about preferences)</li> </ul>	<ul style="list-style-type: none"> <li>• Two segments</li> <li>• Customer unsure of which segment they are in (i.e., uncertain about preferences)</li> </ul>
Salespeople	<ul style="list-style-type: none"> <li>• Two ability levels: High- and low-skill</li> <li>• Effort either “high” or “low”</li> <li>• Ability is observable but effort is not</li> <li>• One rep of each skill level available</li> </ul>	<ul style="list-style-type: none"> <li>• Two ability levels: High- and low-skill</li> <li>• Effort either “high” or “low”</li> <li>• Ability is observable but effort is not</li> <li>• Two low-skill and one high-skill reps available</li> </ul>	<ul style="list-style-type: none"> <li>• Two ability levels: High- and low-skill</li> <li>• Effort either “high” or “low”</li> <li>• Ability is observable but effort is not</li> <li>• One rep of each skill level available</li> </ul>
Role of Salespeople	<ul style="list-style-type: none"> <li>• Risk neutral</li> <li>• To “add value” by selecting a product/service enhancement for a given customer</li> </ul>	<ul style="list-style-type: none"> <li>• Risk neutral</li> <li>• To “add value” by selecting a product/service enhancement for a given customer</li> </ul>	<ul style="list-style-type: none"> <li>• Risk neutral</li> <li>• To persuade the customer by selecting the message that <i>signals</i> that the product is valuable for them</li> </ul>
Contracting	<ul style="list-style-type: none"> <li>• Firm offers contract comprised of salary and incentive</li> <li>• High-skill reps have a higher reservation wage and thus earn more in equilibrium</li> </ul>	<ul style="list-style-type: none"> <li>• Firms bid for reps’ services</li> <li>• No explicit reservation wages</li> </ul>	<ul style="list-style-type: none"> <li>• Firm offers contract comprised of salary and incentive</li> <li>• High-skill reps have a higher reservation wage and thus earn more in equilibrium</li> </ul>

### 3. The Intrafirm Hiring, Allocation, and Training of Value-Creating Salespeople

In this section, we consider a sales force with the responsibility for *creating value* for the customer. Salespeople add value to products in many ways, such as figuring out the “solution” to the customer’s problem or communicating a customer’s project specifications to the firm’s back shop. The difference between value-creating salespeople and value-signaling (or persuasive) salespeople is that the salesperson’s actions in the former add tangible value. In the latter, which we will analyze in §5, her actions only *signal* value. To aid the reader in comparing the three main models we analyze in the following sections, Table 1 summarizes the key assumptions made in each.

#### 3.1. Firm and Customer Match

Similar to the set-up used by Moorthy and Srinivasan (1995), we will assume the firm is endowed with one of two equally likely types:  $T \in \{B, W\}$ . The firm, but not the customer, observes  $T$ . A single customer is also endowed with one of two types:  $t \in \{B, W\}$ . The

customer “matches” the firm when  $T = t$ . The match will be denoted by the parameter  $\mu$ .<sup>7</sup>

$$\mu \equiv \begin{cases} m & T = t, \\ n & T \neq t. \end{cases}$$

These “types” are meant to capture heterogeneity in customer needs. Some customers have needs that are more easily met by Firm B’s products and others by Firm W’s products. Whether or not the product actually does so is a function of the salesperson’s actions. The distribution of customer needs is captured by  $\phi$ : With probability  $\phi \in (\frac{1}{2}, 1)$ ,  $t = B$ . Neither the firm nor the customer observes  $t$  ex ante, and thus they both have prior belief  $\phi$  that  $t = B$ . We call Firm type B “better” since it is positioned more favorably with respect to market needs.

The salesperson chooses one out of  $K$  actions, which we will refer to as “solutions.” Each either adds value or not. As in §2,  $a_k = 1$  if solution  $k$  adds value

<sup>7</sup> This concept of a discrete match between the firm and the customer is also used by Wernerfelt (1994).

for the customer and  $a_k = 0$  if it does not. The utility the customer derives from solution  $k$  is:

$$U = a_k - P^T, \quad (3.1)$$

where  $P^T$  is the price. The complexity of choosing a solution varies across customers and industries. We assume that it is easier to choose a valuable solution for customers that match the product:  $\delta^m < \delta^n$ . In addition, we will think about the *levels* of  $\delta^m$  and  $\delta^n$  as capturing the fact that it is harder to add value in some industries than in others. We assume that the customer effects are relatively small.<sup>8</sup>

### 3.2. The Salesperson

Salespeople are either “high skill” or “low skill”, corresponding to their ability to choose a valuable solution conditional on putting forth the effort:  $\sigma \in \{\bar{\sigma}, \underline{\sigma}\}$ , where  $\bar{\sigma} > \underline{\sigma}$ . We assume here that the firm can discriminate between salespeople of different types via noncontractual methods such as interviewing, customer referrals, or direct observation. We show in §D of the Technical Appendix that our core results also hold if the firm is uncertain about the salesperson’s skill level. Because the firm cannot observe the salesperson’s effort, it must design a contract to ensure that she takes the actions that the firm prefers she take. The only observable outcome on which this contract can be based is whether there has been a sale or not. Further, since this outcome is binary, the only contract available to the firm is to select a salary  $\alpha$ , which the salesperson receives regardless of outcome, and a commission  $\beta$  paid for a successful sale. This is an example of a “forcing contract,” as introduced by Mirrlees (1974), in that the salesperson receives one amount if and only if her production reaches a certain level and a much lower (and possibly negative) amount if it does not. She is assumed to be risk neutral and, thus, has the following utility:

$$V(e) \equiv \alpha + \beta \cdot 1(a_{k^*} = 1 | e) - C(e), \quad (3.2)$$

<sup>8</sup> Let  $\delta^n = \delta^m + \varepsilon$ . Then, we can think of  $\delta^m$  as the “industry effect” because it affects all customers the same way. We will think of  $\varepsilon$  as the “customer effect” because it captures the difference between matching and nonmatching customers.

where  $C(e)$  is the disutility associated with her effort  $e$ . Two comments are in order with respect to Equation (3.2). First, the assumption of risk neutrality, while not a realistic one, is not restrictive. We make it here purely for ease of exposition. We show in the Technical Appendix that the qualitative nature of these results carries over to the risk-averse case. Moreover, this simplifying assumption is not entirely uncommon. See, for example, Rao (1990). Second, taken literally, the equation implies that the salesperson is paid for selecting a valuable solution rather than making a sale. The former is, of course, likely to be unobservable in all but a very small subset of sales contexts. However, the firm will always set the price to one, so a sale is made if and only if the solution adds value. To see why this is true, refer back to Equation (3.1). Each customer’s utility from a purchase, following the sales interaction, can take one of two values, either  $-P^T$  or  $1 - P^T$ , regardless of type. In the former state, the firm cannot earn positive profits and, thus, will set the price to maximize profit in the latter state. This is accomplished by setting  $P^T = 1$ , which yields a sale if and only if the solution adds value. Thus, the selection of a valuable solution and the closing of a sale are equivalent.

We will assume that the salesperson’s reservation wage  $V^{\min}(\sigma)$  is nondecreasing in her ability:  $V^{\min}(\bar{\sigma}) \geq V^{\min}(\underline{\sigma}) \equiv 0$ .<sup>9</sup> Assuming that the reservation wage may differ across agents is not uncommon, particularly in the labor economics literature (see, for example, Albrecht and Vroman 1992, Mailath and Postlewaite 1990, Booth 1984, Malcomson 1981, and Weiss 1980). We will specify the high-skill rep’s reservation wage as a linear function of the marginal return on her ability:

$$V^{\min}(\bar{\sigma}) \equiv \gamma \cdot \max_{T \in \{B, W\}} \{P^T \cdot [\Pr[a_{k^*} = 1 | \bar{e}, \bar{\sigma}, T] - \Pr[a_{k^*} = 1 | \bar{e}, \underline{\sigma}, T]]\}, \quad (3.3)$$

<sup>9</sup> The assumption that  $V^{\min}(\underline{\sigma}) = 0$  is purely without loss of generality given our working assumption of risk neutrality. Moreover, under the standard exponential utility formulation, an assumption that  $V^{\min}(\underline{\sigma}) > 0$  would simply involve the carrying of a constant term in our equations and would have no qualitative impact on the results.



where  $\gamma \in [0, 1)$  captures the relative power of the salesperson vis-à-vis the firm. So, the salesperson captures a proportion  $\gamma$  of the marginal return available to the firm type that experiences the *highest* potential returns on selling ability. Because the firm types differ in the returns they enjoy from selling skill, we have chosen to simply anchor the reservation wage on the maximum. Higher  $\gamma$ 's mean she is able to capture more of the value associated with her skill. Note that since we are dealing with only two types of salespeople, nothing would change if the specified  $\gamma$  is a function of the salesperson's ability:  $\gamma(\sigma)$ . Because better salespeople are probably better at negotiating their compensation packages, we might further restrict this to be increasing:  $\gamma' \geq 0$ . For notational brevity, we will maintain the simpler specification of  $\gamma$  while keeping in mind that nothing changes for  $\gamma = \gamma(\sigma)$ .<sup>10</sup>

Selling is a difficult task. The process of reading the customer, asking the correct set of questions to ascertain his needs and formulating an optimal solution to his problem requires significant effort on behalf of the salesperson. It is reasonable to assume that higher effort will improve the likelihood that a valuable solution will be selected. We capture this by modeling the salesperson as choosing an effort level  $e \in \{e, \bar{e}\}$ , "low effort" or "high effort," respectively, such that:

$$\begin{aligned} \Pr[\hat{a}_k = a_k | \bar{e}, \bar{\sigma}] &= \bar{\sigma}, & \Pr[\hat{a}_k = a_k | \bar{e}, \underline{\sigma}] &= \underline{\sigma}, \\ \Pr[\hat{a}_k = a_k | e, \bar{\sigma}] &= \frac{1}{2}, & \Pr[\hat{a}_k = a_k | e, \underline{\sigma}] &= \frac{1}{2}. \end{aligned} \quad (3.4)$$

Defining  $\underline{C} \equiv C(e)$ ,  $\bar{C} \equiv C(\bar{e})$ , and  $\Delta C \equiv \bar{C} - \underline{C}$ , the salesperson puts forth  $\bar{e}$  if:<sup>11</sup>

$$\beta \geq \frac{\Delta C}{\Pr[a_{k^*} = 1 | T, \sigma, \bar{e}] - \Pr[a_{k^*} = 1 | T, \sigma, e]}. \quad (3.5)$$

<sup>10</sup> We are grateful to an anonymous referee for this insight.

<sup>11</sup> We are very grateful to an anonymous referee for pointing out the similarity between this incentive compatibility constraint and that in the traditional agency model. In the latter, output as a function of effort is  $x = f(e) + \varepsilon$  where  $\varepsilon \sim N(0, \sigma^2)$ . The agent's certainty equivalent can be represented as  $\alpha + \beta f(e) - C(e) - \frac{1}{2} r \beta^2 \sigma^2$ , where  $r$  is the agent's coefficient of absolute risk aversion. The agent's optimal effort choice must then satisfy  $\beta = (\partial C / \partial e) / (\partial f / \partial e)$ , which is effectively a continuous-space analog of Equation (3.5).

Further, the firm will choose to implement  $\bar{e}$  as long as:

$$\begin{aligned} P^T \cdot \{ \Pr[a_{k^*} = 1 | T, \sigma, \bar{e}] - \Pr[a_{k^*} = 1 | T, \sigma, e] \} \\ \geq \Delta C. \end{aligned} \quad (3.6)$$

We will assume that  $\Delta C$  is low enough or  $\sigma$  high enough that Inequality (3.6) holds and that the firm will therefore select a high enough  $\beta$  to satisfy (3.5) and implement high effort.<sup>12</sup> Note that since (3.5) is an inequality, there is a continuum of  $\beta$ 's that will satisfy it. Moreover,  $\alpha$  is set to the value that ensures that the salesperson earns exactly the amount given by (3.3) in expectation:

$$\alpha = V^{\min}(\sigma) - \beta \cdot \Pr[a_{k^*} = 1 | T, \sigma, \bar{e}] - \bar{C}. \quad (3.7)$$

Thus, there is a continuum of  $\{\alpha, \beta\}$  pairs, each of which (a) ensures that the salesperson puts forth high effort and (b) pays the salesperson exactly her reservation wage in expectation. Combined, these two imply directly that each of the  $\{\alpha, \beta\}$  pairs along this continuum achieves exactly the same expected profit outcome for the firm. It is interesting to note that one such pair is that which "sells her the firm." This is accomplished by setting  $\beta = 1$  and  $\alpha = V^{\min}(\sigma) - \Pr[a_{k^*} = 1] - \bar{C}$ . Again, this yields the same profit outcome for the firm as all of the others along the continuum.

Finally, we have assumed that the salesperson has no better information about the customer's type than does the customer himself. One could certainly imagine that she first receives a signal about the customer and *then* makes her effort decision. It is straightforward to show that this would have no qualitative

<sup>12</sup> It is very likely that a high-skill rep will be better than a low-skill rep even at low effort levels. The assumption that  $\Pr[\hat{a}_k = a_k | e, \sigma] = \frac{1}{2} \forall \sigma$  is made simply for ease of exposition and is not important to our results. On one hand, it ensures that the high-skill rep has a higher marginal impact of effort, which is a desirable property. More important, allowing this to vary across salespeople would only change Equations (3.5) and (3.6). The effect on the former would simply be to change the incentive portion of the compensation scheme. However, under risk neutrality, there would be a change to the salesperson's salary that would exactly offset this, leaving the firm's total compensation costs unaffected. As for the latter, assuming that  $\Delta C$  is small would still be sufficient to ensure that the firm prefers high effort.

impact on the results that follow because it would simply change the incentive compatibility constraint in (3.5). Specifically, the denominator of (3.5) now reflects the fact that the salesperson has no additional information about  $\mu$ . This could easily be restated to capture the fact that  $\beta$  will have to be high enough to encourage her to put forth  $\bar{e}$  regardless of her belief about  $\mu$  after receiving a signal.

### 3.3. Optimal Policy for Hiring and Allocating Salespeople

One question facing a firm is what type of salespeople it will employ. This involves a trade-off between the marginal return on selling skill and the ability of the salesperson to extract that return. Formally, the firm's decision is:

$$\text{Max}_{\{h^T \in \{\bar{\sigma}, \underline{\sigma}\}, \alpha, \beta, P^T\}} (P^T - \beta) \cdot \Pr[a_{k^*} = 1 \mid T, h^T, \bar{e}] - \alpha, \quad (3.8)$$

subject to:

$$\alpha + \beta \cdot \Pr[a_{k^*} = 1 \mid T, h^T, \bar{e}] - \bar{C} \geq V^{\min}(h^T) \quad (3.9)$$

and Equation (3.5), where firm  $T$ 's sales,  $\Pr[a_{k^*} = 1 \mid T, h^T, \bar{e}] = p(h^T, \delta^m) \cdot \Pr[\mu = m \mid T] + p(h^T, \delta^n) \cdot \Pr[\mu = n \mid T]$  with  $p(\cdot)$  given by (2.5). Equations (3.5) and (3.9) are the standard incentive compatibility and participation constraints, respectively. Since conditional on (3.5), (3.8) is strictly decreasing in  $\alpha$  and  $\beta$ , there is no loss of generality in setting (3.9) to equality, which allows us to then substitute it back into (3.8). The firm's objective is then to maximize  $P^T \cdot \Pr[a_{k^*} = 1 \mid T, h^T, \bar{e}] - V^{\min}(\bar{\sigma}) - \bar{C}$  subject to Equation (3.5). As Proposition 1 shows, the solution is moderated by the task's complexity:

**PROPOSITION 1.** Let  $R \equiv (p(\bar{\sigma}, \delta^m) - p(\underline{\sigma}, \delta^m)) / (p(\bar{\sigma}, \delta^n) - p(\underline{\sigma}, \delta^n))$  and  $\Phi \equiv \phi / (1 - \phi)$ . Then, (a) when  $\delta^m$  and  $\delta^n$  are high,  $h^B = \bar{\sigma}$  and  $h^W = \underline{\sigma}$  if and only if  $\gamma \geq (\Phi + R) / (\Phi R + 1)$ . Otherwise,  $h^B = \bar{\sigma}$  and  $h^W = \bar{\sigma}$ . (b) When  $\delta^m$  and  $\delta^n$  are low,  $h^W = \bar{\sigma}$  and  $h^B = \underline{\sigma}$  if and only if  $\gamma \geq (\Phi R + 1) / (\Phi + R)$ . Otherwise,  $h^W = \bar{\sigma}$  and  $h^B = \bar{\sigma}$ .

Proposition 1 is an application of Lemma 1. Here, firm type  $B$ 's task is always easier because  $\phi > \frac{1}{2}$ . Thus, it has a higher (lower) marginal value for selling ability compared with firm type  $W$  when the

task is complex (simple). When the task is simple, as long as the high-skill rep is expensive enough, then only  $W$  hires her. Of course, it is useful to ask how expensive is "expensive enough?" By Lemma 2,  $R$  approaches infinity for high  $\delta$ 's and vanishes for low  $\delta$ 's. Thus, as  $\phi$  approaches 1 and  $\delta^m$  approaches 0,  $(\Phi R + 1) / (\Phi + R)$  vanishes. So, when the products are sufficiently different in the extent of their appeal and the task is sufficiently simple, *only* firm type  $W$  will hire the highly skilled salesperson as long as she earns an *epsilon* more than the low-skill salesperson. The analogous result holds for complex products. Since  $R > 0$ , both firms will always hire the high-skill rep at  $\gamma = 0$ . Note also that, because  $\gamma$  appears only on the LHS of these conditions, it would be straightforward to allow for the more general specification of  $\gamma(\sigma)$  by simple substitution. Specifically, if the high-skill rep has ability  $\bar{\sigma}'$ , then the condition becomes  $\gamma(\bar{\sigma}') \geq (\Phi + R(\bar{\sigma}')) / (\Phi \cdot R(\bar{\sigma}') + 1)$ .

We believe the form of the reservation wage specified in Equation (3.3) is an attractive and realistic one for several reasons. First, it seems reasonable to expect that a salesperson's reservation wage will vary across industries. Thus, by making it a function of  $\phi$  and the  $\delta$ 's, we allow for this. Specifically, there are some industries (values of  $\delta$ ) in which the impact of ability is more valuable. We would expect—and this specification captures the fact that—the premium paid to high-skill reps would be highest in these industries. Moreover, this specification allows us to analyze the optimal hiring policy in terms of a single—and an interesting—parameter: the sales rep's negotiating power. However, it is essential to emphasize that the qualitative nature of the result—that we expect to see more highly skilled reps selling for better firms in complex industries, but worse firms in simple industries—is not dependent on this specification. The Technical Appendix (Section C, General Reservation Wage) provides a proof to this effect. Specifically, Proposition C.1 shows that for a general set of reservation wage specifications,  $h^W = \bar{\sigma}$  and  $h^B = \underline{\sigma}$  only when  $\delta$  is low, and  $h^W = \underline{\sigma}$  and  $h^B = \bar{\sigma}$  only when  $\delta$  is high. The conditions that we place on  $V^{\min}(\bar{\sigma})$  in this general proposition are similar to those we use for the specification in (3.3) and have the effect that the reservation wage is "moderate," not too high

or too low. If  $V^{\min}(\bar{\sigma})$  is too high, neither firm type would ever hire her. We would expect that the sales rep would most likely leave the industry or work for less money in this case. This restriction is equivalent to our assumption here that  $\gamma < 1$ . If  $V^{\min}(\bar{\sigma})$  is too low, there will not be cases in which firms with differential value will make different hiring decisions. In the limit, if her reservation wage approaches the low-skill rep's, all firm types will want to hire the high-skill rep regardless of the value that she would bring to their business. This restriction is equivalent to the conditions on  $\gamma$  stated in Proposition 1.

It is straightforward to extend this analysis to a multiproduct setting. Consider a firm with a fixed staff comprised of one high-skill and one low-skill salesperson. It must decide to which of two products each rep will be assigned: Product  $B$  is better (easier to sell) than product  $W$ . Here, the labor market considerations are moot because the firm's expected total labor cost is independent of the allocation of salespeople:

$$E[V(\bar{\sigma}) + V(\underline{\sigma})] \equiv V^{\min}(\bar{\sigma}) + V^{\min}(\underline{\sigma}) + 2\bar{C}. \quad (3.10)$$

Proposition 2 shows that the intuition is similar whether we are considering a single-product firm making a hiring decision or a multiple-product firm organizing its reps.

**PROPOSITION 2.** *In the multiproduct model, when  $\delta^m$  and  $\delta^n$  are high,  $h^B = \bar{\sigma}$  and  $h^W = \underline{\sigma}$ . When  $\delta^m$  and  $\delta^n$  are low,  $h^B = \underline{\sigma}$  and  $h^W = \bar{\sigma}$ .*

### 3.4. Investment in Salesforce Training

We can use a similar model to address an analogous question: "In which sales forces should the firm invest training resources?" We will consider a firm with two products,  $B$  and  $W$ , in the same industry with product  $B$  being better than  $W$  in the sense that we have defined. The firm currently employs two salespeople of equal skill  $\sigma$ , one on each product, and has the option of investing an amount  $I$  to increase the skill of a salesperson to  $\sigma' > \sigma$ . Assuming that switching costs are high enough that salespeople will not be more likely to change jobs as a result of the training, the firm is only concerned about the marginal increase in sales force effectiveness as compared with

the investment. Proposition 3 shows that the intuition we have developed thus far also holds in the allocation of training resources:

**PROPOSITION 3.** (a) *The firm invests only in the better product's sales force when*

$$\begin{aligned} \phi(p_g^m - p_b^m) + (1 - \phi)(p_g^n - p_b^n) \\ \geq I \geq (1 - \phi)(p_g^m - p_b^m) + \phi(p_g^n - p_b^n). \end{aligned}$$

(b) *The firm invests only in the worse product's sales force when*

$$\begin{aligned} \phi(p_g^m - p_b^m) + (1 - \phi)(p_g^n - p_b^n) \\ \leq I \leq (1 - \phi)(p_g^m - p_b^m) + \phi(p_g^n - p_b^n). \end{aligned}$$

Moreover, there exists a single  $\hat{\delta}^m$  such that for all  $\delta^m > \hat{\delta}^m$  ( $\delta^m < \hat{\delta}^m$ ), the region given in condition (a) is nonempty (empty), and the region given in condition (b) is empty (nonempty).

Proposition 3 says that there are two conditions for the firm to optimally invest only in the better product's sales force. First, it must be the case that the return on investment training must be higher in product  $B$  than product  $W$ . This occurs when the products are complex enough or when the  $\delta$ 's are high enough. Assuming that this holds, it must also be the case that the cost of training is neither too high nor too low. Even though product  $B$  would benefit more from training, it might be the case that the cost is so high that *neither* product would find it profitable to invest in training. Alternatively, if the cost is so low, investment into *both* products would be optimal.

## 4. The Market for Value-Creating Salespeople

It is very common for a firm's salespeople, perhaps more than most of its employees, to be hired away by a competitor. However, given that this process often amounts to a bidding war, it is not obvious why poaching exists in equilibrium. For example, assume that if firm  $W$  hired firm  $B$ 's salesperson, it would add an amount  $X$  in incremental sales as compared with a salesperson of lesser ability. Further, this poaching would decrease firm  $B$ 's sales by the same amount. A bidding war between the firms should result in the

salesperson being paid a premium of  $X$  over what less-skilled salespeople are paid. Thus, firm  $W$  should be *at best* indifferent between poaching or not poaching. In fact, there are at least two additional factors, which we will not model here, that might work in the other direction. First, there might be a cost associated with hiring (for example, management time for interviewing, training, etc.). Also, there is likely to be informational asymmetry: The current employer knows best what the salesperson's true ability is. This "winner's curse" should further decrease the probability of poaching. While this argument may be convincing theoretically, the data are hardly supportive. In 1999, U.S. firms lost an average of almost 12% of their sales reps. Retail experienced the highest turnover rates: 19% (Marchetti 2000). In this section, we argue that poaching may be optimal because selling skill has a differential impact across firms. Our model also predicts the *direction* of the poaching as a function of the complexity of the industry: When will the "better" firms steal reps from their rivals and when will they lose reps to their rivals?

#### 4.1. Competitive Model

As Figure 1 shows, the game has two essentially identical periods separated by an information revelation stage. We consider a context with two firms  $T = B, W$ . The firms compete for customers *and* salespeople. The product market competition we will model is similar to that analyzed by Narasimhan (1988). As in §3, the firms differ in that the customer matches firm  $B$ 's product with probability  $\phi > \frac{1}{2}$ , and  $W$ 's product with probability  $1 - \phi$ . The customer does not know which firm he matches. Also, as above, it is *ceteris paribus* easier for the salesperson to choose a solution for a matching customer. This has the implication that it is easier for firm  $B$ 's product to be sold. So, we will again refer to firm  $B$ 's as the "better"

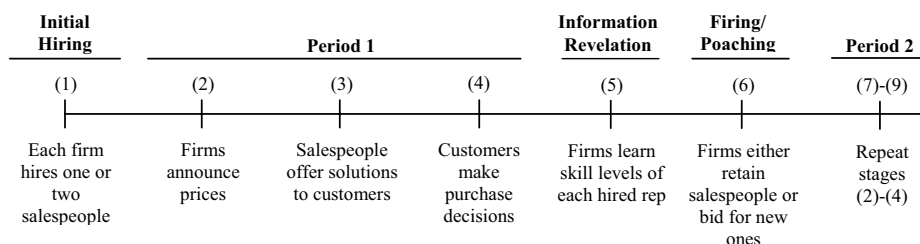
product. In this model, some customers shop both firms and others only shop one of the firms. This might be due to many reasons, including informational asymmetries (not all customers *know* about the existence of all stores), differences in search costs, or positive past experiences with other products. Each firm  $T$  faces a single customer. With probability  $\lambda \in (0, 1)$ , he shops *only* store  $T$  and with probability  $1 - \lambda$  he also shops at the other firm. It is important to draw a distinction between  $\phi$  and  $\lambda$ . On one hand,  $\phi$  captures the symmetry between the firms in terms of their market appeal. The higher  $\phi$  is, the higher the probability that the customers' problems are more easily solved by Firm  $B$ 's products than by Firm  $W$ 's. Thus, market needs are more favorable to the better firm than to the worse. Higher  $\lambda$ 's mean that fewer customers receive proposals from both firms. Thus, we interpret  $\lambda$  as capturing the intensity of competition in the market. It might be useful to compare these two parameters to their analogs in the standard Hotelling models. The "transportation cost" in those models is analogous to our  $\lambda$  since it moderates the degree to which people will consider both products. The distribution of consumers along the line is analogous to our  $\phi$  parameter since it captures the relative preferences of customers vis-à-vis the firms.

As above, the salesperson selects one out of  $K$  solutions that either adds value for the customer or not:

$$a_k^T = \begin{cases} 0 & \text{Solution } k \text{ adds value for} \\ & \text{the customer,} \\ 1 & \text{Solution } k \text{ does not add value} \\ & \text{for the customer.} \end{cases} \quad (4.1)$$

Here, we use superscript  $T$  because some customers receive proposed solutions from *both* firms. The customer receives no utility if  $a_{k^*}^T = 0$ , while the utility

Figure 1 Competitive Model



from a valuable solution now depends in part on whether or not the customer matches. We will use the following utility function for product  $T$ :

$$U^T \equiv a_{k^*}^T(1 + \rho \cdot 1(T = t)) - P^T, \quad (4.2)$$

where  $\rho > 0$ . We will assume that  $\rho$  is “small.” So, the utility from a valuable solution is 1 if the customer does not match and  $1 + \rho$  if he does. The customer chooses the product offering the highest positive utility. As in Narasimhan’s (1988) extension to his model,  $\rho$  serves as a “tie-breaker” when he receives a valuable solution from *both* firms. For example, imagine that after shopping for a new pickup truck at both Ford and Chevy dealerships someone found that both offered trucks that suited his needs ( $a^{\text{Ford}} = a^{\text{Chevy}} = 1$ ). If he only saw the Chevy, he would have purchased it. However, having seen both, he *ceteris paribus* prefers the Ford.

Before proceeding, it is useful to present the following lemma that shows that this model has a unique symmetric pure-strategy price equilibrium for small  $\rho$ . Note that we *require* neither symmetry nor pure strategies for our results. However, since the price equilibrium is not our focus we will maintain the assumptions that give rise to these prices for ease of exposition.

LEMMA 3. (a) *For complex enough tasks, the unique pure-strategy price equilibrium is  $P^W = P^B = 1 + \rho$ ; (b) for simple enough tasks, when  $\rho$  is low enough, the unique pure-strategy price equilibrium is  $P^B = P^W = 1$ .*

In Period 1, the true skill level of each salesperson  $\sigma \in \{\bar{\sigma}, \underline{\sigma}\}$  is not perfectly discernable by either the firm or the salesperson: Nobody is exactly sure how well she will do. However, based on interviews, reference checks, or the “word on the street,” the parties are able to distinguish those that are “likely to be highly skilled” from those that are not. Specifically, salespeople are distinguishable by two types  $\hat{\sigma} \in \{\bar{\sigma}, \underline{\sigma}\}$  such that  $\Pr[\sigma = \bar{\sigma} \mid \hat{\sigma} = \bar{\sigma}] = 1 - \nu > \frac{1}{2}$  and  $\Pr[\sigma = \bar{\sigma} \mid \hat{\sigma} = \underline{\sigma}] = \nu < \frac{1}{2}$ . The parameter  $\nu$  captures the difficulty in skill assessment. There are three salespeople, one of type  $\bar{\sigma}$  and two of type  $\underline{\sigma}$ . The relative scarcity of type  $\bar{\sigma}$  salespeople captures the fact that high-skill reps are generally a rare commodity. Since each firm faces a single customer and each salesperson can “serve” one customer, there is an adequate

supply of salespeople. Below, we will consider a context in which there are more customers and thus a shortage of salespeople. After the initial hiring process, the firms compete in the product market. Besides generating revenue, this also provides information about the salespeople. We will assume that the true skill level  $\sigma^T \in \{\bar{\sigma}, \underline{\sigma}\}$  of the rep hired by firm  $T$  is perfectly revealed to all parties in this process. The abilities of the nonhired reps are not revealed.

#### 4.2. Hiring Process

In this model the demand for, and supply of, selling talent determines the reps’ compensation. Thus, there is no need for explicit reservation wages. Since the presence of two firms provides the rep with an *explicit* outside option, we do not need a reservation wage to provide an *implicit* one. The hiring process begins in each period with the firms simultaneously announcing their initial bids. Each bid is a triplet  $b^T = (b_\alpha^T, b_\beta^T, b_\sigma^T)$ , where  $b_\alpha^T, b_\beta^T$  are the salary and commission, respectively, and  $b_\sigma^T \in \{\bar{\sigma}, \underline{\sigma}\}$  the salesperson on whom they are bidding. The rep always accepts the bid with the highest expected value. We will again assume that  $\Delta C$  is low enough that the firm wants to implement the high effort level no matter the level of skill of the salesperson hired. To ensure incentive compatibility, then, any bid for salesperson  $\bar{\sigma}$  must include a high-enough commission  $b_\beta^T$ .

$$b_\beta^T \geq \Delta C \cdot \{(1 - \nu)(\Pr[\text{Sale} \mid T, \bar{\sigma}, \bar{e}] - \Pr[\text{Sale} \mid T, \bar{\sigma}, e]) + \nu(\Pr[\text{Sale} \mid T, \sigma, \bar{e}] - \Pr[\text{Sale} \mid T, \sigma, e])\}^{-1}. \quad (4.3)$$

An analogous constraint exists for bids for salesperson  $\underline{\sigma}$ . Besides this, we place no restriction on the bids. In particular, there is no constraint on  $b_\alpha^T$ . As we saw in §3, the principal role of the intercept is to scale the contract to ensure that the agent earns her reservation wage in expectation. Because we have no reservation wage in this model, we need place no restriction in it. In fact, due to risk neutrality, we could assume that  $b_\alpha^T = 0$  with absolutely no loss of generality.

If  $b_\sigma^T \neq b_\sigma^W$ , each firm hires salesperson  $b_\sigma^T$  at salary  $b_\alpha^T$  and commission  $b_\beta^T$ , and the hiring process ends. If  $b_\sigma^T = b_\sigma^W$ , the firms engage in a bidding war for  $\hat{\sigma}$ . This commences with a randomly chosen firm submitting another bid and continues until one of the

firms drops out by submitting a bid for a different salesperson. We place two restrictions on the bidding. First, each bid must be strictly higher than the previous bid by at least  $\varepsilon$ . This reflects both a practical limitation (firms cannot offer fractions of a cent) as well as industry norms (it is unlikely that a bid of \$1 more than a rival's would be looked upon favorably by the candidate). This also ensures that bidding ends in finite time. Second, a firm may not renege once the other drops out. The outcome of this bidding process is the set of hiring policies  $h_1^T \in \{\bar{\sigma}, \underline{\sigma}\}$  and  $h_2^T(\sigma^B, \sigma^W) \in \{\bar{\sigma}, \underline{\sigma}, \bar{\sigma}, \underline{\sigma}\}$  for Periods 1 and 2, respectively. The Period 2 hiring policy is a function of the true skill levels of the Period 1 hires.

#### 4.3. Results

We will address two questions with this model. First, we would like to check that the intuition developed in the previous sections holds in a competitive setting. Do high-skill reps sell better (worse) products when the industry is complex (simple)? Is this result robust to uncertainty about the reps' skill? Second, we would like to see what the model says about the interfirm movement of salespeople.

**PROPOSITION 4.** *For  $\varepsilon$  small enough,<sup>13</sup> when the task is complex enough, all subgame perfect equilibria have  $h_1^B = \bar{\sigma}$  and  $h_1^W = \underline{\sigma}$ . When the task is simple enough, all subgame perfect equilibria have  $h_1^B = \underline{\sigma}$  and  $h_1^W = \bar{\sigma}$ . Period 2 policies are as follows:*

	$\sigma^B = \underline{\sigma}$	$\sigma^B = \bar{\sigma}$
$\sigma^W = \underline{\sigma}$	$h_2^B = \underline{\sigma}, h_2^W = \underline{\sigma}$	$h_2^B = \bar{\sigma}, h_2^W = \underline{\sigma}$
$\sigma^W = \bar{\sigma}$	$h_2^B = \bar{\sigma}, h_2^W = \underline{\sigma}$	$h_2^B = \bar{\sigma}, h_2^W = \bar{\sigma}$
Complex Task		
	$\sigma^B = \underline{\sigma}$	$\sigma^B = \bar{\sigma}$
$\sigma^W = \underline{\sigma}$	$h_2^B = \underline{\sigma}, h_2^W = \underline{\sigma}$	$h_2^B = \underline{\sigma}, h_2^W = \bar{\sigma}$
$\sigma^W = \bar{\sigma}$	$h_2^B = \underline{\sigma}, h_2^W = \bar{\sigma}$	$h_2^B = \bar{\sigma}, h_2^W = \bar{\sigma}$
Simple Task		

<sup>13</sup> We require small enough  $\varepsilon$  to avoid outcomes in which one firm places a value of  $X$  on the high-skill salesperson while the other firm places a value of  $Y < X$ , but the latter firm wins her services with a bid of  $b \in (X - \varepsilon, Y)$ .

The rows and columns of each of these matrices denote the realization of the hired salespeople's skill levels. So, in each matrix, the lower left corner describes the case in which the salesperson hired by firm  $W$  turned out to be highly skilled while the salesperson hired by firm  $B$  did not. Proposition 4 reinforces the idea that selling skill may be increasing or decreasing in product quality depending on the complexity of the task. This is seen first in the Period 1 hiring policies, in which the better (worse) firm hires the salesperson that "looks" like she has high (low) skill when the industry is complex (simple). Thus, our results are robust to uncertainty. Most interesting is what happens when the skills are revealed. When the task is complex, a high-skill salesperson is always hired in Period 2 by the better firm, as long as such a salesperson exists. With probability  $\nu^2$ , the worse firm hires a high-skill salesperson and the better firm hires a low-skill salesperson in Period 1. This is the shaded cell in the complex task matrix on the left. In this case, the better firm poaches the worse firm's rep in Period 2. When the industry is simple enough, the poaching is in exactly the opposite direction. This suggests that hiring error offers a parsimonious explanation for the existence of poaching. Poaching occurs only (but not necessarily) by the better (worse) firm when the industry is complex (simple). This leads to the testable hypothesis that poaching levels are increasing in the difficulty of ex ante skill evaluation. A final important implication of Proposition 4 is that the firm one works for, or the quality of the product one sells, may be an informative signal of one's selling skill. In complex industries, the fact that one sells the better product is a signal of high skill, while the opposite is true in a simple industry.<sup>14</sup>

#### 4.4. Undersupply of Salespeople

The model presented above suggests that salespeople change firms due to the resolution of uncertainty surrounding a salesperson's ability. Of course, in practice, interfirm movement is likely to occur for other reasons as well. In particular, it is likely that salespeople move between firms due to the general scarcity

<sup>14</sup> For example, on average,  $\Pr[\sigma = \bar{\sigma} \mid T = B, \text{high } \delta] = (2 - \nu)/2 > \frac{1}{2}$ .

of sales talent.<sup>15</sup> This might be due, for example, to a short-term mismatch between the product market demand and the supply of trained salespeople. This phenomenon could arise from either demand-side factors (for example, in markets experiencing high growth) or supply-side factors (due to the exit of salespeople from the prospect pool). We will capture the mismatch in terms of the former, but expect that the results would not differ in a supply-side model. We will allow for the exogenous growth of potential demand in Period 2, while assuming that the supply of salespeople remains constant. Specifically, we will continue to assume that each salesperson can service a single customer and look at the case in which each firm faces one customer in Period 1 and *two* customers in Period 2. For ease of exposition, we will assume that the customers are indistinguishable in the sense that the firms cannot tell *ex ante* whether they have visited or will visit the other firm or not. Thus, we abstract away from strategic considerations associated with the firm's equilibrium assignment of salespeople to specific customers.<sup>16</sup> Each firm  $T$ 's hiring strategy in Period 1 is now a pair:  $\tilde{h}_1^T = (h_{11}^T, h_{12}^T)$  where  $h_{1i}^T \in \{\bar{\sigma}, \underline{\sigma}, \emptyset\}$ ,  $i = 1, 2$ . The second subscript  $i$  denotes the "slot," or opening, for salespeople; each firm has two potential slots. Without loss of generality, we will assume that each firm fills Slot 1 first. An empty slot is denoted by  $\emptyset$ . Similarly, in Period 2:  $\tilde{h}_2^T = (h_{21}^T(\bar{\sigma}^B, \bar{\sigma}^W), h_{22}^T(\bar{\sigma}^B, \bar{\sigma}^W))$  where  $h_{2i}^T \in \{\bar{\sigma}, \underline{\sigma}, \bar{\sigma}, \underline{\sigma}, \emptyset\}$ . In the second period, the firms' decisions are a function of the realized skill levels of the hired salespeople from Period 1. To highlight the key intuition of this extension, we will focus on the context in which the realized skills are  $\bar{\sigma}$  and  $\underline{\sigma}$ . This is the most interesting because it offers insight into what happens when there is a broad distribution of talent available on the market. We will also simplify the problem somewhat by focusing on the limiting cases:  $\delta^n \rightarrow 1$  for very complex products and

$\delta^m \rightarrow 0$  for very simple products. The following corollary shows that there may be poaching even in the absence of hiring errors and that poaching due to undersupply may occur in either the same or the opposite direction as poaching due to hiring errors. Further, we find that there may be poaching of reps in *both directions* under some conditions.

**COROLLARY 1.** Assume that the revealed skill levels after Period 1 are  $\bar{\sigma}$  and  $\underline{\sigma}$ . (a) For  $\varepsilon$  and  $\rho$  small, as  $\delta^n \rightarrow 1$ , all subgame perfect equilibria have  $P^W = P^B = 1 + \rho$ ,  $h_1^B = (\bar{\sigma}, \emptyset)$ , and  $h_1^W = (\underline{\sigma}, \emptyset)$ . Further,  $\tilde{h}_2^B(\bar{\sigma}, \underline{\sigma}) = \tilde{h}_2^B(\underline{\sigma}, \bar{\sigma}) = (\bar{\sigma}, \underline{\sigma})$  and  $\tilde{h}_2^W(\bar{\sigma}, \underline{\sigma}) = \tilde{h}_2^W(\underline{\sigma}, \bar{\sigma}) = (\underline{\sigma}, \emptyset)$ . (b) For  $\varepsilon$  and  $\rho$  small, as  $\delta^m \rightarrow 0$ , all subgame perfect equilibria have  $P^W = P^B = 1$ ,  $h_1^W = (\bar{\sigma}, \emptyset)$ , and  $h_1^B = (\underline{\sigma}, \emptyset)$ . Further, for each set of parameters  $\{\delta^m, \delta^n, \bar{\sigma}, \underline{\sigma}, \nu, \phi, \lambda\}$ , there exists a triplet  $\{F', F'', Z\}$  such that the Period 2 strategies are:

- (I) Either  $F'' < \Phi < F'$ , or  $\Phi < \min\{F', F''\}$  and  $Z < 0$ :  $h_2^B = (\bar{\sigma}, \underline{\sigma})$ ,  $h_2^W = (\underline{\sigma}, \emptyset)$ ,
- (II)  $\Phi > \max\{F', F''\}$ :  $h_2^B = (\underline{\sigma}, \underline{\sigma})$ ,  $h_2^W = (\bar{\sigma}, \emptyset)$ ,
- (III)  $F' < \Phi < F''$ :  $h_2^B = (\bar{\sigma}, \underline{\sigma})$ ,  $h_2^W = (\underline{\sigma}, \emptyset)$ ,
- (IV)  $\Phi < \min\{F', F''\}$  and  $Z > 0$ :  $h_2^B = (\underline{\sigma}, \emptyset)$ ,  $h_2^W = (\bar{\sigma}, \underline{\sigma})$ .

The proof of this corollary is available in the Technical Appendix. Before analyzing the intuition behind the results, note the fascinating implications. First, we will get poaching even when hiring errors are arbitrarily small. This occurs in the case of simple products in Equilibria I and III. In both cases, the worse firm hires the rep that "looks" good in Period 1. Assuming that these expectations are met, the better firm poaches the high-skill rep in Period 2. Moreover, this implies that, in these regions, the better firm will be doing the poaching even for simple products. This is the opposite of what we found when there was no undersupply of reps. Of course, the worse firm still poaches in the other regions—II and IV—when there are hiring errors. Finally, these results suggest that we may simultaneously observe poaching by *both* firms. This occurs, for example, in Region II with probability  $\nu^2$ . If both firms made a hiring error, then  $\sigma^B = \bar{\sigma}$  and  $\sigma^W = \underline{\sigma}$ . In this region, the worse firm will hire away the high-skill rep from the better *and* the better firm will hire away the low-skill rep from the worse.

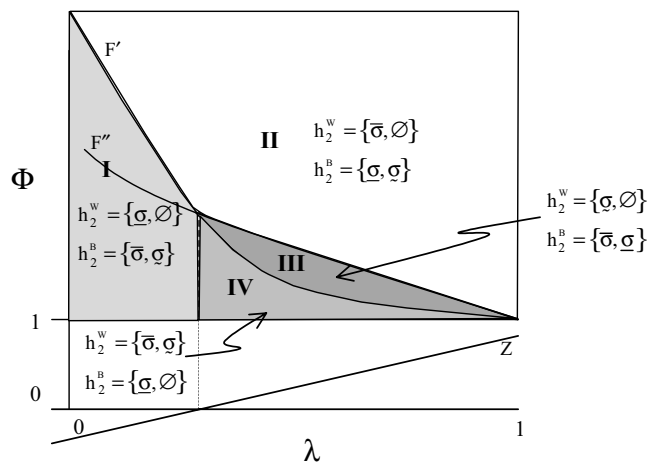
<sup>15</sup> We are very grateful to an anonymous reviewer for the suggestion that led to this extension.

<sup>16</sup> Since the firm does not know whether a customer has seen another salesperson or whether he will do so, it cannot condition the assignment of a specific salesperson (or no salesperson at all) on this fact.

The divergence between these results and those preceding derives from the fact that our question to this point has been *which* salespeople the firm should hire. Here, we have added a second concern: *How many*? Note that our results for complex products are identical to those in Proposition 4. Thus, these results are robust to a temporary undersupply of salespeople. Moreover, the *which* and the *how many* questions yield symmetric answers here: The firm that hires the most also hires the most highly skilled. This is not, however, generally true for simple products. The Period 2 hiring policies in the simple context are shown graphically in Figure 2. One striking feature of these policies is that in some regions (I and IV), the firm that hires the most salespeople also hires the most highly skilled salespeople. On the other hand, in other regions (II and III), this is not true. Consider a case in which  $W$  and  $B$  have each hired a salesperson of ability  $\sigma^W$  and  $\sigma^B$ , respectively, and the firms bid for the final rep of ability  $\sigma_3$ . The condition under which  $W$  will hire  $\sigma_3$  is given in Figure 3. The LHS of this inequality represents the marginal impact on  $W$ 's profit if it hires  $\sigma_3$  as compared with the case in which  $B$  hires her. Thus, this is the maximum  $W$  would be willing to pay to hire her (and keep  $B$  from hiring her). The RHS is the analogous quantity for  $B$ . We can decompose this marginal impact into two effects, a direct one and an indirect one. As shown in Figure 3, the direct effect of hiring  $\sigma_3$  has two components. First, the firm that hires her will have an

additional salesperson on the street who, given that the price is 1, will yield sales to all the firm's matching customers in the limit. This is the first term. The second term represents the direct impact on nonmatching customers: Salesperson  $\sigma_3$  will yield sales to the proportion  $\lambda$  customers that shop only  $W$  as well as the half of the  $1 - \lambda$  customers that shop both firms but are not "served" by  $B$  because it is understaffed. Finally, the third term captures the "indirect effect" of hiring the additional rep: She will *improve the effectiveness of the other salesperson* by reducing the level of competition. Specifically, if  $W$  hires two salespeople, then half the time<sup>17</sup> each of  $W$ 's salespeople will face no competition for the customer that shops  $B$  and  $W$ . By taking a salesperson, any salesperson, off of the market the firm lessens the competition its other reps will face in this undersupply context. One can see from inspection that the better firm always experiences a higher direct effect than the worse because it has more matching customers. This advantage is increasing in  $\phi$ . Less obvious is the fact that the better firm's advantage in the direct effect is *decreasing* in  $\sigma_3$ . This is because, when the product is very simple to sell, all salespeople can sell a matching customer but salespeople differ in their ability to sell to nonmatching customers. Because  $W$  has more nonmatching customers, an increase in the ability of the salesperson will have more of a beneficial effect on the worse firm. The indirect effect depends both on  $\phi$  and on the abilities of the reps already hired,  $\sigma^B$  and  $\sigma^W$ , but not on  $\sigma_3$ . It is increasing in  $\phi$  because when the product is simple, this effect is mainly experienced in terms of the salesperson's impact on selling to the firm's nonmatching customers and  $W$  has more of them. It depends on  $\sigma^B$  and  $\sigma^W$  because the indirect effect represents the fact that the other rep faces less competition. Thus, the higher the skill level of the other rep, the higher the indirect effect. So, if  $\sigma^B$  is close to  $\sigma^W$ , then  $W$  experiences a higher indirect effect since it has more nonmatching customers. On

Figure 2 Hiring Regions in Corollary 1—Simple Products



<sup>17</sup> Firm  $B$  has one salesperson and thus can only serve one of the two customers. Our indistinguishability assumption leads to the random outcome in which the probability that a given customer that arrives at  $W$  will have seen (or will see) a salesperson at  $B$  is  $1/2$ .



Figure 3 Necessary Condition for  $W$  to Hire the “Third Rep”

$$\begin{array}{c}
 \text{Marginal Impact on } W\text{'s Profit} \qquad \qquad \text{Marginal Impact on } B\text{'s Profit} \\
 \underbrace{\left( (1-\phi) + \phi p_{\sigma_3}^n \frac{(1+\lambda)}{2} \right)}_{\text{“Direct Effect”}} + \underbrace{\phi p_{\sigma^W}^n \frac{(1-\lambda)}{2}}_{\text{“Indirect Effect”}} > \phi + (1-\phi) p_{\sigma_3}^n \frac{(1+\lambda)}{2} + (1-\phi) p_{\sigma^B}^n \frac{(1-\lambda)}{2}
 \end{array}$$

the other hand, for a given  $\phi$ , if  $\sigma^B$  is much higher (lower) than  $\sigma^W$ ,  $B$  will experience a higher (lower) indirect effect. The indirect effect does not depend on  $\sigma_3$  because essentially this impact is seen in the fact that she “frees up” the other rep. There is, by assumption, no difference between a high-skill and a low-skill rep’s ability to do so. The ultimate hiring equilibrium will depend in large part on the relative levels of the direct and indirect effect. When  $\phi$  is relatively low—when the firms are similar in terms of their market appeal—the direct effect for the two firms is very similar because the firms have very similar proportions of matching and nonmatching customers. Thus, the indirect effect will dominate in this region. This has the important implication that the firm that hires the *most* salespeople will also hire the *highest-skilled* salespeople. We see this in both of the equilibria when  $\phi$  is low: I and IV.

While we have analyzed why and when the firm that hires the most will also hire the most highly skilled, we have not discussed which firm will do so. As shown in Figure 2, sometimes it will be the better firm (I) while other times it will be the worse (IV). A necessary condition for Region IV is:

$$\begin{aligned}
 0 &< \left[ (p(\bar{\sigma}, \delta^n) + p(\underline{\sigma}, \delta^n)) \frac{1+\lambda}{2} - p(\underline{\sigma}, \delta^n) \lambda \right] - 1 \\
 &\equiv Z.
 \end{aligned} \tag{4.4}$$

The first term in brackets is the relative impact of employing the two highest-skilled reps—those with abilities  $\bar{\sigma}$  and  $\underline{\sigma}$ —on nonmatching customers as compared with just the lowest-skilled rep. This relative benefit is increasing in  $\lambda$  because it captures the degree of competition the reps face. Higher  $\lambda$  means that they face fewer situations in which the customer

will receive a proposed solution from the other firm as well. The second term on the RHS is the relative impact on matching customers. This approaches one in the limit because the probability that the matching customer is sold by any salesperson approaches one. That is, having two reps sell to matching customers instead of one delivers in the limit additional profits of one. When (4.4) is positive, which occurs when  $\lambda$  is high enough, then the relative impact is higher on nonmatching customers. In this case, if either firm will hire the most *and* the best it will be  $W$ . When  $\lambda$  is low, so that the relative impact is higher on the *matching* customer, if either firm will hire the most and the best it will be  $B$ .

Combining these two analyses—in (4.4) and Figure 3—we find that when the firms are relatively similar (low values of  $\phi$ ), then the firm that hires the most salespeople will also hire the most highly skilled salespeople. So, one firm will hire  $\bar{\sigma}$  and  $\underline{\sigma}$  while the other just hires  $\underline{\sigma}$ . When competition is not too intense ( $\lambda$  is high), the relative benefit of hiring  $\bar{\sigma}$  and  $\underline{\sigma}$  as compared with just  $\underline{\sigma}$  is experienced more on the nonmatching customers. The matching customers are not subject to competition because  $\rho > 0$ ; the firm always sells to a matching customer if the rep delivers him a valuable solution. On the other hand, even a valuable solution may not be enough to win a nonmatching customer. It depends on what the other firm’s success was. As competition becomes less intense, the impact of hiring policies on nonmatching customers becomes more prominent. Thus, the worse firm—which has more nonmatching customers—hires the most *and* the most highly skilled reps. When competition is intense, the opposite is true.

Now, when  $\phi$  is high—when the firms are relatively asymmetric—we find that the better firm *always* hires

the most salespeople. This is seen in Regions I, II, and III. As discussed above, this is driven by the fact that in this region the direct effect of hiring the additional rep dominates the outcome and that the better firm always has a higher direct effect. Given that the *how many* question is thus determined in this region, we can confine ourselves to the *which* question as we have throughout the rest of the paper. The difference here, of course, is that the better firm will hire two salespeople and the worse firm will hire only one. Assume that firm *B* has hired a rep of ability  $\sigma^B$ , leaving two reps of ability  $\sigma_2$  and  $\sigma_3$ , where  $\sigma_2 > \sigma_3$ . Firm *W* will hire  $\sigma_2$  if her marginal impact on *W*'s profit is higher than it would be on firm *B*. This condition is shown in Figure 4. On a marginal basis, the firms experience only direct effects. There are no marginal indirect effects because the better firm will experience the exact same indirect effect regardless of the ability of the second rep hired. Recall from above that the magnitude of the indirect effect depends on the ability of the other reps, not the marginal rep. In Figure 4, we have decomposed the marginal impact into customers that shop only one firm and customers that shop both firms. The latter term reflects the sales generated from nonmatching customers that shop both firms. In this simple context, the firm is always able to deliver a valuable solution to its matching customers. Customers that shop both firms will only buy from their nonmatching firm when they are not offered a solution at their matching firm. Thus, the worse firm is not able to generate such sales. The inequality in Figure 4 reduces to  $\Phi > (1 + \lambda)/2\lambda$ , which holds for high  $\phi$  and  $\lambda$ . Consistent with our intuition to this

point, the marginal impact of selling ability is highest on the nonmatching customer when products are relatively simple. The worse firm always has more nonmatching customers and this disparity grows in  $\phi$ . As  $\lambda$  grows, competition becomes less intense. This benefits the worse firm more than the better firm in this context because the better firm's larger sales staff allows it to generate sales from its nonmatching customers that shop both firms, but the worse firm cannot.

So, in summary, we have found that relaxing the restriction that the firms will hire the same number of salespeople has added an important layer of complexity on top of our existing characterization of the problem. In particular, we have been able to investigate the interplay of two important questions facing the firm in its sales force planning: Which salespeople should we hire and how many salespeople should we hire? We have found that, in equilibrium, the firm that hires the most does not necessarily hire the most highly skilled. This is evidence that the forces behind the *how many* decisions are not necessarily the same as those behind the *which* decision. Specifically, the former was driven mainly by the extent to which the impact of having a larger sales force would be more direct or indirect. That is, would an additional rep have more impact on the firm's profit by bringing additional sales or by lessening the competition faced by the other rep? When the former was the case, the better firm always hired the most reps because any rep of any ability will deliver more incremental sales of the better firm's product as compared to the worse. Even when this is the case, however, the better firm

Figure 4 Necessary Condition for *W* to Hire the Second-Best Rep When  $\phi$  Is High

$$\begin{array}{c}
 \begin{array}{cc}
 \text{Marginal Impact on W's Profit} & \text{Marginal Impact on B's Profit}
 \end{array} \\
 \begin{array}{c}
 \underbrace{\hspace{10em}} \\
 \phi(p_{\sigma_2}^n - p_{\sigma_3}^n)\lambda > (1-\phi)(p_{\sigma_2}^n - p_{\sigma_3}^n)\lambda + (1-\phi)(p_{\sigma_2}^n - p_{\sigma_3}^n)\frac{1-\lambda}{2} \\
 \underbrace{\hspace{10em}}
 \end{array} \\
 \begin{array}{ccc}
 \underbrace{\hspace{3em}} & \underbrace{\hspace{3em}} & \underbrace{\hspace{3em}} \\
 \text{Customers that Shop Only W} & \text{Customers that Shop Only B} & \text{Customers that Shop Both Firms}
 \end{array}
 \end{array}$$

may or may not hire the most highly skilled salespeople. When  $\phi$  and  $\lambda$  are both high—when the firms are relatively asymmetric and competition is not very intense—the worse firm's single rep will be the best rep available.

## 5. Persuasive Selling

So far, we have assumed that the salesperson adds actual, tangible value to the product by solving a customer's problem. However, in many fields the selling task is more a *persuasive* one than a value-adding one. Thus, as a final inquiry, we will consider a context in which the rep's job is to select pieces of information to reveal to the customer to convince him that the product is "right" for him. Besides prescribing the optimal allocation of persuasive selling skill, this model will also offer insight into why a rational customer should believe a salesperson in this context even though he knows she is biased.

### 5.1. Model

For parsimony, we will return to the single-product monopolist model which we describe in §§3.1 and 3.2. We will make two important changes to that model. First, a customer of type  $t$  is now assumed to experience utility from product type  $T$  given by:

$$U^t \equiv 1(\mu = m) - P^T. \quad (5.1)$$

The customer's utility is now a function of whether or not he matches the product. That is, the underlying value is fixed but unobservable. Because he does not observe either  $t$  or  $T$ , his prior belief is that  $\mu = m$  is  $\frac{1}{2}$ . The match captures whether the product is "right" for him. If it is not, he gets no utility from it. Comparing this to (3.1), we see that the role of  $\mu$  was less important to the customer when the salesperson's job was to create value. This is because the customer saw what he was getting before he bought it. Here, he is not certain about the underlying value of the product unless he buys it.

The second change we will make to the model is in the nature of the action being chosen by the salesperson. Here, her job is to select the right sales pitch to convince the customer that he matches the product. Each of the salesperson's  $K$  pitches consists solely

of information: product demonstrations, "test drives," testimonials, references, specifications, etc. The key aspect of a pitch as compared with a solution is that it adds no value except as a *signal* of a match. Each pitch  $k$  is evaluated by the customer as being either convincing or not:

$$a_k = \begin{cases} 1 & \text{pitch } k \text{ is convincing for} \\ & \text{the customer,} \\ 0 & \text{pitch } k \text{ is not convincing} \\ & \text{for the customer.} \end{cases} \quad (5.2)$$

As above, we define  $\delta^m$  and  $\delta^n$  as the complexity of the task of convincing matching and nonmatching customers, respectively. We assume that  $\delta^m < \delta^n$ : It is easier to convince a customer that he matches if, in fact, he matches. To see why this is different from value creation, compare the customer's expected utility, conditional on an effective action, in each case:

$$\text{Value-Added Selling : } E[U \mid a = 1] = 1 - P^T,$$

$$\text{Persuasive Selling : } E[U \mid a = 1] = \Pr[\mu = m \mid a = 1] - P^T. \quad (5.3)$$

As in the value-added case, the customer observes perfectly whether  $a = 1$ . However, he can no longer deduce perfectly the *implication* of the effective action. That is because the sales pitch, unlike the solution, is only a *signal* of value, he must make an inference about the utility he will experience in the latter case. This has an important implication for the value of a salesperson: The customer will optimally discount the value of a convincing pitch as a function of the inferred ability of the salesperson. When a customer receives a convincing pitch, he knows it is part signal and part noise. The signal derives from the fact that  $\delta^m < \delta^n$ , and so  $\Pr[a = 1 \mid \mu = m] > \Pr[a = 1 \mid \mu = n]$ . The noise, on the other hand, derives from the fact that  $\Pr[a = 1 \mid \bar{\sigma}] > \Pr[a = 1 \mid \sigma]$ . This means that he is more likely to receive a convincing pitch the higher the salesperson's skill, *regardless of the product's value*.

### 5.2. Beliefs

Unlike in previous sections where pricing was relatively nonstrategic, here we must consider the fact that price might signal the firm's type. If the firms

price differently in equilibrium, then the customer infers  $T$  (though not  $\mu$ ) perfectly. When  $\phi$  is high, this may hurt  $W$  (and benefit  $B$ ) because the customer would begin with a very pessimistic prior belief. For this reason, we will use the Bayesian Nash Equilibrium concept (see Fudenberg and Tirole 1991), which requires that we define off-path beliefs. Let  $P^* \equiv \{P^{B*}, P^{W*}\}$  be the pair of equilibrium prices and  $P'$  the price that the customer observes. If  $P' \notin P^*$ , we need additional assumptions about beliefs which we restrict to the following structure:

$$\begin{aligned}\omega^h &\equiv \Pr[T = B \mid P' > \text{Max}\{P^{B*}, P^{W*}\}], \\ \omega^m &\equiv \Pr[T = B \mid \text{Max}\{P^{B*}, P^{W*}\} \\ &> P' > \text{Min}\{P^{B*}, P^{W*}\}], \\ \omega^l &\equiv \Pr[T = B \mid \text{Min}\{P^{B*}, P^{W*}\} > P'].\end{aligned}\quad (5.4)$$

This is a rather coarse structure for off-path beliefs. We could have specified the more general distribution  $\omega(P')$ . Since the specification in (5.4) is a special case of this general distribution, we are biasing against the existence of any equilibrium that must be supported by specific off-path beliefs. Thus, we believe that this represents a reasonably conservative restriction.

### 5.3. Core Results

As is the case with most models of this type, our model gives rise to multiple equilibria. Since our goal is to understand the relative valuation of high-skill salespeople across firms of different types, we will focus on cases in which (a) the firms are very different ( $\phi$  is high) and (b) selling skill (the  $\sigma$ 's) is high enough that at least one of the firms optimally hires a high-skill salesperson. As we will show below, these two restrictions limit our analysis to two classes of equilibria. We will briefly discuss the other equilibria in §5.4. The following proposition shows that the intuition that we developed above in a context of value-creating salespeople applies to purely value-signaling salespeople as well.

**PROPOSITION 5.** (a) *Given the following conditions, there exists a range of prices  $[\bar{P}, \underline{P}]$  such that, for each  $P \in [\bar{P}, \underline{P}]$ , there exist off-path beliefs  $\{\omega^h, \omega^m, \omega^l\}$  to*

*support an equilibrium in which  $h^B = \underline{\sigma}$ ,  $h^W = \bar{\sigma}$  and  $P^B = P^W = P$ :*

- (i)  $\phi$  is high enough,
- (ii)  $\gamma \geq (\Phi \cdot R + 1)/(R + \Phi)$ .

(b) *Given the following conditions, there exists a range of prices  $[\bar{P}, \underline{P}]$  such that for each  $P \in [\bar{P}, \underline{P}]$ , there exist off-path beliefs  $\{\omega^h, \omega^m, \omega^l\}$  to support an equilibrium in which  $h^B = \bar{\sigma}$ ,  $h^W = \underline{\sigma}$ , and  $P^B = P^W = P$ :*

- (i)  $\phi$  is high enough,
- (ii)  $\gamma \geq (R + \Phi)/(\Phi \cdot R + 1)$ .

First, note that the conditions for this equilibrium are stronger than for Proposition 1 because condition (ii) is identical to the condition for Proposition 1 and condition (i) is new. This is because the persuasive-selling equilibrium is driven not only by the difference in the selling ability, but also by the willingness of the two firms to charge prices that support this equilibrium. As with Proposition 1, the main intuition behind these equilibria lies in the conditions (ii). In equilibrium (a), since  $\Phi > 1$ , we need for  $R$  to be small. By Lemma 2, Proposition 5(a) says that when the task is simple enough, then the hiring policy is  $h^B = \sigma$  and  $h^W = \bar{\sigma}$ . Conversely, when the task is complex enough, Proposition 5(b) says that the hiring policy is  $h^B = \bar{\sigma}$  and  $h^W = \underline{\sigma}$ . It is also interesting to note that, in equilibrium, the firms place a premium on the high-skill rep even though the rational customer optimally discounts the information she provides. Put differently, even though the customer *knows* that the salesperson is biased, he puts some stock in the fact that the information that she delivered was convincing. This is due to the stochastic dominance of the outcome distribution for a matching customer as compared with a non-matching customer. It is straightforward to show that, for any  $\phi$ , it must always be true that  $E[U \mid a = 1] > E[U \mid a = 0]$ . Thus, the salesperson's convincing pitch is informative (and her skills therefore valuable) precisely because it is easier for her to convince us that we match a product if it is true than if it is false. We believe that the idea that we are *rationally* persuaded by a biased party like a salesperson is an interesting one and is perhaps worthy of future elaboration.

### 5.4. Other Equilibria

As noted above, this model also gives rise to other equilibria. We will briefly discuss the qualitative

nature of these equilibria and refer the interested reader to the Technical Appendix for a more thorough treatment. This section concludes with Proposition 6, showing that the equilibria in Proposition 5 are unique under certain conditions.

**Pooling Equilibria.** There exist two pooling equilibria in this model in which the firms hire the same salesperson, either high- or low-skill. In the former, the firm types pool on a high price and depend on the rep to deliver a convincing pitch. This equilibrium only exists when  $\gamma$  is small, and therefore does not overlap with either of the equilibria in Proposition 5. They may also pool on a low price, sell their product all of the time, and both hire the low-skill rep. It is necessary for this latter equilibrium that the salespeople not be “too valuable.” If the  $\sigma$ ’s get too high, then the firms are not satisfied with low-skill salespeople and the equilibrium disappears.

**Separating Equilibria.** Here again, we have two additional equilibria. In one, the better firm hires the high-skill salesperson and prices high while the worse firm hires the low-skill salesperson and prices low. This captures an important aspect of selling skill, which is that *ceteris paribus* higher-priced products ought to be able to attract higher-skill salespeople. While this equilibrium is certainly compelling, it vanishes as the firms get different enough (as  $\phi \rightarrow 1$ ). In this region, the worse firm is no longer willing to signal its lower quality. Another separating equilibrium in this model is one in which the firms both hire low-skill salespeople. Again, firm *B* prices high while firm *W* prices low. Since  $\gamma < 1$ , this can only occur when *W* has a higher marginal return on selling ability than *B*, which is when the products are simple. Following the same logic as above, this equilibrium disappears when the firms are different enough.

We conclude with the following proposition that gives conditions for the uniqueness of the equilibria in Proposition 5. This requires only one qualitatively new condition: that the low-skill salesperson is not “too low-skilled.” The other conditions, that  $\phi$  and  $\gamma$  be high enough and that the tasks were either simple or complex enough, were required for the original equilibria (though the specific cut-offs for existence and uniqueness are not generally the same).

**PROPOSITION 6.** Assume  $\phi$ ,  $\gamma$ , and  $\sigma$  are all high enough; then the equilibrium in Proposition 5(a) (5(b)) is unique if the task is simple (complex) enough.

## 6. Conclusion and Future Directions

This paper deals with a fundamental issue facing sales-driven firms: How much is selling ability worth? The principal result is that, when the task is relatively complex, selling ability has a bigger impact on products that are easier to sell. When the task is relatively simple, selling ability has a bigger impact on products that are harder to sell. This result has many important implications for the firm. The firm selling two complex products should allocate its highest-skill salespeople to its easier-to-sell product. If the environment was relatively simple, the firm should allocate this rep to its harder-to-sell product. Not surprisingly, the firm’s allocation of training resources should be similar: more on the harder-to-sell (easier-to-sell) product if the tasks are relatively easy (difficult). We have shown that this insight also explains the movement between firms in a competitive environment. So, in a complex selling environment, better firms should poach highly skilled reps from worse firms. The opposite should be true in simple contexts. Further, the insight holds under different assumptions about the nature of the selling task. In §§3 and 4, we looked at selling as a value-adding function and in §5 we looked at selling as a persuasive, or value-signaling, function. Finally, we have generated interesting insight into the difference between the questions “which salespeople should we hire?” and “how many salespeople should we hire?” We found that the firm that hires the most salespeople does not always hire the best salespeople. Specifically, when the firms are sufficiently different in terms of their market appeal, the better firm always hires more salespeople. However, when the product market competition is not too intense, the worse firm hires the highest-skilled rep. This is because the returns to having a salesperson are higher for the better firm but the returns to having a high level of *selling ability* were higher for the worse firm.

We believe that this paper makes several significant and unique contributions to the literature. First, we have addressed a problem that, to our knowledge, has not been addressed explicitly: the value of salespeople. This is a critically important problem. Firms spend billions of dollars annually on their sales forces. It is essential that we understand the nature of the impact of this resource, where the impact is likely to be greatest, and how the impact—and therefore value—differs across contexts. Moreover, we believe that we have shed some light on the complicated pattern of movement of salespeople across firms. Up to one in five salespeople leaves a firm annually. It is essential that we understand the forces behind this movement and determine whether and when it is preventable, or inevitable.

Second, we have addressed the problem by developing a unique approach to modeling selling phenomena. It differs from other approaches to sales force modeling in that we have presented a model of the selling process at a microlevel. We began in §2 with a specification of *what selling is*. Other than Wernerfelt (1994), we know of no other attempt to capture the complex nature of the selling function in an analytical environment at such a structural level. The motivating factor for us to do so was that it allowed us to look at the interaction between selling ability and task complexity. Without modeling selling explicitly, we would not have been able to investigate the complicated relationship between how good one is and how hard the task is. We believe that this has the potential to generate significant future research in many domains. We have modeled selling skill as the ability to select the best of several alternatives: High-skill reps are better at selecting an effective action. This approach may be applicable to a number of other areas. For example, CEOs are paid to make good decisions, which essentially amounts to their ability to select the best from among a set of possible solutions. Our analysis suggests that when the industry is characterized by relatively “difficult” problems (those that are harder to solve), then an increase in a CEO’s ability will have a bigger impact on the easier problems than on the more difficult problems. Thus, those firms with relatively easier problems will place a higher value

on a CEO candidate. We might speculate that “better” firms have more problems of the “simple” type. Thus, we would predict that in complex industries, the most highly skilled CEOs work for better firms, all else equal. We would predict the opposite when the industry is characterized by relatively simple problems. This line of analysis might also be extended into analogous areas in which “skill” translates into one’s success at selecting the “best of K.” Moreover, while we tend to think of “skill” as an *individual-level* construct, we might also interpret our model as capturing *institution-level* skill as well. For example, advertising agencies perform a function similar to personal selling: They select the best way to design an ad. If (a) highly skilled agencies are paid more than lesser-skilled agencies and (b) it is harder to come up with an ad that “works” for some clients than others, then our model might suggest that, in complex markets, the easier products engage the high-skill agencies while the opposite is true in simple markets.

The paper generates several empirically testable propositions. Most notably, *ceteris paribus*, when the industry is characterized by a complex selling task we should observe the better firms employing more highly skilled reps. When the task is relatively simple, they should employ less-skilled reps. The challenges of such a study are not trivial. For example, while one would be tempted to focus on a firm or an industry to hold as many covariates constant as possible, this approach would most likely not ensure sufficient variance on the industry complexity measure to test the propositions. Still, a well-designed survey into this question would be an important next step in this research. One would imagine that the complexity construct could be operationalized as a multi-item measure comprised of questions about the diversity of customer needs, the range of possible solutions, and the breadth of uses of their products. The ability construct might be captured directly, inferred, or measured. For example, we might infer that a salesperson hired away from a competitor for more money than he was making is a “high-skill” rep, while one moving between firms for the same or less money is not. On the other hand, measurement might again be accomplished via a multi-item construct comprised of questions about years of experience, offers received

by other firms, industry recognition, customer satisfaction, and awards received.

Finally, we believe that the results we have presented offer significant insight for the manager. They provide guidance in the allocation of salespeople across product lines as well as the determination of the potential value of salespeople from other firms. While it is tempting for a manager to argue that *all* of her salespeople should be the “best,” this cannot always be the case. The firm operates under significant constraints on both its financial and human resources. Thus, managers must make tough decisions about where the best reps should reside in the organization, which reps in which to invest, and which reps to hire away from competitors. Our analysis points out that some products are so good that they do not *need* the best salespeople. Others are so bad that they cannot *afford* the best salespeople. Of course, it is likely that internal political forces—e.g., the impact the product sells on one’s image or reputation—may interact with the underlying forces of valuation that we model. At the very least, we hope that the paper spurs managers to think through the important, but perhaps seldom asked, question, “is this rep *really* more valuable to us than she would be to another company?” If the answer is “no,” then the firm is necessarily overpaying. Of course, the question is not an easy one to answer. We show that it requires an understanding of the industry as well as of the relative difficulty of the firm’s products as compared with the competition.

## Acknowledgments

The author is grateful to Leonard Lee, Dina Mayzlin, Duncan Simester, Olivier Toubia, Wei Wu, and especially Birger Wernerfelt for comments on earlier drafts of this paper and to Brian Ratchford, Steve Shugan, the area editor, and three anonymous referees for their many helpful suggestions. He also thanks seminar participants at Berkeley, Carnegie Mellon, Chicago, Florida, Harvard, MIT, Michigan, Northwestern, Stanford, UCLA, USC, Washington, Wharton, and Yale.

## Appendix A: Proofs

Note: We will denote  $\bar{S} \equiv \bar{\sigma}/(1 - \bar{\sigma})$ ,  $\underline{S} \equiv \underline{\sigma}/(1 - \underline{\sigma})$ ,  $D^\mu \equiv \delta^\mu/(1 - \delta^\mu)$ , and  $\Omega^j \equiv \omega^j/(1 - \omega^j)$ . Further, we will define  $\bar{p}^\mu \equiv p(\bar{\sigma}, \delta^\mu)$ ,  $\underline{p}^\mu \equiv p(\underline{\sigma}, \delta^\mu)$ ,  $\bar{p}^\mu \equiv p(\bar{\sigma}, \delta^\mu)$ , and  $\underline{p}^\mu \equiv p(\underline{\sigma}, \delta^\mu)$ .

PROOF OF LEMMA 1. The prior that  $a_k = 1$  for any  $k$  is  $1 - \delta$ . Since the agent optimally selects randomly from among those solutions

with  $\hat{a}_k = 1$ , the probability of selecting a valuable solution is simply the posterior probability that this solution is, in fact, valuable:  $\Pr[a_k = 1 | \hat{a}_k = 1]$ . This is found by application of Bayes’ Rule:

$$\begin{aligned} \Pr[a_k = 1 | \hat{a}_k = 1] &= \frac{\Pr[\hat{a}_k = 1 | a_k = 1] \cdot \Pr[a_k = 1]}{\Pr[\hat{a}_k = 1 | a_k = 1] \cdot \Pr[a_k = 1] + \Pr[\hat{a}_k = 1 | a_k = 0] \cdot \Pr[a_k = 0]} \\ &= \frac{\sigma(1 - \delta)}{\sigma(1 - \delta) + (1 - \sigma)\delta} = \left[1 + \frac{\delta}{1 - \delta} \cdot \frac{1 - \sigma}{\sigma}\right]^{-1}. \end{aligned}$$

By inspection,  $\partial p / \partial \delta < 0$  and  $\partial p / \partial \sigma > 0$ .  $\partial^2 p / \partial \delta^2 = K \cdot (3\sigma - 1 - 2\sigma^2) / (\sigma(2\delta - 1) - \delta)^3$ ,  $K > 0$ . The numerator is strictly positive on  $\sigma \in (\frac{1}{2}, 1)$ .  $\sigma(2\delta - 1) - \delta$  is strictly increasing in  $\delta$  and is negative at  $\delta = 1$ . Also,  $\partial^2 p / \partial \sigma^2 = 2\delta(2\delta^2 - 3\delta + 1) / (\delta(2\sigma - 1) - \sigma)^3$  where the denominator is strictly negative and the numerator is negative for  $\delta \in (\frac{1}{2}, 1)$ . Finally,  $\partial^2 p / \partial \delta \partial \sigma = (\delta - \sigma) / (\delta(2\sigma - 1) - \sigma)^3$  where, again, the denominator is negative and the numerator is positive (negative) for high (low)  $\delta$ .  $\square$

PROOF OF LEMMA 2. (a) First, letting  $\delta^i = \delta^j - e$ , note that  $R \rightarrow_{\delta^i \rightarrow 1} +\infty$  since  $\bar{p}^n$  and  $\underline{p}^n$  converge to 0 at the same rate and  $\bar{p}^m > \underline{p}^m$ . Further, we can show that

$$\left. \frac{\partial R}{\partial \delta^n} \right|_{\delta^m = \delta^n - e} = \frac{A[B/C - D]}{E},$$

where

$$A \equiv (\delta^n + \underline{\sigma} - 2\delta^n \underline{\sigma})^2 (\delta^n + \bar{\sigma} - 2\delta^n \bar{\sigma})^2,$$

$$B \equiv (\delta^n(1 - \delta^n) + e - 2\delta^n e + e^2)(\delta^n + \underline{\sigma} - 2\delta^n \underline{\sigma} + e(2\underline{\sigma} - 1))$$

$$\cdot (\delta^n + \bar{\sigma} - 2\delta^n \bar{\sigma} + e(2\bar{\sigma} - 1)) \cdot (\underline{\sigma} \bar{\sigma}(1 - 2\delta^n) + \delta^{n2}(\underline{\sigma} + \bar{\sigma} - 1)),$$

$$C \equiv (\delta^n + \underline{\sigma} - 2\delta^n \underline{\sigma})^2 (\delta^n + \bar{\sigma} - 2\delta^n \bar{\sigma})^2,$$

$$D \equiv (\delta^{n2} - \delta^n) \cdot [\underline{\sigma} \bar{\sigma}(1 + 2e) + (\delta^{n2} + e^2)(\underline{\sigma}_b + \bar{\sigma} - 1) - 2\delta^n(\underline{\sigma} \bar{\sigma} + e(\underline{\sigma} + \bar{\sigma} - 1))],$$

and

$$E \equiv (\delta^n - 1)^2 \delta^{n2} \cdot (\delta^n + \underline{\sigma} - 2\delta^n \underline{\sigma} + e(2\underline{\sigma} - 1))^2 \cdot (\delta^n + \bar{\sigma} - 2\delta^n \bar{\sigma} + e(2\bar{\sigma} - 1))^2.$$

Since  $E \rightarrow_{\delta^n \rightarrow 1} 0^+$ , we will focus on identifying the sign of the numerator. It is clear that  $A \rightarrow K_1 > 0$ ,  $C \rightarrow K_2 > 0$ , and that  $D \rightarrow 0$ , so we need to identify the sign of  $B$  which approaches  $K \cdot [1 - \underline{\sigma} + e(2\underline{\sigma} - 1)] \cdot [1 - \bar{\sigma} + e(2\bar{\sigma} - 1)]$ ,  $K > 0$ . Since  $\sigma \in (\frac{1}{2}, 1)$ ,  $B \rightarrow K_3 > 0$  and we are done. By the same argument, when  $\delta^n = \delta^m + f$ ,  $R \rightarrow_{\delta^m \rightarrow 0} 0$ . We can, in a similar fashion, show that  $\partial R / \partial \delta^m \rightarrow_{\delta^m \rightarrow 0} K_4$ , such that  $\text{sgn}[K_4] = \text{sgn}[(f(2\sigma - 1) - \sigma)(f(2\bar{\sigma} - 1) - \bar{\sigma})]$ . So, as long as  $\text{sgn}[f(2\sigma - 1) - \sigma] = \text{sgn}[f(2\bar{\sigma} - 1) - \bar{\sigma}]$ , we are all done. Assuming that this does not hold, this would imply, for  $f < \frac{1}{2}$ , that  $f(2\sigma - 1) - \sigma > 0 > f(2\bar{\sigma} - 1) - \bar{\sigma}$ , which implies that  $f > \underline{\sigma} / (2\sigma - 1)$ , which implies that  $\frac{1}{2} > \underline{\sigma} / (2\sigma - 1) \Leftrightarrow 2\sigma - 1 > 2\underline{\sigma}$ , which is impossible.  $\square$

LEMMA 4. Let  $e \equiv \delta^n - \delta^m$ . There exists a single  $\hat{\delta}^m \in (0, 1 - e)$  such that for all  $\delta^m > (<) \hat{\delta}^m$

$$\phi(p_g^m - p_b^m) + (1 - \phi)(p_g^n - p_b^n) > (<)(1 - \phi)(p_g^m - p_b^m) + \phi(p_g^n - p_b^n).$$

PROOF OF LEMMA 4. Rearranging the equality yields  $(p_g^m - p_b^m) > (<) (p_g^n - p_b^n)$ . Straightforward application of the quadratic formula shows that this equation has two roots in  $\delta^m$ :

$$\delta^{m+} \equiv \frac{2\bar{\sigma}\underline{\sigma} - e(\bar{\sigma} + \underline{\sigma} - 1) + \sqrt{4(e-1)\bar{\sigma}\underline{\sigma}(\bar{\sigma} + \underline{\sigma} - 1) + [2\bar{\sigma}\underline{\sigma} - e(\bar{\sigma} + \underline{\sigma} - 1)]^2}}{2(\bar{\sigma} + \underline{\sigma} - 1)},$$

$$\delta^{m-} \equiv \frac{2\bar{\sigma}\underline{\sigma} - e(\bar{\sigma} + \underline{\sigma} - 1) - \sqrt{4(e-1)\bar{\sigma}\underline{\sigma}(\bar{\sigma} + \underline{\sigma} - 1) + [2\bar{\sigma}\underline{\sigma} - e(\bar{\sigma} + \underline{\sigma} - 1)]^2}}{2(\bar{\sigma} + \underline{\sigma} - 1)}.$$

Assume that  $\delta^{m+}$  takes values less than 1. This implies that, at these values,

$$\sqrt{4(e-1)\bar{\sigma}\underline{\sigma}(\bar{\sigma} + \underline{\sigma} - 1) + [2\bar{\sigma}\underline{\sigma} - e(\bar{\sigma} + \underline{\sigma} - 1)]^2} < 2(\bar{\sigma} + \underline{\sigma} - 1) - [2\bar{\sigma}\underline{\sigma} - e(\bar{\sigma} + \underline{\sigma} - 1)].$$

If the RHS is negative, we have a contradiction. If it is positive, we can square both sides and maintain the inequality. Doing so, and rearranging yields:

$$4(e+1)(\bar{\sigma}-1)(\underline{\sigma}-1)(\bar{\sigma}+\underline{\sigma}-1) < 0,$$

which is a contradiction since  $\bar{\sigma} + \underline{\sigma} - 1 > 0$  by virtue of the fact that  $\bar{\sigma}, \underline{\sigma} > \frac{1}{2}$ . So, since we constrain the  $\delta$ 's to be less than one, only  $\delta^{m-}$  is relevant. Moreover, by inspection, since  $4(e-1)\bar{\sigma}\underline{\sigma}(\bar{\sigma} + \underline{\sigma} - 1) < 0$ ,  $\delta^{m-}$  is positive. Since  $(p_g^n - p_b^n) \rightarrow_{\delta^m \rightarrow 1-\epsilon} 0$  and  $(p_g^m - p_b^m) \rightarrow_{\delta^m \rightarrow 0} 0$ , we are done.  $\square$

PROOF OF PROPOSITION 1. See Lemma 4 above. As shown there, when the task is complex (when  $\delta^m > \hat{\delta}^m$ ), firm  $B$  has a higher return on selling ability. In this case,  $h^B = \bar{\sigma}$  iff the improvement in firm  $B$ 's profit from hiring the high-skill rep is at least equal to the high-skill rep's reservation wage:  $\phi\bar{p}^m + (1-\phi)\bar{p}^n - V^{\min}(\bar{\sigma}) \geq \phi p^m + (1-\phi)p^n$ . By expanding the expression for  $V^{\min}(\bar{\sigma})$ , we get

$$\phi\bar{p}^m + (1-\phi)\bar{p}^n - \gamma[\phi(\bar{p}^m - p^m) + (1-\phi)(\bar{p}^n - p^n)] \geq \phi p^m + (1-\phi)p^n$$

which *always* holds for  $\gamma < 1$ . On the other hand,  $h^W = \underline{\sigma}$  iff the improvement in firm  $W$ 's profits from hiring the high-skill rep is no more than her reservation wage:

$$\phi\bar{p}^n + (1-\phi)\bar{p}^m - V^{\min}(\underline{\sigma}) \leq \phi p^n + (1-\phi)p^m$$

or

$$\phi\bar{p}^n + (1-\phi)\bar{p}^m - \gamma[\phi(\bar{p}^m - p^m) + (1-\phi)(\bar{p}^n - p^n)] \leq \phi p^n + (1-\phi)p^m,$$

which reduces to  $\gamma \geq (\Phi + R)/(\Phi R + 1)$ . By Lemma 2,  $R$  approaches  $\infty$  as the  $\delta$ 's get very high, and thus the RHS of this inequality approaches  $\Phi^{-1}$ . Since  $\Phi^{-1} < 1$ , this holds for high-enough  $\gamma$ . The low  $\delta$  case is proven analogously.  $\square$

PROOF OF PROPOSITION 2. The firm should assign the high-skill rep to product  $B$  when the incremental impact of her skill is higher there than on product  $W$ :

$$\phi(p_g^m - p_b^m) + (1-\phi)(p_g^n - p_b^n) > \phi(p_g^n - p_b^n) + (1-\phi)(p_g^m - p_b^m),$$

which rearranges to  $(p_g^m - p_b^m) > (p_g^n - p_b^n)$ . By Lemma 4, this occurs when  $\delta^m > \hat{\delta}^m$ , or when the task is complex enough. The opposite case is proven analogously.  $\square$

PROOF OF PROPOSITION 3. For it to be the case that the firm invests in  $B$ 's sales force only, it must be true that two conditions are met: (1) The incremental impact on  $B$ 's profit is higher than the investment amount:  $\phi p_g^m + (1-\phi)p_g^n - I \geq \phi p_b^m + (1-\phi)p_b^n$ , and (2) the incremental impact on  $W$ 's profit is lower than the investment amount:  $(1-\phi)p_g^m + \phi p_g^n - I \leq (1-\phi)p_b^m + \phi p_b^n$ . Combined, these imply that we need

$$\phi(p_g^m - p_b^m) + (1-\phi)(p_g^n - p_b^n) \geq I \geq (1-\phi)(p_g^m - p_b^m) + \phi(p_g^n - p_b^n),$$

which is condition (a). We find condition (b) in the same manner. A necessary condition for (a) ((b)) is

$$\phi(p_g^m - p_b^m) + (1-\phi)(p_g^n - p_b^n) \geq (\leq) (1-\phi)(p_g^m - p_b^m) + \phi(p_g^n - p_b^n),$$

which rearranges to  $(p_g^m - p_b^m) \geq (\leq) (p_g^n - p_b^n)$ . By Lemma 4, this only holds for  $\delta^m \geq \hat{\delta}^m$ .  $\square$

PROOF OF LEMMA 3. Abusing our notation somewhat, let  $p_{\sigma T}^T = \Pr[a_{k*}^T = 1 \mid \sigma^T \in \{g, b\}]$ . (a) High  $\delta$ 's:  $P^B = P^W = 1 + \rho$ . Equilibrium profits are  $\Pi^B = (1+\rho)\phi p_{\sigma B}^m$  and  $\Pi^W = (1+\rho)(1-\phi)p_{\sigma W}^m$ . For existence, first check  $B$ 's deviations. A deviation to  $P^B = 1$  yields a profit of

$$\lambda(\phi p_{\sigma B}^m + (1-\phi)p_{\sigma B}^n) + (1-\lambda)(\phi p_{\sigma B}^m + (1-\phi)p_{\sigma B}^n) = \phi p_{\sigma B}^m + (1-\phi)p_{\sigma B}^n$$

and is thus unprofitable if  $\rho > \Phi^{-1}(p_{\sigma B}^n/p_{\sigma B}^m)$ , which vanishes as  $\delta^n \rightarrow 1$ . No other deviations are threats. For  $W$ , a deviation to  $P^W = 1$  yields a profit of  $\phi p_{\sigma W}^m + (1-\phi)p_{\sigma W}^n$  and is thus unprofitable if  $\rho > \Phi(p_{\sigma W}^n/p_{\sigma W}^m)$ , which again vanishes as  $\delta^n \rightarrow 1$ . For uniqueness, first check  $P^B = P^W = 1$ . Profits in this equilibrium would be

$$\Pi^B = \lambda(\phi p_{\sigma B}^m + (1-\phi)p_{\sigma B}^n) + (1-\lambda)(\phi p_{\sigma B}^m + (1-\phi)p_{\sigma B}^n(1-p_{\sigma W}^m)),$$

and

$$\Pi^W = \lambda(\phi p_{\sigma W}^m + (1-\phi)p_{\sigma W}^n) + (1-\lambda)((1-\phi)p_{\sigma W}^m + \phi p_{\sigma W}^n(1-p_{\sigma B}^m)).$$

A deviation to  $P^B = 1 + \rho$  yields profits of  $(1+\rho)\phi p_{\sigma B}^m$  and would be profitable if

$$1 + \rho > \frac{\lambda(\phi p_{\sigma B}^m + (1-\phi)p_{\sigma B}^n) + (1-\lambda)(\phi p_{\sigma B}^m + (1-\phi)p_{\sigma B}^n(1-p_{\sigma W}^m))}{\phi p_{\sigma B}^m},$$

which approaches  $\rho > 0$  as  $\delta^n \rightarrow 1$ , so this deviation is always profitable at high  $\delta$ 's. Check  $P^B = 1 + \rho$ ,  $P^W = 1$ . Here, the firms "tie" when  $t = B$  and  $a_{k*}^B = a_{k*}^W = 1$ , so both firms will deviate to prices that are  $\epsilon$  lower. Analogously, we cannot have an equilibrium where  $P^W = 1 + \rho$ ,  $P^B = 1$ . (b) Low  $\delta$ 's:  $P^B = P^W = 1$ . Equilibrium profits are

$$\Pi^B = \lambda(\phi p_{\sigma B}^m + (1-\phi)p_{\sigma B}^n) + (1-\lambda)(\phi p_{\sigma B}^m + (1-\phi)p_{\sigma B}^n(1-p_{\sigma W}^m))$$

and

$$\Pi^W = \lambda(\phi p_{\sigma W}^m + (1-\phi)p_{\sigma W}^n) + (1-\lambda)((1-\phi)p_{\sigma W}^m + \phi p_{\sigma W}^n(1-p_{\sigma B}^m)).$$



First check  $B$ 's deviations. A deviation to  $P^B = 1 + \rho$  yields a profit of  $(1 + \rho)\phi p_{\sigma^B}^m$  and is thus unprofitable if  $\rho < \Phi^{-1}\lambda p_{\sigma^B}^n$  as  $\delta^m \rightarrow 0$ . So, the equilibrium holds for small  $\rho$ . No other deviations are threats. For  $W$ , a deviation to  $P^W = 1 + \rho$  yields a profit of  $(1 - \phi)(1 + \rho)p_{\sigma^W}^m$  and is thus unprofitable if  $\rho < \Phi\lambda p_{\sigma^W}^n$  as  $\delta^m \rightarrow 0$ . So, again, this exists for low  $\rho$ . For uniqueness, check  $P^B = P^W = 1 + \rho$ . As above,  $B$ 's deviation to  $P^B = 1$  is profitable if  $\rho < \Phi^{-1}(p_{\sigma^B}^n/p_{\sigma^B}^m)$ , which approaches  $\Phi^{-1}p_{\sigma^B}^n$  as  $\delta^m \rightarrow 0$ . So, this equilibrium disappears for small  $\rho$ . As above, the asymmetric equilibria fail because the firms have an incentive to undercut to avoid the "tie."  $\square$

**PROOF OF PROPOSITION 4.** We will begin with complex tasks. By Lemma 3, the prices are  $P^W = P^B = 1 + \rho$ . Equilibrium profits are:  $\Pi^B = (1 + \rho)\phi p_{\sigma^B}^m$  and  $\Pi^W = (1 + \rho)(1 - \phi)p_{\sigma^W}^m$ . Since the abilities are revealed perfectly to both parties, we can separate the problem and solve Period 1 first. First, assume there exists a SPE such that  $h_1^B = (\underline{\sigma}, \emptyset)$  and  $h_1^W = (\tilde{\sigma}, \emptyset)$  when the task is complex. By assumption, each bid must be strictly greater than the previous bid by at least  $\varepsilon$ . Thus, there are at most  $J$  bidding rounds. Let  $b^T(j)$  be firm  $T$ 's round  $j$  bid vector with elements  $b_\alpha^T(j)$ ,  $b_\beta^T(j)$ , and  $b_\sigma^T(j)$ . Further, let  $S^T(j) \equiv b_\alpha^T + b_\beta^T \cdot \Pr[\text{Sale} | T]$  which is the expected value of the contract bid by firm  $T$  in round  $j$ . We condition this on  $b_\sigma^T = \tilde{\sigma}$  because the optimal expected value of a contract offer to  $\underline{\sigma}$  is always 0. Note that, when bidding for  $\tilde{\sigma}$ , either firm could always drop out and offer  $\underline{\sigma}$  a zero expected wage contract. We will prove that this SPE cannot exist by showing that for sufficiently complex tasks (defined as high enough  $\delta$ 's) and sufficiently low  $\varepsilon$ , it is *always* optimal for firm  $B$  to outbid firm  $W$  for  $\tilde{\sigma}$ . Assume that firm  $B$  drops out of the bidding for  $\tilde{\sigma}$  in round  $j$ . This implies that  $b_\sigma^B(j) = \underline{\sigma}$  and thus the expected profit from hiring  $\tilde{\sigma}$  at  $\varepsilon$  more than the previous bid is smaller than the expected profit of offering a zero expected wage contract to  $\underline{\sigma}$ :

$$(1 + \rho)(\Pr[\text{Sale} | B, h^B = \tilde{\sigma}] - \Pr[\text{Sale} | B, h^B = \underline{\sigma}]) \leq S^W(j - 1) + \varepsilon,$$

or

$$(1 + \rho)\phi\{(1 - \nu)p_g^m + \nu p_b^m\} - [(1 - \nu)p_b^m + \nu p_g^m] \leq S^W(j - 1) + \varepsilon$$

$$\Leftrightarrow (1 + \rho)\phi(1 - 2\nu)(p_g^m - p_b^m) \leq S^W(j - 1) + \varepsilon.$$

Since  $W$  did not drop out in round  $j - 1$ , we have  $b_\sigma^W(j - 1) = \tilde{\sigma}$  and  $(1 + \rho)(1 - \phi)(1 - 2\nu)(p_g^m - p_b^m) \geq S^W(j - 1)$ . Together, these two conditions imply that  $(1 + \rho)(1 - \phi)(1 - 2\nu)(p_g^m - p_b^m) \geq (1 + \rho)\phi(1 - 2\nu)(p_g^m - p_b^m) - \varepsilon$  or

$$2\phi - 1 \leq \frac{\varepsilon}{(1 + \rho)(1 - 2\nu)(p_g^m - p_b^m)},$$

which fails for small  $\varepsilon$  since  $\phi > \frac{1}{2}$ . Now, for simple tasks. Here, equilibrium profits are:

$$\Pi^B = \lambda(\phi p_{\sigma^B}^m + (1 - \phi)p_{\sigma^B}^n) + (1 - \lambda)(\phi p_{\sigma^B}^m + (1 - \phi)p_{\sigma^B}^n(1 - p_{\sigma^W}^m))$$

and

$$\Pi^W = \lambda(\phi p_{\sigma^W}^m + (1 - \phi)p_{\sigma^W}^n) + (1 - \lambda)((1 - \phi)p_{\sigma^W}^m + \phi p_{\sigma^W}^n(1 - p_{\sigma^B}^m)).$$

We will assume there exists a SPE such that  $h_1^B = (\tilde{\sigma}, \emptyset)$  and  $h_1^W = (\underline{\sigma}, \emptyset)$  when the task is simple. Assume then that firm  $W$  drops out of the bidding for  $\tilde{\sigma}$  in round  $j$ . This implies that  $b_\sigma^W(j) = \underline{\sigma}$ , and thus the expected profit from hiring  $\tilde{\sigma}$  at  $\varepsilon$  more than the previous bid is smaller than the expected profit of offering a zero expected wage contract to  $\underline{\sigma}$ :

$$\Pr[\text{Sale} | W, h^W = \tilde{\sigma}] - \Pr[\text{Sale} | W, h^W = \underline{\sigma}] \leq S^B(j - 1) + \varepsilon$$

or

$$\{(1 - \phi)p_{\sigma^W}^m + \phi p_{\sigma^W}^n[\lambda + (1 - \lambda)(1 - p_{\sigma^W}^m)]\}$$

$$- \{(1 - \phi)p_{\sigma^W}^m + \phi p_{\sigma^W}^n[\lambda + (1 - \lambda)(1 - p_{\sigma^W}^m)]\} \leq S^B(j - 1) + \varepsilon.$$

Since  $B$  did not drop out in round  $j - 1$ , we have  $b_\sigma^B(j - 1) = \tilde{\sigma}$  and

$$\{\phi p_{\sigma^B}^m + (1 - \phi)p_{\sigma^B}^n[\lambda + (1 - \lambda)(1 - p_{\sigma^W}^m)]\}$$

$$- \{\phi p_{\sigma^W}^m + (1 - \phi)p_{\sigma^W}^n[\lambda + (1 - \lambda)(1 - p_{\sigma^W}^m)]\} \geq S^B(j - 1).$$

Together, these two conditions imply that

$$\{\phi p_{\sigma^B}^m + (1 - \phi)p_{\sigma^B}^n[\lambda + (1 - \lambda)(1 - p_{\sigma^W}^m)]\}$$

$$- \{\phi p_{\sigma^W}^m + (1 - \phi)p_{\sigma^W}^n[\lambda + (1 - \lambda)(1 - p_{\sigma^W}^m)]\}$$

$$\geq \{(1 - \phi)p_{\sigma^B}^m + \phi p_{\sigma^B}^n[\lambda + (1 - \lambda)(1 - p_{\sigma^W}^m)]\}$$

$$- \{(1 - \phi)p_{\sigma^W}^m + \phi p_{\sigma^W}^n[\lambda + (1 - \lambda)(1 - p_{\sigma^W}^m)]\} - \varepsilon.$$

If we take limits on both sides as  $\delta^m \rightarrow 0$ , we get

$$\{\phi + (1 - \phi)p_{\sigma^B}^n\lambda\} - \{\phi + (1 - \phi)p_{\sigma^W}^n\lambda\}$$

$$\geq \{(1 - \phi) + \phi p_{\sigma^B}^n\lambda\} - \{(1 - \phi) + \phi p_{\sigma^W}^n\lambda\} - \varepsilon$$

$$\Leftrightarrow \lambda(1 - \phi)(p_{\sigma^B}^n - p_{\sigma^W}^n) \geq \lambda\phi(p_{\sigma^B}^n - p_{\sigma^W}^n) - \varepsilon$$

$$\Leftrightarrow (2\phi - 1)[-\lambda(p_{\sigma^B}^n - p_{\sigma^W}^n)] \geq -\varepsilon$$

$$\Leftrightarrow (2\phi - 1)[\lambda(p_{\sigma^B}^n - p_{\sigma^W}^n)] \leq \varepsilon,$$

which fails for low  $\varepsilon$ . Thus, we have a contradiction which proves the proposition.

**PROOF OF PROPOSITION 5.** Assume that the conditions of equilibrium 5(a) are met and that the equilibrium in which  $h^B = \underline{\sigma}$  and  $h^W = \tilde{\sigma}$  and  $P^B = P^W = P^*$  does not exist. This equilibrium requires that  $E[U | a = 1] \geq P^* > E[U | a = 0]$  so that  $E[\Pi^{B*}] = P^* \cdot \Pr[a = 1 | h^B = \underline{\sigma}, B] - \Delta C$  and  $E[\Pi^{W*}] = P^* \cdot \{(1 - \gamma)\Pr[a = 1 | h^W = \tilde{\sigma}, W] + \gamma \cdot \Pr[a = 1 | h^W = \underline{\sigma}, W]\} - \Delta C$ . For this equilibrium *not* to exist, we must have a strictly profitable deviation by one of the types on  $P^T$ , on  $h^T$ , or on both. Begin with  $B$ 's deviation to  $P' > P^*$ . Since

$$E[U | a = 1, P' > P^*] = \left[1 + \frac{\Omega^h p_{\sigma^B}^n + \Phi \bar{p}^n}{\Omega^h \Phi p_{\sigma^B}^m + \bar{p}^m}\right]^{-1},$$

which is increasing in  $\Omega^h$ ,  $\Pi^{B'} > \Pi^{B*} \Rightarrow E[U | a = 1, P' > P^*] > P^*$ , which is prevented for low-enough  $\Omega^h$ . Since the  $E[U | a = 1,$

$P' > P^* \rightarrow \omega^h \rightarrow 0 = [1 + \Phi \bar{p}^n / \bar{p}^m]^{-1}$  the latter quantity is  $\underline{P}$ . Deviation to  $P^* > P' > E[U | a = 0]$  only decreases price, but not  $\Pr[a = 1]$ , so is strictly worse. Deviation to  $P' \leq E[U | a = 0]$  is profitable if

$$P^*[\phi p^m + (1 - \phi)\underline{p}^n] - \Delta C < \left[1 + \frac{\Omega^l(1 - \underline{p}^n) + \Phi(1 - \bar{p}^n)}{\Omega^l\Phi(1 - \underline{p}^m) + (1 - \bar{p}^m)}\right]^{-1},$$

which fails for  $\phi$  high enough and  $\alpha^l$  low enough.<sup>18</sup>  $B$  deviates to  $h^B = \bar{\sigma}$  if  $\phi p^m + (1 - \phi)\underline{p}^n < \phi \bar{p}^m + (1 - \phi)\bar{p}^n - \gamma[\phi(\bar{p}^m - \underline{p}^n) + (1 - \phi)(\bar{p}^n - \underline{p}^m)]$  or  $\gamma < (\Phi R + 1)/(\Phi + R)$ , which violates the condition.  $B$  will not deviate to  $h^B = \bar{\sigma}$  and  $P' > P^* \geq \underline{P}$  for  $\Omega^h$  low enough or to  $h^B = \bar{\sigma}$  and  $P^* > P' > E[U | a = 0]$  for  $\phi$  high enough and  $\omega^l$  low enough. Deviation to  $P' \leq E[U | a = 0]$  and  $h^B = \bar{\sigma}$  is always strictly worse than deviating to  $P' \leq E[U | a = 0]$  alone because the customer always buys anyway.  $W$  will not deviate to  $P' > P^* \geq \underline{P}$  or  $P^* > P' > E[U | a = 0]$  for the same reason as  $B$ . Deviation to  $P' \leq E[U | a = 0]$  is always strictly worse than deviating to  $P' \leq E[U | a = 0]$  and  $h^W = \underline{\sigma}$ .  $W$  deviates to  $h^W = \underline{\sigma}$  if

$$\phi \underline{p}^n + (1 - \phi)\underline{p}^m > \phi \bar{p}^n + (1 - \phi)\bar{p}^m - \gamma[\phi(\bar{p}^n - \underline{p}^n) + (1 - \phi)(\bar{p}^m - \underline{p}^m)],$$

which fails  $\forall \gamma < 1$ . Deviations to  $h^W = \underline{\sigma}$  and  $P' > P^* \geq \underline{P}$  are unprofitable at low  $\Omega^h$ . Deviation to  $P^* > P' > E[U | a = 0]$  and  $h^W = \underline{\sigma}$  is unprofitable for the same reason that each of the two deviations individually is unprofitable. Finally, deviation to  $P' \leq E[U | a = 0]$  and  $h^W = \underline{\sigma}$  is profitable if

$$P^*\{(1 - \gamma)[\phi \bar{p}^n + (1 - \phi)\bar{p}^m] + \gamma[\phi \underline{p}^n + (1 - \phi)\underline{p}^m]\} - \Delta C < \left[1 + \frac{\Omega^l(1 - \underline{p}^n) + \Phi(1 - \bar{p}^n)}{\Omega^l\Phi(1 - \underline{p}^m) + (1 - \bar{p}^m)}\right]^{-1},$$

which fails for  $\phi$  high enough and  $\alpha^l$  low enough. Proving equilibrium 5(b) is essentially analogous except for the fact that the off-path posteriors are not necessarily monotonic in the beliefs. They will depend on the signs of the following two quantities:  $M \equiv \bar{p}^m \bar{p}^n - \Phi \underline{p}^n \bar{p}^m$  and  $N \equiv (1 - \underline{p}^m)(1 - \bar{p}^n) - \Phi(1 - \underline{p}^n)(1 - \bar{p}^m)$ . Here,

$$E[U | a = 1, P' > P^*] = \left[1 + \frac{\Omega^h \bar{p}^n + \Phi \underline{p}^n}{\Omega^h \Phi \bar{p}^m + \underline{p}^m}\right]^{-1},$$

which is increasing in  $\omega^h$  if  $M < 0$  and decreasing if  $M > 0$ . Assuming that the equilibrium does not exist, deviations by  $B$  to  $P' > P^*$  are prevented by high-enough  $\Omega^h$  when  $M > 0$ . Since

$$E[U | a = 1, P' > P^*] \xrightarrow{\omega^h \rightarrow 1} \left[1 + \frac{\Phi \bar{p}^n}{\bar{p}^m}\right]^{-1},$$

the latter quantity is  $\underline{P}$ . When  $M < 0$ , deviations are prevented for high-enough  $\Omega^h$ . In this case,  $\underline{P} = [1 + \Phi \bar{p}^n / \bar{p}^m]^{-1}$ . Deviation to  $P^* >$

<sup>18</sup> Of course it also requires that  $\Delta C$  is not too high, which we have already assumed to be true throughout the paper. That is, we have assumed that  $\Delta C$  is such that the firm always prefers to implement  $\bar{e}$ , which implies that  $P^*[\phi \bar{p}^m + (1 - \phi)\bar{p}^n] - \Delta C > P^*[\phi(1 - \delta^m) + (1 - \phi)(1 - \delta^n)] > 0$ .

$P' > E[U | a = 0]$  only decreases price, but not  $\Pr[a = 1]$ , so is strictly worse. Deviation to  $P' \leq E[U | a = 0]$  is strictly worse than this deviation and deviating to  $h^B = \underline{\sigma}$ . When  $\gamma < 1$ , deviating on sales only by  $B$  never makes sense when  $\delta$ 's are high.  $B$  will not deviate to  $h^B = \underline{\sigma}$  and  $P' > P^* \geq \underline{P}$  for  $\Omega^h$  low enough or to  $h^B = \underline{\sigma}$  and  $P^* > P' > E[U | a = 0]$  for  $\gamma < 1$ . Deviation to some  $P' \leq E[U | a = 0]$  and  $h^B = \underline{\sigma}$  is weakly worse than deviating to

$$P' = E[U | a = 0] = \left[1 + \frac{\Omega^l(1 - \bar{p}^n) + \Phi(1 - \underline{p}^n)}{\Omega^l\Phi(1 - \bar{p}^m) + (1 - \underline{p}^m)}\right]^{-1}.$$

When  $N > (<) 0$ , this is decreasing (increasing) in  $\Omega^l$ . Such a deviation is prevented for  $\omega^l$  and  $\phi$  high enough since

$$\left[1 + \frac{\Omega^l(1 - \bar{p}^n) + \Phi(1 - \underline{p}^n)}{\Omega^l\Phi(1 - \bar{p}^m) + (1 - \underline{p}^m)}\right]^{-1} \xrightarrow{\phi \rightarrow 1, \omega^l \rightarrow 1} 0.$$

When  $N < 0$ , the analogous logic holds for low-enough  $\omega^l$  and high  $\phi$ . Now,  $W$  will not deviate to  $P' > P^* \geq \underline{P}$  or  $P^* > P' > E[U | a = 0]$  for the same reason as  $B$ . Deviation to  $P' \leq E[U | a = 0]$  is profitable if  $P^*[\phi \underline{p}^n + (1 - \phi)\underline{p}^m] - \Delta C < E[U | a = 0]$ . When  $N > (<) 0$ , this fails for  $\omega^l$  high (low) enough and  $\phi$  high enough.  $W$  deviates to  $h^W = \bar{\sigma}$  if  $\phi \bar{p}^n + (1 - \phi)\bar{p}^m - \gamma[\phi(\bar{p}^m - \underline{p}^m) + (1 - \phi)(\bar{p}^n - \underline{p}^n)] > \phi \underline{p}^n + (1 - \phi)\underline{p}^m$  or  $\gamma < (\Phi + R)/(\Phi R + 1)$ , which contradicts the condition. Deviations to  $h^W = \bar{\sigma}$  and  $P' > P^* \geq \underline{P}$  are unprofitable at low  $\Omega^h$  and high  $\gamma$ . Deviation to  $P^* > P' > E[U | a = 0]$  and  $h^W = \bar{\sigma}$  is unprofitable for the same reason that each of the two deviations individually is unprofitable. Finally, deviation to  $P' \leq E[U | a = 0]$  and  $h^W = \bar{\sigma}$  is strictly worse than deviating only to  $h^W = \bar{\sigma}$ , which we have already shown is unprofitable.  $\square$

**PROOF OF PROPOSITION 6.** First, assume there exist equilibria in which the firms separate on price:  $P^{B*} \neq P^{W*}$ . To prevent  $W$  from pretending to be  $B$ , we require that

$$\begin{aligned} P^{W*} \cdot [1(P^{W*} \leq E[U | a = 1, W, h^W]) \Pr[a = 1 | W, h^W] \\ + 1(P^{W*} \leq E[U | a = 0, W, h^W]) \cdot \Pr[a = 0 | W, h^W]] \\ \geq P^{B*} \cdot [1(P^{B*} \leq E[U | a = 1, B, h^B]) \Pr[a = 1 | B, h^B] \\ + 1(P^{B*} \leq E[U | a = 0, B, h^B]) \cdot \Pr[a = 0 | B, h^B]]. \end{aligned}$$

Note that

$$E[U | a = 1, W, h^W] = \left[1 + \Phi \frac{p(h^W, \delta^n)}{p(h^W, \delta^m)}\right]^{-1} \xrightarrow{\phi \rightarrow 1} 0.$$

The same is true of  $E[U | \theta = 0, W, h^W]$  forcing  $P^{W*}$  to 0, so no such equilibrium can exist for high-enough  $\phi$ . Now, assume that  $P^{B*} = P^{W*} \geq E[U | a = 0, h^B = \bar{\sigma}, h^W = \bar{\sigma}]$ . When the  $\delta$ 's are high, this requires that  $[\phi \bar{p}^n + (1 - \phi)\bar{p}^m] - \gamma[\phi(\bar{p}^m - \underline{p}^m) + (1 - \phi)(\bar{p}^n - \underline{p}^n)] \geq [\phi \underline{p}^n + (1 - \phi)\underline{p}^m]$ , which fails for high  $\gamma$  and  $\delta$ . The analogous condition fails for low  $\delta$ 's and high  $\gamma$ . Further, assume that  $h^B = h^W = \underline{\sigma}$  and  $P^{B*} = P^{W*} > E[U | a = 0, h^B = h^W = \underline{\sigma}]$ . When the  $\delta$ 's are high,  $B$  always deviates to  $h^B = \bar{\sigma}$  because  $\gamma < 1$ . The analogous deviation exists for  $W$  when the  $\delta$ 's are low. If  $h^B = h^W = \bar{\sigma}$  and  $P^{B*} = P^{W*} > E[U | a = 0, h^B = h^W = \bar{\sigma}]$ ,  $W$  deviates to  $h^W = \underline{\sigma}$  when

$\gamma$  and  $\delta$  are high enough and  $B$  deviates to  $h^B = \sigma$  when  $\gamma$  is high enough and  $\delta$  is low enough. Finally, the firms may pool on  $P^{B*} = P^{W*} \leq E[U | a = 0, h^B, h^W]$ . First, note that neither firm may hire the high-skill salesperson in any equilibrium because the firms have priced to sell regardless of the outcome. Thus, we will only consider the equilibrium where  $h^B = h^W = \sigma$ . Assume first that the  $\delta$ 's are high (the task is complex) so that  $\bar{p}^m - p^m > \bar{p}^n - p^n$ . A necessary condition for this equilibrium is that

$$\begin{aligned} P^* &= E[U | a = 0, h^B = h^W = \sigma] \\ &\geq E[U | P' > P^*, h^B = h^W = \sigma, \omega^h] \\ &\quad * \{(1 - \gamma)[\phi \bar{p}^m + (1 - \phi) \bar{p}^n] + \gamma[\phi p^m + (1 - \phi) p^n]\} - \Delta C, \end{aligned}$$

which prevents  $B$  from deviating to a higher price and hiring a high-skill salesperson. We have considered the highest possible price for this equilibrium because if this one is not sustainable, none of the lower-priced ones are. Taking the limit of both sides as  $\gamma \rightarrow 1$ ,  $\phi \rightarrow 1$ , and  $\delta^n \rightarrow 1$ , the condition becomes  $(1 - p^m)/(2 - p^m) \geq p^m - \Delta C$ , which fails for high  $\sigma$  since  $p^m \rightarrow_{\sigma \rightarrow 1} 0 \geq 1 - \Delta C$  for low  $\Delta C$ . The same approach holds for the case where  $\bar{p}^m - p^m < \bar{p}^n - p^n$ .  $\square$

## Appendix B: Notation

$T \in \{B, W\}$ : Firm type.  $B$  is better firm,  $W$  is worse firm

$t \in \{b, w\}$ : Customer type

$\mu \in \{m, n\}$ :  $m$  if customer "matches" ( $T = t$ ),  $n$  if not

$\phi \in (\frac{1}{2}, 1)$ : Proportion of customers that match firm  $B$

$\Phi \in (1, \infty)$ :  $\phi/(1 - \phi)$

$K$ : The cardinality of the set of actions

$a_k \in \{0, 1\}$ : 1 if action  $k$  is effective, 0 otherwise

$\delta^\mu \in (0, 1)$ : Complexity of the task of selling to customer with matching parameter  $\mu$

$\hat{a}$ :  $K$ -dimensional vector of impression-formation data

$\hat{a}_k \in \{0, 1\}$ : Signal of effectiveness of action  $k$

$\sigma \in \{\bar{\sigma}, \underline{\sigma}\}$ : Salesperson ability: "high-skill" or "low-skill"

$\hat{\sigma} \in \{\bar{\sigma}, \underline{\sigma}\}$ : Observable salesperson type: "probably high-skill" or "probably low-skill"

$\sigma^T \in \{\bar{\sigma}, \underline{\sigma}\}$ : Type of salesperson employed by firm  $T$

$\bar{p}^\mu, p^\mu \in (0, 1)$ : Probability that a high-skill (low-skill) salesperson chooses an effective action for customer with matching parameter  $\mu$

$R \in (0, \infty)$ : Relative impact of selling ability on matching customers as compared with nonmatching

$\bar{\alpha}, \alpha \in \Re$ : Salary component of compensation

$\bar{\beta}, \beta \in [0, 1]$ : Incentive paid to salesperson

$V(\sigma) \geq 0$ : Salesperson of type  $\sigma$ 's utility

$V^{\min}(\sigma) \geq 0$ : Salesperson of type  $\sigma$ 's reservation utility

$e \in [\bar{e}, \underline{e}]$ : Effort level: "high effort" or "low effort"

$C(e) \geq 0$ : Cost of putting forth effort level  $e$

$\Delta C > 0$ :  $C(\bar{e}) - C(\underline{e})$

$\gamma$ : Proportion of the marginal value delivered by a salesperson that she is able to capture

$h^T \in \{\bar{\sigma}, \underline{\sigma}\}$ ,  $h^T \in \{\bar{\sigma}, \underline{\sigma}\}$ : Firm (type)  $T$ 's hiring policy when there is one customer

$h_{1i}^T \in \{\bar{\sigma}, \underline{\sigma}, \tilde{\sigma}, \underline{\sigma}, \emptyset\}$ ,  $i = 1, 2$ : Firm  $T$ 's hiring policy when there are two customers

$I > 0$ : Investment in sales force training

$\lambda \in (0, 1)$ : Probability that a customer shops only a single firm

$\rho$ : Added to customer utility if there is a valuable solution and they match the company

$\nu \in [0, 1]$ : Hiring error

$b^T$ : Vector of bid components

$b_\alpha^T$ : Salary component of firm  $T$ 's bid

$b_\beta^T \in [0, 1]$ : Incentive component of firm  $T$ 's bid

$b_\sigma^T \in (\bar{\sigma}, \underline{\sigma})$ ,  $b_\sigma^T \in \{\bar{\sigma}, \underline{\sigma}\}$ : Salesperson to whom firm  $T$ 's bid is offered

$\varepsilon > 0$ : Required difference between current and previous bid

$\omega^h, \omega^m, \omega^l \in [0, 1]$ : Off-path beliefs: probability the firm is  $B$  given that it has offered a price higher than the highest equilibrium price, between the equilibrium prices, or lower than the lowest equilibrium price, respectively

## References

- Albrecht, J., S. Vroman. 1992. Nonexistence of single-wage equilibria in search models with adverse selection. *Rev. Econom. Stud.* **59** 617–624.
- Booth, A. 1984. A public choice model of trade union behaviour and membership. *Econom. J.* **94**(December) 883–898.
- Fudenberg, D., J. Tirole. 1991. *Game Theory*, Chapter 6. MIT Press.
- Kotler, P. 2000. *Marketing Management*, Millenium Ed. Prentice-Hall, Upper Saddle River, NJ.
- Little, J. D. C. 1970. Models and managers: The concept of a decision calculus. *Management Sci.* **16**(April) 466–485.
- Lodish, L. 1971. CALLPLAN: An interactive salesman's call planning system. *Management Sci.* **18**(December) 25–40.
- Mailath, G., A. Postlewaite. 1990. Workers versus firms: Bargaining over a firm's value. *Rev. Econom. Stud.* **57** 369–380.
- Malcomson, J. 1981. Unemployment and the efficiency wage hypothesis. *Econom. J.* **91**(December) 848–866.
- Marchetti, M. 1997. Hey buddy, can you spare \$113.25? *Sales and Marketing Management* **149**(8) 69–72.
- . 2000. What a sales call costs. *Sales and Marketing Management* (September) 80–82.
- Mirrlees, James. 1974. Notes on welfare economics, information, and uncertainty. M. Balch, D. McFadden, S. Wu, eds. *Essays on Economic Behavior Under Uncertainty*. North-Holland, Amsterdam, The Netherlands.
- Montgomery D., A. Silk, C. Zaragoza. 1971. A multiple-product sales force allocation model. *Management Sci.* **18**(December) 3–24.
- Moorthy, K. S., K. Srinivasan. 1995. Signaling quality with a money back guarantee: The role of transaction costs. *Marketing Sci.* **14**(Spring) 442–466.
- Narasimhan, Chakravarthi. 1988. Competitive promotional strategies. *J. Bus.* **61** 427–449.
- Rao, R. 1990. Compensating heterogeneous salesforces: Some explicit solutions. *Marketing Sci.* **9**(Fall) 319–341.
- Schlosser, J. 1999. ABC sweeps up in Nov. *Broadcasting and Cable* (December 6) 8.

## CODES

*Eye of the Beholder: An Analysis of the Relative Value of a Top Sales Rep Across Firms and Products*

---

- Weiss, A. 1980. Job queues and layoffs in labor markets with flexible wages. *J. Political Econom.* **88**(3) 526–538.
- Weitz, B. 1981. Effectiveness in sales interactions: A contingency framework. *J. Marketing* **45**(Winter) 85–103.
- Wernerfelt, B. 1994. On the function of sales assistance. *Marketing Sci.* **13** 68–82.
- Zoltners, A., P. Sinha. 1980. Integer programming models for sales resource allocation. *Management Sci.* **26**(March) 242–260.

*This paper was received May 29, 2000, and was with the authors 13 months for 4 revisions; processed by Charles Weinberg.*