University of Cyprus
Department of Economics

Working Paper 02-2018

Informal Delegation and Training

Emre Ekinci and Nikos Theodoropoulos

Department of Economics, University of Cyprus, P.O. Box 20537, 1678 Nicosia, Cyprus
Tel.: +357-22893700, Fax: +357-22895028, Web site: http://www.ucy.ac.cy/econ/en

Informal Delegation and Training*

Emre Ekinci[†] **Nikos Theodoropoulos**[‡]

Universidad Carlos III de Madrid University of Cyprus

February 2018

Abstract

This paper investigates the relationship between the firm's incentives to provide training and to delegate authority. We consider a principal-agent model in which the firm is not able to commit to delegation contractually and the conflict of interest between the firm and the worker arises both because the latter is biased towards certain decisions and because players interpret information differently (i.e., they have differences of opinion). Our theoretical analysis consists of two parts. First, we examine the equilibrium behavior when the degree of incongruence between the firm and the worker is public information. Second, we analyze the equilibrium behavior when the firm is privately informed about its type wherein the type refers to the level of differences in opinion between the firm and the worker. This exercise shows the extent to which the firm can use training provision to convey its private information to the worker, thereby committing not to retract the agent's authority it initially granted. In our empirical analysis, we use a cross section of matched employer-employee data of British establishments to examine the extent to which the model's predictions are supported by data.

*We thank Christos Bilanakos and Jungho Lee for their helpful comments. The financial support of Spain's Ministry of Economy and Competitiveness (ECO2016-75961-R) is greatly acknowledged.

[†]Universidad Carlos III de Madrid; emre.ekinci@uc3m.es

[‡]University of Cyprus; n.theodoropoulos@ucy.ac.cy.

1 Introduction

Two fundamental challenges faced by firms concern the allocation of decision rights and the provision of training. The (skill-biased) technological change experienced in industrialized countries has expanded the menu of organizational choices, and this altered the way firms allocate authority. For example, Caroli and Van Reenen (2001) find that an increase in the supply of human capital leads firms to decentralize authority. Similarly, Bloom et al. (2012) find that firms with higher proportions of college graduates are also more likely to delegate. In related work, Acemoglu et al. (2007) analyze how the diffusion of new technologies affects the allocation of authority within firms. They find that delegation is more prevalent in heterogenous industries and in firms that are either younger or closer to the technological frontier.

Despite extensive research that investigates human capital as a determinant of delegation, this literature has mostly overlooked the relationship between the firm's incentives to delegate authority and to provide training.[1] The goal of our analysis is to shed light on the interaction between the firm's delegation and training decisions when the firm's delegation decision is not credible and when the firm and the worker do not have fully congruent interests over the selection of an optimal action. In particular, we address the following questions: Does the firm provide higher training to the worker to whom it delegates authority? If so, what is the rationale behind this behavior? When the worker conjectures that the firm may take his authority back before exercising it, can the firm use training provision as a commitment device? That is, can the firm provide more training to convince the worker that it is not going to retract his authority?

To explore these questions, we develop a single-period principal-agent model whereby the firm makes choices on whether to delegate authority to the worker and how much training to provide to the same worker, whereas the worker chooses how much effort to exert in information acquisition. The goal of this employment relationship is to take an action (or to make a decision) which then determines the payoff of each player depending on the state of the world initially unknown to either player.

The key ingredients of our model are the following. First, decision rights are not contractible as in Baker et al. (1999). That is, even though the firm can informally delegate authority to a worker, before the worker exercises his authority, the firm can revoke it, at

[1] See Aghion et al. (2013) and Bolton and Dewatripont (2013) for recent surveys of the literature. For a textbook treatment and references to earlier analyses, see Laffont and Martimort (2009).

a cost, if it is its interest to do so.[2] Second, along the lines of Che and Kartik (2009), we allow the firm and the worker to have different prior beliefs about the state. This leads each player to interpret any information about the state differently even though they obtain exactly the same information. Therefore, the conflict of interest between the firm and the worker does not arise only because they have different underlying preferences but also because they disagree on how to interpret information. Third, the firm provides training which is complementary to effort in the worker's cost function. That is, additional training reduces the worker's marginal cost of effort in information acquisition.[3]

To highlight the role of differences of opinion, we consider two types of principals. The first type of principal is referred to as "low type" because the disagreement between this type of principal and the agent arises solely from the agent's bias for smaller actions. In contrast, the second type of principal is referred to as "high type" because in addition to the agent's preference for smaller actions, this type of principal also disagrees with the agent about how to interpret any information about the state. Therefore, the conflict of interest over the optimal action is more pronounced when the agent is hired by the high type principal than when he is hired by the low type principal.

Central to our analysis is the principal's temptation to renege on her promise by taking back the agent's authority. We show that there exists a critical cutoff such that a principal finds it optimal to retract the agent's authority she initially granted if and only if her cost of revoking delegation is lower than the cutoff. The critical cutoff depends on the principal's type in the following ways. First, the cutoff is always lower for the low type principal than for the high type principal. When the principal of a given type lets the agent exercise his authority, she incurs disutility due to the agent's biased decision. Because the magnitude of disutility is smaller for the low type principal, it is cheaper for her to let the agent choose an action. Second, observing a signal increases the cutoff for the high type principal, whereas it does not affect it for the low type principal. Acquiring information about the state

[2] As discussed by Bolton and Dewatripont (2013), delegation within organizations may not be credible due to the so-called business judgement rules which induce courts not to enforce any contracts written between parties in a single organization. The cases to which those rules are not applicable include contracts between an organization and its financial investors, and contracts within universities. See Aghion et al. (2008) for an analysis of how the difference in terms of enforceability of delegation contracts affects a scientist's project choice.

[3] To keep our analysis tractable, we do not take into account how the transferability of human capital generated by training alters firms' incentives to provide training. In other words, because our analysis focuses on a single period, the firm is not concerned about not reaping the returns from training if the worker departs from the firm. See Acemoglu and Pischke (1998, 1999) for related analyses.

mitigates incongruity between the high type principal and the agent, and consequently, makes it cheaper for the high type principal to let the agent exercise his authority. In contrast, because the degree of disagreement between the low type principal and the agent is not driven by any differences of opinion but the agent's exogenous bias, acquiring information does not alter the low type principal's decision to revoke delegation decision.

To investigate how the principal's inability to commit to delegation alters her incentives to allocate decision rights and provide training, our analysis focuses on parameterizations of the model under which the cost of revoking delegation is sufficiently large that the low type of principal never takes the authority back from the agent, whereas the high type principal retracts the agent's authority only when she does not observe any signal.

The type of the principal determines the extent to which the delegation decision involves a tradeoff between increased initiative and loss of control in the spirit of Aghion and Tirole (1997). On the one hand, because the agent will pursue his own interests rather than those of the principal in exercising his authority, delegation entails a costly loss of control for either type. But, because the degree of incongruence is more pronounced when the principal is of high type, this type of principal incurs a larger cost from the loss of control. On the other hand, the agent will have stronger incentives to acquire information if he is delegated authority by the high type principal, whereas this effect will be absent if the low type delegates authority to the agent. Hence, depending on the relative magnitude of these two effects, the high type principal may find it optimal to delegate authority, whereas the low type principal has no incentive to delegate.

In the first part of our theoretical analysis, we consider a symmetric-information structure in which the type of the principal is public information. In equilibrium, we observe that the low type principal retains authority, whereas the high type principal delegates authority to the agent if the uncertainty in the environment is sufficiently high or the agent's signal is sufficiently precise. The low type principal opts for centralization because, as indicated, delegation does not lead the agent to exert higher effort but brings about a cost stemming from loss of control. Note that the agent has lack of additional effort incentives under delegation because he chooses his preferred action regardless of whether he generates a signal, since the low type principal never retracts the agent's authority. In contrast, the high type principal has an incentive to delegate because she elicits higher effort from the agent under delegation. Unlike the former case, the agent now anticipates that he will be able to exercise his authority only if his efforts generate a signal about the state; thus, delegation provides additional incentives for effort in information acquisition. We show that the high type principal's rents from delegation increase with the uncertainty in the environment and the precision of the agent's signal. Intuitively, as either the environment becomes more volatile or the agent's signal becomes more informative, the value of acquiring information increases for the principal. This renders delegating authority more attractive for the principal of high type because the worker exerts higher effort under delegation.

Our analysis also shows that the principal who delegates authority provides higher training to the agent than the principal who retains authority. The logic behind this result is the following. Since training reduces the worker's marginal cost of effort and all players are better off when a decision is made after a signal is observed, principals have incentives to provide training under both organizational modes. But, the principal who delegates authority has additional incentives to provide training, since the value of observing a signal is higher under delegation. Intuitively, if the agent generates a signal, he can exercise his authority in which case the principal avoids the cost of revoking delegation but incurs disutility from the agent's biased decision. Because the positive effect of avoiding the revoking cost outweighs the disutility due to incongruence (otherwise it would be optimal for the principal to retract the agent's authority even after he generates a signal), the marginal benefit of providing training is larger for the principal who delegates authority than for the principal who retains authority. Hence, delegation is associated with higher levels of training.

In the second part of our theoretical analysis, we examine the asymmetric-information case to investigate the extent to which principals, privately informed about their own types, can use training provision to convey their private information to the agent.[4] Our analysis shows that the high type principal can increase training provision to a level at which the low type principal does not find it optimal to provide; thus, it is possible for the high type principal to signal her type to the agent. However, the high type principal does not find it optimal to do so because the cost of signaling through additional training offsets the benefits from delegation. In other words, the size of delegation rents that the high type principal would earn is not sufficiently large to make it optimal to signal her type by providing additional training. Hence, we observe pooling behavior in equilibrium. Specifically, if the agent's prior belief that a given principal is of high type is sufficiently high, both types of

[4]The idea that the informed principal can take a costly action to signal her type to the agent has been examined in other contexts. For example, Fuchs (2015) considers a setting in which the firm privately observes the quality of the match between the worker and the firm, and he shows that the firm can credibly convey this information to the worker by paying discretionary bonuses.

principals delegate authority; otherwise, they retain authority. Also, as in the symmetric-information case, delegation, if observed, is associated with higher levels of training.

In the second part of the paper, we investigate the extent to which the implications of our theory are supported by the data. To test the theory's predictions, we use data drawn from a large, nationally-representative cross-section of British workplaces, the 2004 wave of the Workplace Employee Relations Survey (WERS).[5] We investigate the following three predictions that our theory yields: i) delegation and training are positively related; ii) firms are more likely to delegate authority as the uncertainty in the environment increases; iii) firms are more likely to delegate authority as the worker's signal becomes more precise. The empirical results provide support for our theory.

The outline for the paper is as follows. The next section discusses related work in the literature. Section 3 presents our theory. In subsections,we first present the setup of our model and provide preliminary results. Next, we present an analysis of the equilibrium of the model and then close this section with a discussion of the model's implications. Section 4 presents the data and discusses the empirical results. Section 5 presents concluding remarks.

2 Related Literature

Our paper falls into the extensive literature examining various aspects of delegation of authority within organizations. Starting with the seminal work of Aghion and Tirole (1997), a growing number of papers have examined the interplay between delegation and incentives for information acquisition (e.g., Zabojnik, 2002; Bester and Krähmer, 2008; Rantakari, 2012; Shin and Strausz, 2014). As in these papers, the question of how delegating authority alters the agent's incentives for information acquisition is essential to our analysis; however, our analysis differs in several important ways. First, we consider an environment in which delegation is not contractible.[6] Second, in our setting the conflict of interest between the firm and the worker does not only arise from their different underlying preferences but also because they interpret information differently.

As in Baker et al. (1999), our model is built on the assumption that the principal's

[5]Other studies that examine the delegation of authority and employ this dataset include Acemoglu et al. (2007), DeVaro and Kurtulus (2010), DeVaro and Prasad (2015) and Bilanakos et al. (2017).

[6]Baker et al. (1999) and Aghion et al. (2008) also consider settings in which decision rights are not contractible. Another interesting paper in this strand of the literature is by Marino et al. (2009). These authors examine the allocation of authority when the principal's ability to centralize decision making is limited by the agent's willingness to disobey the order made by the principal.

delegation decision is not credible, but it differs from theirs in the following ways. First, we examine how the principal's commitment problem affects her delegation and training decisions in a single-period model, whereas Baker et al. (1999) study an infinitely repeated game and show that the principal can informally delegate authority through self-enforcing contracts. Second, we consider informational environments that differ from those in their analysis. In particular, the asymmetric-information structure in our setting pertains to the case in which the agent lacks the information about the degree of incongruence between the principal and himself, whereas in their setting the principal does not have the same information as the agent about the returns generated by implementing the project proposed by the agent. In other words, we analyze an informed principal's problem about conveying her type to the agent, whereas they examine if informal delegation can be sustained through repeated interactions when the agent is better informed than the principal. Similarly, Alonso and Matouschek (2007) consider an infinitely repeated game in which the agent is better informed than the principal; but, they examine the agent's incentives to reveal information under different organizational modes the principal can commit to.[7]

In terms of modeling the conflict of interest between the principal and the agent, our paper is closest to that of Che and Kartik (2009). These authors investigate how differences of opinion affect the agent's incentives to acquire information and the strategic disclosure of any information he has acquired. Their main result is that difference of opinion generates incentives for effort in information acquisition but also worsens the agent's information disclosure. In a part of their analysis, Che and Kartik (2009) modify the Aghion-Tirole model to consider the relationship between differences of opinion and the principal's decision to allocate authority. However, the delegation decision in their setting is irreversible (thus, the principal has no commitment problem) and they do not consider training provision. Hence, our theoretical analysis is complementary to their analysis concerning the principal's decision to allocate authority.

Finally, Bilanakos et al. (2017) examine the relationship between delegation and training by incorporating the firm's training decision into the Aghion-Tirole framework. As in our analysis, they assume that training and worker effort are complementary; however, because our setup differs significantly from the one considered by Aghion and Tirole (1997), the implications of our model are different from those of Bilanakos et al. (2017). First, even

[7]This strand of the literature in which the agent communicates his information strategically stems from the seminal work by Dessein (2002). See Bolton and Dewatripont (2013) or Gibbons et al. (2013) for a more detailed discussion.

though their model also implies that firms provide higher levels of training under delegation, their rationale behind this prediction is different from ours. Specifically, they show that the firm opts for delegation and provides more training than it provides under centralization if the preferences of the firm and those of the worker are sufficiently congruent. In contrast, we show that the firm has an incentive to delegate authority if the conflict of interest between the firm and the worker over the optimal decision is sufficiently pronounced. Second, their model is silent about the effects of uncertainty in the environment and the precision of the worker's information on the firm's incentives to delegate.[8]

3 Theoretical Framework

This section presents our theoretical framework we build upon Che and Kartik (2009). We first provide the setup of our model and preliminary results. Next, we analyze symmetric- and asymmetric-information equilibria, respectively. In the last part, we discuss the model's empirical implications. All proofs are presented in the Appendix.

3.1 Model Setup

We consider a delegation game with two players: the principal ("P", "she" and "firm") and the agent ("A", "he" and "worker"). The principal (or the agent if he has the authority) must take an action, $a \in \mathbb{R}$, and the payoffs of the players depend on the unknown state of the world, $\theta \in \mathbb{R}$.[9] When action a is implemented, the utilities of the principal and the agent, respectively, are

$$u_P(a, \theta) = -(a-\theta)^2 \tag{1}$$

and

$$u_A(a, \theta, b) = -(a-\theta-b)^2, \tag{2}$$

where $b \leq 0$ denotes the agent's bias towards actions with smaller values.[10] As indicated, the true value of the state variable is initially unknown. The agent believes that the state variable is normally distributed with mean μ_A and variance σ_0^2. The principal, however, may

[8] Also, even though they employ the same dataset as we do, they do not test our second and third hypotheses.

[9] Throughout the analysis, we use "action" and "decision" interchangeably.

[10] This preference structure is common in the delegation literature. See, for example, Alonso and Matouschek (2007), Che and Kartik (2009).

have a different opinion concerning the state variable. Denoting the type of the principal by q, the agent knows that the principal is a low type ($q = L$) with probability γ and a high type ($q = H$) with probability $1 - \gamma$. The prior belief of the type q principal about the state variable is given by $\theta \sim N\left(\mu_q, \sigma_0^2\right)$. Without loss of generality, we assume that $\mu_H > \mu_L = \mu_A = 0$, that is, P_L and A agree on the prior distribution of the state variable, whereas P_H and A have a disagreement over this. As will be discussed in the next section, the difference of opinion between the principal and the agent alters how each player interprets any signal about the state, and consequently, their preferred actions.

As in Aghion and Tirole (1997), information acquisition plays a central role on delegation decision. Because the agent has expertise in the tasks assigned to him, he can exert costly effort to acquire a signal about the state. In particular, he exerts effort ψ to observe signal s with probability ψ, where $s \sim N\left(\theta, \sigma_s^2\right)$. The cost of exerting effort ψ, $\psi \in \left[\underline{\psi}, \overline{\psi}\right]$, is given by function $g(\psi, t)$, where t is the amount of training the agent has received at the firm. We assume that $g(\psi, t)$ is strictly increasing and convex with respect to ψ, and additional training reduces the marginal cost of effort, i.e., $\frac{\partial^2 g(\psi,t)}{\partial\psi\partial t} < 0$ for any ψ. The agent decides how much effort to expend in acquiring a signal and his effort choice is not observed by the principal. But, because training reduces the marginal cost of effort borne by the agent, the principal can induce the agent to exert higher levels of effort by providing more training. More specifically, the principal chooses the level of training t, $t \in \left[0, \bar{t}\right]$, and incurs the cost of training given by $c(t)$, which is strictly increasing and convex.[11] Finally, the outcome of the worker's investigation is publicly observed. That is, the principal observes the signal if the agent's efforts generate one.[12]

We adopt the incomplete contracting approach (Grossman and Hart, 1986) by postulating that the principal cannot write incentive contracts contingent on the agent's actions, the signal or payoffs.[13] Further, along the lines of Baker et al. (1999), the delegation decision is assumed to be not contractible. That is, the principal has the ability to delegate authority to the agent; however, her delegation decision is not credible because she can revoke the

[11]We assume that both $g(.)$ and $c(.)$ satisfy the Inada conditions, i.e., $g(.)$ satisfies $\frac{\partial g(\underline{\psi},t)}{\partial\psi} = 0$ and $\frac{\partial g(\overline{\psi},t)}{\partial\psi} = \infty$ for any t, and $c(.)$ satisfies $c'(0) = 0$ and $c'\left(\bar{t}\right) = \infty$. In addition, to derive comparative statics results, we assume that the third order partial derivatives of $g(.)$ are zero and that $c'''(.) \leq 0$.

[12]This feature of our model contrasts with that of Che and Kartik (2009) who assume that the agent can withhold a signal strategically.

[13]For an analysis of the relationship between incentive pay and delegation, see, for instance, Prendergast (2002), DeVaro and Kurtulus (2010), DeVaro and Prasad (2015) or Lo et al. (2016).

authority before the agent takes any action if she finds it optimal to do so. Conditional on having granted authority to the agent, the principal incurs a cost of R, $R > 0$, if she chooses to revoke the agent's authority. The cost of retracting the agent's authority may arise, for example, from the firm's reputational concerns as in Baker et al. (1999), or it may arise because the agent whose authority is revoked may engage in activities harmful to the principal (Hart and Holmstrom, 2010). Finally, the principal also incurs an implementation cost κ, $\kappa > 0$, when she, rather than the agent, chooses the action to be implemented.

The timing of the events is as follows.[14] The type of the principal is determined by the nature and privately observed by the principal herself. Next, the principal decides whether to delegate authority to the agent and how much training to provide to the agent.[15] The agent then decides how much effort to exert in acquiring information. Afterwards, the outcome of the agent's investigation is publicly observed. If the agent was initially granted delegation, the principal then decides whether to retract authority from the agent. Finally, the player with the authority takes an action and payoffs are determined according to the chosen action and the realized value of the state variable.

3.2 Preliminary Analysis

Before deriving the equilibrium, we discuss some features of the model to provide insight about each player's optimal behavior.

Given that the utility of each player decreases with the distance between the action taken and the realized value of the state, the players use all information available to update their beliefs concerning the state variable. When player i, $i = L, H, A$, observes signal s, his or her posterior concerning the state variable, θ, is normally distributed with mean $\widetilde{\mu}_i = \rho s + (1-\rho)\mu_i$ and variance $\widetilde{\sigma}^2 = \frac{\sigma_0^2 \sigma_s^2}{\sigma_0^2 + \sigma_s^2}$, where $\rho = \frac{\sigma_0^2}{\sigma_0^2 + \sigma_s^2}$ (DeGroot, 2004). Note that because each player interprets the signal according to his or her opinion (i.e., his or her own prior belief), their posterior means differ.

Using state-dependent utilities on (1) and (2), one can write each player's expected utility

[14] Analyzing a similar setup, Che and Kartik (2009) show that the presence of participation constraints does not affect the agent's incentives for information acquisition, but may affect the principal's choice over agent type (where the type refers to the degree of initial disagreement between the principal and the agent). Therefore, we could assume that the worker is paid a constant wage which equals his reservation wage.

[15] Dustmann and Schönberg (2012) argue that the firm cannot commit to training provision because training provided by the firm is usually not verifiable by third parties. Following their approach, we assume that training is not contractible.

conditional on signal s as follows:

$$\mathbb{E}\left[u_{P_q}(a,\theta,b)|s\right] = -\left(a - \mathbb{E}\left[\theta|s\right]\right)^2 - \mathbb{V}\left[\theta|s\right], \tag{3}$$

and

$$\mathbb{E}\left[u_A(a,\theta,b)|s\right] = -\left(a - \mathbb{E}\left[\theta|s\right] - b\right)^2 - \mathbb{V}\left[\theta|s\right]. \tag{4}$$

Let $\alpha_i(\Omega_i)$ denote the optimal action of player i conditional on information set Ω_i, where $\Omega_i = s$ if player i observes signal s and $\Omega_i = \emptyset$ if not. From (3) and (4), it is easy to see that $\alpha_A(\Omega_A) = \mathbb{E}\left[\theta|\Omega_A\right] + b$ and $\alpha_{P_q}(\Omega_{P_q}) = \mathbb{E}\left[\theta|\Omega_{P_q}\right]$, where $q = L, H$.

The players' optimal actions display two sources for a conflict of interest. First, because the agent is biased towards smaller actions, he scales down his forecast of the state by b. Second, even when the type H principal and the agent have the same information about θ, their optimal actions differ not only because the agent is biased towards smaller actions, but also because each player interprets the signal differently. To see this more clearly, suppose that neither of the players observes signal s, i.e., $\Omega_{P_H} = \Omega_A = \emptyset$. In this case, the optimal action for P_H is μ_H, whereas the optimal action for the agent is b (since his prior mean is normalized to zero). More interestingly, consider their preferred actions upon observing a signal, i.e., $\Omega_{P_H} = \Omega_A = s$. In this case, the optimal action is $\widetilde{\mu}_H$ for P_H and $\widetilde{\mu}_A + b$ for the agent. Hence, the difference between their optimal actions is $(1-\rho)\mu_H - b$. Note that signal s appears in the optimal action of both players with the same weight. Therefore, the exact value of the signal does not cause any additional incongruity between the type H principal and the agent.[16] Finally, note that the conflict of interest between the type L principal and the agent arises solely from the agent's bias for smaller actions, since they agree over how to interpret any information about the state.

To summarize the discussion, let the agent's bias from the type H principal's standpoint be denoted by $B_s = (1-\rho)\mu_H - b$ if the agent's efforts generate a signal and $B_\emptyset = \mu_H - b$ if not. Similarly, let the agent's bias from the type L principal's standpoint be denoted by $B = -b$. Following the terminology of Che and Kartik (2009), we will refer to B_s as the interim bias. We make three observations here. First, as long as signal s is informative (i.e., $\sigma_s^2 < \infty$), the bias is lower when the agent generates a signal, that is, $B_s < B_\emptyset$. Second, the magnitude of the interim bias is decreasing in the initial uncertainty in the environment

[16] Note that this result depends on the assumption that players' priors beliefs have the same variance, σ_0^2. See Che and Kartik (2009) who employ the same assumption.

and the precision of the signal, i.e., B_s decreases with σ_0^2 and increases with σ_s^2.[17] Third, regardless of whether a signal has been generated by the agent, the type H principal is exposed to larger bias than the type L principal if the agent is given the authority to take an action, i.e., $B < B_s < B_\emptyset$.

3.3 Equilibrium Analysis

In this section, we derive the equilibrium behavior. We begin our analysis with an examination of the symmetric information case in which the type of the principal is public information. Next, we introduce an informational asymmetry by assuming that the type of the principal is her private information and then examine the equilibrium behavior under this assumption.

3.3.1 Symmetric Information

Suppose that the agent observes the type of the principal he is hired by. To derive the equilibrium, we first consider each player's optimal actions under centralization and under delegation. Next, we describe the equilibrium and provide comparative statics results.

Centralization Suppose that the principal does not delegate authority to the agent, i.e., decision making is centralized. To derive the players' optimal behavior, we first consider the effort decision of the agent who faces the type q principal, where $q = L, H$. Since the principal has the authority over the decision to be made, the agent anticipates that the principal will choose $\alpha_{P_q}(s)$ if she observes signal s and $\alpha_{P_q}(\emptyset)$ if not. Therefore, given his level of training, t, the agent chooses a level of effort, ψ_q, to maximize his expected payoff:

$$U_A^c(\psi_q; t) = \psi_q \mathbb{E}\left[u_A(\alpha_{P_q}(s))\right] + (1 - \psi_q)\mathbb{E}\left[u_A(\alpha_{P_q}(\emptyset))\right] - g(\psi_q, t). \tag{5}$$

[17]This result hinges on how much weight is assigned to the signal in predicting the state. Because the type H principal and the agent assign the same weight to the signal in forming their preferred actions, their disagreement is mitigated as the weight of the signal increases. For example, the signal is assigned a higher weight in the posterior mean as it becomes more precise. Indeed, in the extreme case in which the signal is perfectly informative about the state, the difference in the players' optimal actions arises solely from the agent's bias, b (since $\widetilde{\mu}_H - \widetilde{\mu}_A \to 0$ as $\sigma_s^2 \to 0$).

The agent's optimal effort, $\psi_q^c(t)$, is then characterized by

$$\mathbb{E}\left[u_A(\alpha_{P_q}(s))\right] - \mathbb{E}\left[u_A(\alpha_{P_q}(\emptyset))\right] = \frac{\partial g\left(\psi_q^c(t), t\right)}{\partial \psi_q}. \quad (6)$$

Two aspects of $\psi_q^c(t)$ are worth noting. First, even though the agent has no power to make any decisions, his optimal effort is strictly higher than the minimum feasible effort. Her effort incentives stem from the fact that by generating a signal, he achieves lower uncertainty about the state and reduces the difference of opinion between the type H principal and himself. Second, because the cost function is strictly convex with respect to effort and the marginal cost of effort decreases with training, the agent's effort choice is increasing in the level of training provided by the principal.

Anticipating how the level of training alters the agent's effort provision, the type q principal chooses t_q to maximize her expected payoff:

$$U_{P_q}^c\left(t_q; \psi_q^c(t_q)\right) = \psi_q^c(t_q)\,\mathbb{E}\left[u_{Pq}(\alpha_{Pq}(s))\right] + \left(1 - \psi_q^C(t_q)\right)\mathbb{E}\left[u_{P_q}(\alpha_{Pq}(\emptyset))\right] - \kappa - c(t_q). \quad (7)$$

The type q principal's optimal training choice, t_q^c, equates the expected marginal benefit of training with the marginal cost, that is,

$$\frac{\partial \psi_q^c\left(t_q^c\right)}{\partial t}\left[\mathbb{E}\left[u_{P_q}(\alpha_{P_q}(s))\right] - \mathbb{E}\left[u_{P_q}(\alpha_{P_q}(\emptyset))\right]\right] = c'\left(t_q^c\right). \quad (8)$$

As reflected on (8), the marginal benefit of training is determined by the product of two terms: the term in the bracket shows how much the principal's payoff changes if the agent generates a signal and the first term measures how a change in the training provision alters the agent's optimal effort choice.

In sum, when the type q principal chooses to retain authority, she provides training t_q^c, given by (8), and the agent exerts effort $\psi_q^c\left(t_q^c\right)$, given by (6).

Informal Delegation In this part of the analysis, we focus on the equilibrium behavior when the principal initially delegates authority to the agent. To solve the game, we start with the principal's decision to retract the agent's authority and proceed backward.

Suppose that the agent was initially given authority to choose an action and consider if the principal finds it optimal to revoke her delegation decision. The principal makes this decision after the outcome of the agent's investigation is publicly observed (i.e., after all players observe whether or not the agent's efforts have generated a signal). This means that the principal's optimal decision to retract the agent's authority hinges on what information is available to her, that is, either $\Omega_{Pq} = \emptyset$ or $\Omega_{P_q} = s$. Let $r_q(\Omega_{P_q})$ denote the type q principal's decision to retract authority delegated to the agent, where $r_q(\Omega_{P_q}) = 1$ if the principal retracts the agent's authority and $r_q(\Omega_{P_q}) = 0$ if not.

First, we consider the type L principal's incentive to take the authority back from the agent. Because there is no difference of opinion between this type of principal and the agent, the conflict of interest between them arises solely from the agent's bias for smaller actions. Therefore, the type L principal finds it optimal to revoke authority if and only if the sum of the cost of revoking delegation decision and the cost of implementation, $R + \kappa$, is smaller than the disutility caused by the agent's biased decision, b^2. Note that observing a signal does not alter the type L principal's decision to revoke authority because the signal itself does not affect her disagreement with the agent over the optimal action.

In contrast, the signal plays a crucial role on the type H principal's decision to retract the agent's authority. To see this, suppose that the agent's efforts did not generate any signal. If the agent retains authority and takes his preferred action, the principal incurs disutility of $B_\emptyset^2$, which captures the agent's bias for smaller actions and the difference of opinion between the principal and the agent. Therefore, the type H principal, who did not observe any additional signal about the state, finds it optimal to revoke her delegation decision if and only if the sum of the revoking cost and the implementation cost, $R + \kappa$, is smaller than the disutility caused by the agent's bias, $B_\emptyset^2$. The same intuition describes the principal's decision to revoke authority upon observing a signal. In sum, the type H principal revokes authority as long as the disutility stemming from the agent's biased decision making (either B_s^2 or $B_\emptyset^2$ depending on whether the agent has generated a signal) is lower than the cost of revoking authority plus the cost of implementing the action, $R + \kappa$.

Note that each type of principal's optimal decision to retract authority is characterized by a cost threshold above which they find it optimal to let the agent exercise the authority they had initially granted to him. The following lemma formally states this result.

Lemma 1. *Suppose that the principal of type q initially delegates authority to the agent. She finds it optimal to revoke authority if and only if the cost of revoking is sufficiently small, that is, $r_q(\Omega_{P_q}) = 1$ if and only if $R < R_q^*(\Omega_{Pq})$, where $\Omega_{Pq} = \emptyset, s$. Moreover, $R_H^*(\Omega_{P_H} = \emptyset) > R_H^*(\Omega_{P_H} = s) > R_L^*(\Omega_{P_L} = \emptyset) = R_L^*(\Omega_{P_L} = s)$.*

To focus on the interaction between informal delegation and training incentives, we assume that $R_H^*(\Omega_{P_H} = s) < R < R_H^*(\Omega_{P_H} = \emptyset)$ in the rest of the analysis. That is, conditional on granting authority to the agent, the type H principal retracts the agent's authority only when she does not observe any signal, whereas the type L principal never takes the authority back from the agent.

After characterizing the principal's decision to revoke authority, we turn to the agent's effort provision in information acquisition. Observing the type of the principal he faces, the agent anticipates perfectly the conditions under which his authority will be retracted by the principal before exercising it. Therefore, given his amount of training, t, the agent chooses level of effort ψ_q to maximize his expected payoff:

$$U_A^d(\psi_q; t) = \begin{cases} \psi_L \mathbb{E}[u_A(\alpha_A(s))] + (1-\psi_L)\mathbb{E}[u_A(\alpha_A(\emptyset))] - g(\psi_L, t), & \text{if } q = L \\ \psi_H \mathbb{E}[u_A(\alpha_A(s))] + (1-\psi_H)\mathbb{E}[u_A(\alpha_{P_H}(\emptyset))] - g(\psi_H, t), & \text{if } q = H \end{cases}. \tag{9}$$

For given level of training t, let $\psi_L^d(t)$ denote the agent's optimal effort choice when she faces the type L principal and $\psi_H^d(t)$ denote her optimal effort choice when she faces the type H principal. The first-order conditions that characterize these optimal effort levels are

$$\mathbb{E}[u_A(\alpha_A(s))] - \mathbb{E}[u_A(\alpha_A(\emptyset))] = \frac{\partial g(\psi_L^d, t)}{\partial \psi_L} \tag{10}$$

and

$$\mathbb{E}[u_A(\alpha_A(s))] - \mathbb{E}[u_A(\alpha_{P_H}(\emptyset))] = \frac{\partial g(\psi_H^d, t)}{\partial \psi_H}. \tag{11}$$

Contrasting (11) with (10) shows how the marginal returns to effort change with the type of principal faced by the agent. Because the agent always exercises his authority when faced by the type L principal, his incentives for effort stem only from the reduction in the uncertainty concerning the state. In contrast, the agent has additional effort incentives when faced by the type H principal. This follows since by generating a signal, the agent not only enjoys a reduction in the uncertainty but also avoids his authority being taken back by the principal. The latter point is crucial for the agent because he is not exposed to the type H principal's biased decision if he, rather than the principal, chooses the action to be implemented. Hence, holding training level constant, difference of opinion generates additional incentives for information acquisition when the agent is delegated authority.

Anticipating the agent's effort choice, each type of principal decides how much training to provide. To that end, the type L principal chooses the amount of training t_L that maximizes

$$U^d_{P_L}\left(t_L;\psi^d_L(t_L)\right) = \psi^d_L(t_L)\,\mathbb{E}\left[u_{P_L}(\alpha_A(s))\right] + \left(1-\psi^d_L(t_L)\right)\mathbb{E}\left[u_{P_L}(\alpha_A(\emptyset))\right] - c(t_L)\,, \quad (12)$$

whereas the type H principal chooses t_H to maximize

$$U^d_{P_H}\left(t_H;\psi^d_H(t_H)\right) = \psi^d_H(t_H)\,\mathbb{E}\left[u_{P_H}(\alpha_A(s))\right] + \left(1-\psi^d_H(t_H)\right)\left[\mathbb{E}\left[u_A(\alpha_{P_H}(\emptyset))\right] - R - \kappa\right] - c(t_H)\,. \quad (13)$$

The optimal training levels are characterized by the following first-order conditions:

$$\frac{\partial\psi^d_L\left(t^d_L\right)}{\partial t_L}\left[\mathbb{E}\left[u_{P_L}(\alpha_A(s))\right] - \mathbb{E}\left[u_{P_L}(\alpha_A(\emptyset))\right]\right] = c'\left(t^d_L\right) \quad (14)$$

and

$$\frac{\partial\psi^d_H\left(t^d_H\right)}{\partial t_H}\left[\mathbb{E}\left[u_{P_H}(\alpha_A(s))\right] - \mathbb{E}\left[u_{P_H}(\alpha_{P_H}(\emptyset))\right] + R + \kappa\right] = c'\left(t^d_H\right). \quad (15)$$

As reflected by (14) and (15), the marginal returns to training depend on the product of two factors: how additional training alters the agent's effort provision and how the principal's expected payoff changes if a decision is made upon observing a signal. Considering the second terms, we see that the type H principal has stronger incentives to provide training. As seen on (14), the increase in the type L principal's payoff upon observing the signal consists only of the reduction in the uncertainty about the state, since the agent exercises his authority regardless of whether he generates a signal. In contrast, in addition to the reduction in the uncertainty, observing a signal increases the type H principal's payoff through an additional channel. Specifically, because the type H principal does not revoke authority when the agent generates a signal, she avoids the cost of revoking authority and the cost of implementation but becomes exposed to the agent's biased decision making. Because the size of these costs outweighs the disutility stemming from the agent's biased decision (as stated in Lemma 1, it would otherwise be optimal to revoke authority even after observing a signal), this type of principal has additional incentives to provide training.

Equilibrium Behavior After deriving the players' optimal behavior both under centralization and under delegation, we now consider the earlier step on which the principal decides whether to delegate authority to the agent. To derive the principal's delegation decision, we examine the players' optimal training and effort choices under both organizational modes.

Lemma 2. *Optimal training and effort choices satisfy the following conditions:*

i) $t_H^d > t_H^c = t_L^d = t_L^c > 0$.

ii) $\psi_H^d\left(t_H^d\right) > \psi_H^c\left(t_H^c\right) > \psi_L^d\left(t_L^d\right) = \psi_L^c\left(t_L^c\right)$.

Lemma 2 indicates that the agent is given the same (and a positive) level of training both when the type L principal chooses either organizational mode and when the type H principal retains authority. This follows because the marginal benefit of training in these cases is the same; more specifically, it is positive and proportionate to the reduction in the uncertainty about the state variable. Conversely, the type H principal provides a higher level of training when she delegates authority to the agent. This is because by providing additional training, the principal elicits higher effort from the agent and thus becomes less likely to revoke authority she initially granted to the agent. The possibility of encouraging effort gives the principal additional incentives to provide training because she avoids the costs of revoking authority and of implementing the action when the agent generates a signal (see (15)). Hence, the type H principal provides higher training when she opts for delegation than when she retains authority.

Lemma 2 also provides insightful results regarding the agent's effort choice. First, the agent's effort choice does not depend on the organizational mode when he is hired by the type L principal. The intuition behind this result is very simple: because the agent anticipates that his authority, if granted in the first place, will never be revoked by the principal (see Lemma 1) and because he is provided with the same level of training under delegation and under centralization (see i) in Lemma 2), both the marginal benefit and the marginal cost of effort are the same under either organizational mode. Hence, the agent exerts the same level of effort when he is hired by the type L principal.

In contrast, the agent always exerts higher effort when he is hired by the type H principal and increases his effort provision even more when he is delegated authority by this type of principal. To see the logic, consider the agent's incentives for effort when he is hired by the type H principal. Under centralization, generating a signal reduces both the uncertainty about the state (this source of incentive always exists regardless of the principal's type and the organizational mode) and the disagreement between the agent and the type H principal. Thus, the agent exerts higher effort even when the type H principal retains authority than the case in which he is hired by the type L principal. Under delegation, the agent's effort incentives are even stronger because he retains authority and avoids the principal's biased decision if he manages to generate a signal. Further, as i) of Lemma 2 indicates, the type H principal also encourages effort under delegation by providing higher training than it provides under centralization.

To derive the principal's optimal delegation decision, we consider how her expected payoff changes by delegating authority. That is, the principal of type q, $q = L, H$, will delegate authority to the agent rather than centralizing the decision making if and only if the relative attractiveness of the former vis-à-vis the latter, $\Delta_q^d = U_{P_q}^d\left(t_q^d; \psi_q^d\left(t_q^d\right)\right) - U_{P_q}^c\left(t_q^c; \psi_q^c\left(t_q^c\right)\right)$, is positive. The following result shows how a principal's rents from delegation change with some parameters of the model.

Lemma 3. *Δ_L^d and Δ_H^d satisfy the following conditions:*

i) $\Delta_L^d < 0$.

ii) Δ_H^d increases with σ_0^2 and decreases with σ_s^2.[18]

The logic behind i) is quite simple. Because the principal types are publicly observed and the type L principal never revokes the agent's authority, she cannot elicit higher effort from the agent by granting him the authority to make a decision (i.e., $\psi_L^d\left(t_L^d\right) = \psi_L^c\left(t_L^c\right)$ as stated in Lemma 2). This means that by delegating authority to the agent, the principal avoids the implementation cost at the expense of incurring the disutility caused by the agent's biased decision. Because the implementation cost is small, by assumption, the type L principal does not earn any rents from delegating authority, i.e., $\Delta_L^d < 0$.[19]

Next, we consider how the incentives of the type H principal to delegate authority are affected as the uncertainty in the environment changes. To this end, we make three observations concerning the effect of σ_0^2 on Δ_H^d. First, as the environment becomes more uncertain (i.e., as σ_0^2 increases), making a decision based on a signal is more beneficial for the principal not only because the reduction in the uncertainty concerning the state is larger (that is, $\sigma_0^2 - \widetilde{\sigma}^2$ increases with σ_0^2) but also because the difference of opinion between the principal and the agent is less pronounced (since B_s decreases with σ_0^2). Hence, since the agent exerts higher effort under delegation (Lemma 2), holding the agent's effort and training levels constant, the principal's rents from delegating authority increases with σ_0^2. Second, as σ_0^2

[18]Sufficient but not necessary conditions for this result are $\left[\frac{\partial^2 g(\psi,t)}{\partial\psi\partial t}\right]^2 - \frac{\partial^2 g(\psi,t)}{\partial\psi^2}c''(t) > 0$ for any $t > 0$ and $c'(.)$ is not large.

[19]Specifically, we assume that $\kappa < b^2$.

increases, the principal increases her training provision more under delegation than under centralization. Third, higher initial uncertainty has two effects on the agent's effort choice. On the one hand, the agent has stronger effort incentives under delegation because, as indicated, the principal increases training provision more when she opts for delegation. On the other hand, because the interim bias is lower, the agent's incentives for effort increases more under centralization. As we show in the Appendix, the first effect dominates the latter as long as the complementarity between training and effort is sufficiently large. Taken together, these three observations imply that the type H principal's rents from delegation increase with the initial uncertainty.[20]

Not surprisingly, the principal's rents from delegation decrease as the signal becomes less informative (i.e., as σ_s^2 increases). To see why the signal precision and the initial uncertainty affect the returns from delegation in opposite directions, consider how they alter the interim bias. As σ_s^2 increases, the players assign a lower weight to the signal in their posteriors. This also means that the weight on their priors, over which the type H principal and the agent have a disagreement, increases. Thus, the magnitude of the interim bias increases with the variance of the signal. The effect of the initial uncertainty on the interim bias is the opposite: as σ_o^2 increases, the weight on the signal increases, and consequently, the interim bias decreases. This explains why the effects of σ_0^2 and σ_s^2 on delegation rents act in opposite ways. In sum, as the signal becomes less informative, the principal reduces training provision more under delegation than under centralization, whereas the agent lowers his effort more under delegation than under centralization.[21] Hence, the principal earns lower rents from delegation as the signal become less informative.

As Lemma 3 shows, the type H principal earns positive rents from delegation for some values of σ_0^2, and σ_s^2. Let Γ be the set of (σ_0^2, σ_s^2) such that $\Delta_H^d > 0$. The following proposition describes the equilibrium.

Proposition 1. *Suppose that the principal's type is public information. The equilibrium behavior is described by i) through iii):*

i) The type L principal retains authority and provides training t_L^c.

[20] We should also note that because the principal increases her training provision more under delegation than under centralization, the cost of training also increases more in the former case. Clearly, this effect reduces the principal's rents from delegation. As discussed in the appendix, we always assume that the effect of training cost is not sufficiently large to alter principals' optimal choices for organizational mode.

[21] In addition to increasing the interim bias, lower signal precision also decreases the reduction in the uncertainty about the state. This effect also reduces the type H principal's incentive to delegate authority.

ii) The type H principal delegates authority if and only if $(\sigma_0^2, \sigma_s^2) \in \Gamma$. *She provides training* t_H^d *under delegation and training* t_H^c *under centralization.*

iii) When the agent is hired by the type H principal, he exerts effort $\psi_H^d\left(t_H^d\right)$ *if he is delegated authority and* $\psi_H^c\left(t_H^c\right)$ *if not. When the agent is hired by the type L principal, he exerts effort* $\psi_L^c\left(t_L^c\right)$.

To complete our discussion of the symmetric-information case, we look at how equilibrium training and effort choices change with the organizational mode.

Corollary 1. *The equilibrium levels of training and effort satisfy the following conditions:*

i) The agent is provided with a higher level of training under delegation.

ii) The agent exerts higher effort when he is delegated authority.

The above corollary summarizes the results in the symmetric-information case. In equilibrium, the level of training provided by the principal is higher under delegation. This implies that we expect a positive correlation between the intensity of training given to a worker and the degree of delegation granted to the same worker. We also expect higher effort provision by the worker when she is delegated authority to make decisions.

3.3.2 Asymmetric Information

We now turn to the asymmetric information case in which the type of the principal is her private information. The main objective of this exercise is to examine the extent to which principals can use training provision to signal their types to the agent. To this end, we first seek a separating equilibrium in which the type H principal provides additional training (i.e., more than she provides in the symmetric-information equilibrium) to convey her type to the agent. We then describe the pooling equilibria in which both types of principals choose either delegation or centralization, and provide the same level of training under either organizational mode.[22]

As discussed in the symmetric-information case, the agent exerts higher effort when he is delegated authority by the type H principal than when he is hired by the type L principal

[22] We focus on perfect Bayesian equilibria (PBE) of the model in which beliefs are derived from Bayes' rule given equilibrium strategies and equilibrium strategies are optimal for the principal and the agent given their beliefs. See Fudenberg and Tirole (1991) for a formal definition.

(see Lemma 2). This means that the type H principal has an incentive to signal her type to the agent, whereas the type L principal has an incentive to mimic the other type.

To examine the existence of a separating equilibrium, suppose that the type H principal delegates authority and provides training $\widetilde{t}^d$, whereas the type L principal retains authority and provides training $\widetilde{t}^c$. Because the agent does not observe which type of principal he is hired by, he uses the principal's organizational mode and training choices to form a belief concerning her type. Specifically, let $\lambda(m,t)$ denote the agent's belief of the probability that the principal who chooses organizational mode m, $m = c, d$, and provides training t is of type H, i.e., $\lambda(m,t) = \Pr(q = H|m,t)$. In a separating equilibrium with $\widetilde{t}^d > \widetilde{t}^c$, the agent's equilibrium belief concerning the type of a principal must satisfy $\lambda\left(d, \widetilde{t}^d\right) = 1$ and $\lambda\left(c, \widetilde{t}^c\right) = 0$. That is, the agent assesses that the principal who delegates authority and provides training $\widetilde{t}^d$ will be of type H with probability one, and that the principal who retains authority and provides training $\widetilde{t}^c$ will be of type L with probability one.

As a first step, we assume that the type L principal envies the type H principal at the symmetric-information equilibrium.[23] This means that the equilibrium described in Proposition 1 cannot be an equilibrium in the asymmetric-information case simply because the type L principal would find it profitable to increase her training to t_H^d to fool the agent that she is of type H. Therefore, when types are not observed by the agent, the type H principal must provide more training than she provides in the symmetric-information equilibrium (i.e., $\widetilde{t}^d > t_H^d$) to credibly convey her type to the agent. More specifically, the type H principal must increase her training provision to a level at which the type L principal does not find it profitable to mimic the other type. This restriction yields the incentive compatibility condition for the type L principal:

$$U_{P_L}^d\left(\widetilde{t}^d; \psi\left(\widetilde{t}^d, \lambda\left(d, \widetilde{t}^d\right)\right)\right) \leq U_{P_L}^c\left(t_L^c; \psi\left(t_L^c, \lambda\left(c, t_L^c\right)\right)\right) \tag{16}$$

Similarly, the type H principal must have no incentive to deviate to an alternative strategy. In particular, she must not find it profitable to retain authority and provide training t_L^c:

$$U_{P_H}^d\left(\widetilde{t}^d; \psi\left(\widetilde{t}^d, \lambda\left(d, \widetilde{t}^d\right)\right)\right) \geq U_{P_H}^c\left(t_L^c; \psi\left(t_L^c, \lambda\left(c, t_L^c\right)\right)\right). \tag{17}$$

[23] Under certain parametric assumptions, it is possible to have the symmetric-information equilibrium as a separating equilibrium. However, this "no-envy" equilibrium is not interesting as it assumes away our objective for analyzing the asymmetric-information case in the first place.

Conditions (16) and (17) constitute necessary conditions for a separating equilibrium to exist. We show that it is not possible to find $\widetilde{t}^d$ satisfying both conditions.

Proposition 2. *There exists no separating equilibrium.*

To see the intuition behind this result, consider how the type H principal may use training provision to credibly convey her type to the agent. As shown in the Proof of Proposition 2, conditional on delegating authority, the type H principal can increase her training provision to a level at which the type L principal does not find it optimal to provide; in particular, at that level of training, the type L principal earns zero rents if she chooses to pretend being of type H (i.e., condition (16) holds with equality). However, while imposing enough costs on the type L principal, the type H principal forgoes some rents from delegating authority. More specifically, the additional training cost she imposes on the type L principal equals the rents due to the reduction in the uncertainty that the type H enjoys when she opts for delegation. However, achieving lower uncertainty by delegating authority is the major reason for the type H having an incentive to signal her type. In other words, if she cannot reap any rents from lower uncertainty under delegation, she has no incentive to delegate authority to the agent in the first place. Hence, at the level of training for which the type L has no incentive for imitation, the type H principal has an incentive to deviate to centralization and to providing training t_L^c, i.e., conditions (16) and (17) do not hold at the same time. Hence, there is no separating equilibrium.

To complete the analysis, we next consider pooling equilibria of the model. Before stating the equilibrium formally, we consider why a pooling equilibrium under delegation may arise. As the analysis of the symmetric-information equilibrium shows, delegating authority may be optimal for the type H principal because the agent, anticipating that he can avoid his authority being taken back by the principal if he generates a signal, exerts higher effort. In contrast, the type L principal has no incentive to delegate authority because she cannot elicit higher effort from the agent under delegation. In a pooling equilibrium where both types of principals delegate authority and provide the same level of training, the agent cannot perfectly infer the type of the principal he faces. Therefore, he uses his prior belief, γ, in selecting the optimal effort to exert. This means when the agent cannot observe the type of the principal, delegating authority becomes more attractive for the type L principal, whereas the type H principal's rents from delegation decrease relative to those she earns in the symmetric-information equilibrium. Also, note that the agent exerts higher effort as the probability of facing the type H principal increases. Hence, for sufficiently high values of γ, delegation becomes optimal for either type of principal.

The following Proposition formally describes the equilibrium behavior.

Proposition 3. *Suppose that the type of a principal is her private information. The equilibrium behavior is characterized by threshold $\gamma^*(\sigma_0^2, \sigma_s^2)$ and described by through i)-iii):*

i) If $\gamma \geq \gamma^*(\sigma_0^2, \sigma_s^2)$, both types of principals delegate authority and provide training t_H^d. The agent exerts effort ψ^d that satisfies $\psi_H^d(t_H^d) > \psi^d(t_H^d) > \psi_L^c(t_L^c)$.

ii) If $\gamma < \gamma^*(\sigma_0^2, \sigma_s^2)$, both types of principals retain authority and provide training t_L^c. The agent exerts effort ψ^c that satisfies $\psi_H^c(t_H^c) > \psi^c(t_L^c) > \psi_L^c(t_L^c)$.

iii) $\gamma^*(\sigma_0^2, \sigma_s^2)$ decreases with σ_0^2 and increases with σ_s^2.

As shown in Lemma 2, the rents from delegation increase with σ_0^2. The main reason behind this result is that as the initial uncertainty increases, the principal provides higher training and the agent, as a response, exerts higher effort. These changes, in turn, increase the attractiveness of delegation, vis-à-vis centralization, for the type H principal. In a pooling equilibrium under delegation, delegation becomes less profitable for either type of principal as γ decreases because the agent lowers his effort as he becomes more likely to face the type L principal. In particular, for given levels of σ_0^2 and σ_s^2, the type H principal becomes indifferent between delegation and centralization when γ is sufficiently small. At that critical value of γ, denoted by $\gamma^*(\sigma_0^2, \sigma_s^2)$ in the above proposition, consider the effect of an increase in σ_0^2. As the initial uncertainty increases, the rents from delegation increase and the tie between delegation and centralization is broken in favor of the former. This means that the critical value at which the type H principal is indifferent between the two organizational modes decreases with σ_0^2. A similar intuition explains why threshold $\gamma^*(\sigma_0^2, \sigma_s^2)$ increases with σ_s^2.

3.4 Empirical Implications

The main prediction of our theory is that delegation is associated with higher levels of training. Our analysis of the symmetric-information case shows that all firms provide training (see Lemma 2), but in equilibrium those firms which delegate authority provide more training than the others (see Corollary 1). In the asymmetric-information case, we observe that either all firms delegate authority or they retain it, but in the former case they provide higher levels of training (see Proposition 3). Hence, we expect a positive correlation between delegation and training.

Our theory also yields testable predictions concerning how the uncertainty in the environment and the precision of the worker's information are related to the firm's propensity to delegate authority. In the symmetric-information case, we show that the rents from delegation increase with the uncertainty in the environment and the precision of the worker's signal (see Lemma 3). In equilibrium, we observe that a given firm delegates authority if and only if the firm operates in a sufficiently uncertain environment or the information the worker can acquire is sufficiently precise (see Proposition 1). In the asymmetric-information case, the greater uncertainty in the environment (or the greater precision of the worker's information) leads to a higher likelihood of delegation through two channels. First, as indicated in Proposition 3, the threshold level of prior belief that a given firm is of high type decreases with the uncertainty in the environment (and with the precision of the worker's signal). This means that holding other factors constant, a pooling equilibrium in which all firms delegate is more likely to occur as the uncertainty in the environment (or the precision of the worker's signal) increases. Second, as in the symmetric-information case, a firm's rents from delegation increase with the uncertainty in the environment and the precision of the worker's signal. Therefore, a given firm's propensity to delegate authority increases with these two parameters.

In summary, we test the following hypotheses: i) delegation and training are positively related; ii) holding other factors constant, a firm is more likely to delegate as the uncertainty in the environment increases; iii) holding other factors constant, a firm is more likely to delegate as the worker's signal becomes more informative. To take these predictions to the data, in the next section we discuss how we proxy for the uncertainty in the environment and the precision of the worker's signal.

4 Empirical Analysis

This section provides our empirical analysis. In the first part of the section, we describe our data and the key variables used in the regression analyses. Afterwards, we discuss the empirical results.

4.1 Description of Data and Key Variables

Our data is drawn from the 2004 wave of the Workplace Employee Relations Survey (WERS), a large, nationally-representative cross-section of British workplaces.[24] In each wave of the survey, a sample of workplaces is randomly drawn from the Interdepartmental Business Register, maintained by the Office of National Statistics. The surveys are stratified by workplace size and industry with larger workplaces and some industries being overrepresented (Chaplin et al., 2005). Consequently, sampling weights for each section of the survey are provided to remove the bias resulting from the sample selection procedures and non-response. The survey population is all British workplaces (with 5 or more employees) except those in primary industries and private households with domestic staff. This dataset is suitable for our empirical analysis because it includes information about delegation and training while providing information on a wide range of employee and employer characteristics that can not only serve as important controls but also be used to proxy for the uncertainty in the environment and the signal precision.

The survey comprises three main sections: the 'Management Questionnaire', the 'Worker Representative Questionnaire' and the 'Employee Questionnaire'. We use information from employee and management questionnaires. The management questionnaire was administered by a face-to-face interview with the most senior manager who had day-to-day responsibility for industrial relations or personnel matters. For the employee questionnaire, a sample of 25 employees (or all employees if the total number of the employees in the workplace is fewer than 25) was randomly selected at all workplaces participating in the management survey.[25] The employee questionnaire was self-administered by the employee without direct interviewer involvement. The response rate is 64% in the management questionnaire and 60% in the employee questionnaire.

Our delegation measure comes from the employee questionnaire. At each workplace, the surveyed employees are asked: "*In general, how much influence do you have about the range of tasks you do in your job?*" Responses are recorded on a four-point scale: 'None', 'A little', 'Some' and 'A lot'.[26] Despite being subjective, this measure is employed to proxy for the delegation of authority to workers (e.g., DeVaro and Kurtulus, 2010; DeVaro and Prasad,

[24] In our discussion, we use "workplace", "establishment" and "firm" interchangeably.

[25] Due to the nature of the multistage sampling procedure used to survey employees, employee questionnaires were distributed only in those workplaces where a management interview had taken place.

[26] The distribution of responses to this question (after applying employee weights) is as follows: 'None' 12.7%, 'A little' 14.9%, 'Some' 36.9%, and 'A lot' 35.4%.

2015; Bilanakos et al., 2017).[27] Our main delegation measure takes on a value of one if the worker's response is 'A lot', 'Some' or 'A little' and zero if the worker's response is 'None'. To measure the amount of training a worker has received at his or her workplace, we use the following question from the employee questionnaire: "*During the last 12 months how much training have you had, either paid for or organized by your employer: include only training away from your normal place of work, but it could be on or off the premises*". The potential responses are "None", "Less than one day", "1 to less than 2 days", "2 to less than 5 days", "5 to less than 10 days" and "10 days or more".

To test the model's predictions, we need to proxy for the uncertainty in the environment (σ_0^2) and the precision of the worker's signal (σ_s^2). To measure the uncertainty in the environment, we use binary indicators capturing the current state of the market in which the employee's workplace operates. Specifically, the manager's response indicates whether the market is growing, declining, turbulent or mature. We expect that the firm's environment is more uncertain if the market is either growing, declining or turbulent than if it is mature. To proxy for the precision of the worker's signal, we use variables correlated with the worker's productivity. In our model, the worker could be regarded as more productive either as he becomes more likely to acquire a signal (holding the cost of acquiring the signal constant) or as his signal becomes more informative. We conjecture that the worker's age, tenure at the firm, and education level are positively correlated with the precision of the signal he may acquire. In addition, we proxy for the match quality between the worker's skills and his current job using the following question from the employee questionnaire: "*How well do the work skills you personally have match the skills you need to do your present job?*". The worker is described as overqualified if his response is either "Much higher" or "A bit higher", as qualified if his response is "About the same", and as underqualified if his response is either "A bit lower" or "Much lower". We conjecture that the worker's signal is more precise if he is either overqualified or qualified.

Our empirical analysis excludes workplaces not in the trading sector (government and non-profit establishments) and those observations for which any variables used in the analysis are missing. Merging the employee questionnaire with the management questionnaire, and

[27] Despite using the same source of data, Acemoglu et al. (2007) employ a different measure for delegation. The reason is that in addition to the data from the WERS, Acemoglu et al. (2007) also use two large datasets of French firms. To be consistent with their analysis of the French data (and to some extent with their theory), the authors measure delegation using a question concerning the manager's autonomy from headquarters in making employment decisions.

treating the data at the worker level leaves us with a sample consisting of 10,983 workers clustered in 920 workplaces. Summary statistics are reported in Table 1.

4.2 Results

To test the predictions of the model, we first estimate probit models where the dependent variable is an indicator variable that takes a value of one if the worker is granted authority and zero if not. Results are reported in Table 2.

Our baseline specification, reported in column 1, includes a binary variable for training, the worker's age and tenure (both with squared terms), indicator variables for education (the omitted category is 'no academic qualifications or other qualification'), job-skill match (the omitted category is underqualified), and indicators capturing the current state of the market in which the workplace operates (the omitted category is operating in a mature market). In addition to these key variables, the baseline specification includes the following controls: firm size (the log of the number of employees), indicators for how many years the workplace has been operating (the omitted category is 'more than 25 years'), if the worker has a permanent or a temporary job with no agreed end date (the omitted category is having a fixed period job with an agreed end date), whether the workplace is a part of a larger organization (the omitted category 'sole UK establishment of a foreign organization), and finally indicators for industry and region.

In column 2, we add indicators for the worker's occupation (manager, professional, associate professional and technical, clerical and secretarial, craft and skilled service, personal and protective service, sales, plant and machine operatives, where the omitted category is 'other occupation'). In column 3, we add indicator variables capturing the occupational composition of the workforce (i.e., the percentage of: managers, professional staff, technical staff, clerical staff, craft (skilled) staff, service staff and sales staff, where the omitted category is the share of 'other occupations'). The specifications in columns 4 through 6 are the same as those in columns 1 through 3, respectively, except in the latter set of specifications, we use six categories, rather than a binary variable, for training. In these richer specifications, the omitted category for training is "None".

We first discuss the relationship between delegation and training. As column 1 indicates, the coefficient for training is positive and statistically significant at the one percent level. When we control for the worker's occupation and the occupational composition at the workplace, the coefficient for training decreases by 31 percent (from 0.228 to 0.155) but remains statistically significant at the one percent level. To get a sense of the magnitude of the effect, consider the coefficient reported in the baseline specification. Accordingly, providing training with a worker is associated, on average, with a 2.8 percentage points increase in the predicted probability that the same worker is delegated authority. Once indicators for occupations are added, the positive effect of training on the predicted probability of being delegated authority decreases from 2.8 percentage points to 1.9 percentage points.

As indicated, we evaluate a less restricted specification for training in columns 4 through 6, where six exclusive categories measure the intensity of training provided by the firm. From these results, we observe three patterns. First, the effect of receiving training less than one day is not statistically different from the effect of not receiving training at all. Even though the coefficients for obtaining training less than a day are positive, they are not estimated with precision. Second, the probability of delegation monotonically increases with training up to training equal to 5 days, but the coefficients for higher training categories are smaller. However, according to the results in column 4, the coefficients for adjacent training categories for levels of training higher than 5 days are not statistically different from each other. Therefore, a more precise statement to describe the relationship is that the probability of delegation monotonically increases with training for levels of training between 1 day and 5 days, and then the positive effect flattens out with additional training. Third, in comparison to the baseline results, coefficient estimates for training categories become smaller as controls for occupations and the occupational composition of the workplace enter the estimating equation.[28]

We now turn to the model's second prediction, that is, a firm becomes more likely to delegate authority as the uncertainty in the environment increases. We consider three binary variables indicating whether the market at which the firm currently operates is growing, declining or turbulent (recall that the benchmark category is market being mature). The results provide mixed support for the prediction. The results from the baseline specifications (columns 1 and 4) show that operating in a turbulent market, as opposed to operating in a mature market, is positively associated with the probability of granting delegation. The coefficients slightly increase and remain statistically significant (at the five percent level) when controls for occupations are added (see columns 2 and 5). In contrast, the coefficients decrease (by 21 percent) but remain positive and statistically significant at the ten percent

[28]The only exception to this pattern is training category for 5 to 10 days. This coefficient decreases from 0.292 to 0.149 when controls for occupations are included and then increases to 0.150 when controls for workforce composition are added.

level when controls for occupational composition are added (see columns 3 and 6). The coefficients in columns 1 and 3 indicate that operating in a turbulent market, as opposed to operating in a mature market, is associated with 2.1 and 1.7 percentage points, respectively, higher likelihood of delegation. The coefficients for growing markets are positive in all specifications, but statistically significant (at the ten percent level) only in columns 2 and 5. These coefficients indicate that operating in a growing market, as opposed to in a mature market, is associated with 1.4 percentage points higher probability of delegation. Finally, even though the coefficients for declining markets are positive in all specifications, they are not statistically different from zero at conventional levels.

Turning to the third prediction, we first consider the effect of educational level on delegation. Column 1 and 4 show that there is no monotonic relationship between the worker's highest educational attainment and the probability of delegation. More specifically, holding an A-level degree and lower has no effect on delegation (coefficients are small and not statistically different from zero), whereas holding either a college degree or a postgraduate degree, as opposed to holding either a lower educational qualification or no degree at all, is associated with a higher likelihood of delegation. Also, even though the coefficient for postgraduate degree is larger than the coefficient for college degree, the difference is not statistically different from zero. These effects, however, turn insignificant when the worker's occupation is controlled for (see, for example, columns 2 and 3). The reasoning behind this result is simple. Because the variation of education levels in a given occupation is not large, we observe that the positive effect of holding a college degree or higher disappears once indicators for occupations enter the specification. Consistent with this reasoning, we also observe that coefficients for certain occupations, such as manager, professional, technical and clerk, are positive and estimated with high precision (all these coefficients are statistically significant at the one percent level).

The results concerning the effect of age and tenure provide support for the third prediction. Since in our specifications we include both age and age squared (divided by 100 for convenience) as explanatory variables, the effect of age on the probability of delegation depends on at what age level the marginal effect is evaluated. We observe qualitatively the same result in all specifications: the coefficient for age is positive and the coefficient for its squared term is negative (and both coefficients are statistically significant at the one percent level). This pattern indicates that the probability of delegation increases (at a decreasing rate) with age and the effect turns zero after a certain age. For example, according to the results from column 6, the effect of age on the probability of delegation becomes zero after around age 36. Similar to age, we include both tenure at firm and its squared term in our estimations. The results show that the coefficient for tenure is positive (and statistically significant at the one percent level) in all specifications, whereas the coefficient for its squared term is negative in all specifications but statistically significant (at the ten percent level) only in columns 2 and 4.[29] According to the coefficients in columns 2 and 4, the effect of tenure on delegation is positive up to tenure for 20 years.

As indicated, we conjecture that higher match quality between the worker's skills and his or her job is associated with higher probability of delegation. The results provide support for this hypothesis. In all specifications, the coefficients for overqualified and qualified are positive and statistically significant at the one percent level, and the difference between the marginal effects of being overqualified and qualified is statistically significant.[30] This implies that the effect of being qualified on the probability of delegation is greater than that of being overqualified. According to the coefficients reported in column 3, being qualified, as opposed to being underqualified, is associated with a 5.5 percentage points increase in the predicted probability of delegation, whereas being overqualified with respect to the same benchmark is associated with a 4 percentage points increase in the predicted probability of delegation.

As our delegation measure is derived from a question in which potential responses have a meaningful (ordinal) ranking, we also estimate an ordered probit model where the dependent variable describes the degree of delegation the worker has been granted with four distinct categories. The results are reported in Table 3. Note that specifications in columns 1 through 4 of Table 3 are analogous to those in columns 3 through 6 of Table 2, respectively.

The results concerning the relationship between training and delegation remain consistent with the theory. When training is measured by a binary variable, the coefficient is positive and statistically significant at the one percent level (see column 1). This means that the intensity of delegation (which is modelled as a latent variable) increases with training. The incremental effect of obtaining training (i.e., when training increases from 0 to 1) is associated with a 9.2 percentage points decrease in the probability that delegation is not granted (i.e., the worker's response is 'None').[31] In columns 2 through 4, we observe that the relationship between obtaining training less than one day and delegation is not statistically significant,

[29] We also estimated our richest specification (column 6) by excluding tenure squared. In that case, the coefficient for tenure turns out to be 0.019 and it is statistically significant at the one percent level.

[30] The p-values associated with these tests are lower than 0.01 in all specifications.

[31] The incremental effect is statistically significant at the one percent level ($z = 3.33$).

whereas obtaining training more than one day is positively related with training. Further, the coefficient for training equal to 1 to 2 days is smaller than the coefficient for training equal to 2 to 5 days in all specifications, but the difference is statistically significant only in columns 2 and 3.[32] Finally, adjacent training categories for levels of training higher than 5 days are not statistically different from each other.

Next, we look at the effect of the uncertainty in the environment, captured by indicators for the state of the market. Consistent with the earlier results, the coefficients for these indicator variables remain positive, as predicted by the theory, but not all of them are statistically significant. We observe the following differences between the results from Table 3 and those from Table 2. First, the coefficients for operating in a turbulent market remain positive, but become statistically not different from zero (this happens because the size of the coefficient decreases by approximately 50 percent, whereas the standard error decreases only marginally). Second, despite being marginally smaller than in the earlier set of results, the effect of operating in a growing market is positive and statistically significant (at the ten percent level in columns 1, 2, and 4 and at the five percent level in column 3).

Finally, we turn to how delegation is related to our proxies for the precision of the worker's signal. We observe that the qualitative results concerning age and tenure at the firm are the same. That is, the main effect is positive, whereas the coefficient for the quadratic term is negative (both coefficients are significant at the one percent level in all specifications). Concerning the effect of educational attainment, we observe the same pattern except for holding a GCSE degree (grades A-C). More precisely, we observe that holding a college degree and holding a postgraduate degree are associated with a higher likelihood of delegation (see column 1); these effects, however, disappear once we control for occupational dummies (as we observe in the probit model). The difference from the earlier set of results is that the coefficient for GCSE degree (grades A-C) turn significant (at the one percent level) in ordered probit models. Last, we obtain qualitatively the same results concerning the effect of the match quality between the worker's skills and his or her job. Overall, the results suggest that the correlation between being either qualified or overqualified and delegation is positive.[33]

[32] More specifically, the p-value for the difference between training categories "one to two days" and "two to five days" is 0.060 in column 2 and 0.630 in column 3.

[33] Note that the coefficient for qualified is greater than the coefficient for overqualified, but the difference is significant at modest levels. Specifically, the p-values from the corresponding tests are 0.122, 0.092 and 0.061 in columns 2, 3, and 4, respectively.

5 Conclusion

In this paper, we undertake a theoretical and empirical investigation of how the firm's incentives to provide training and to delegate authority interact. Two aspects of our setting help us provide insights concerning the firm's choices on training provision and allocation of authority. First, because delegation decision is not contractible, the firm has an incentive to retract the worker's authority after the outcome of the worker's investigation is observed. This raises a commitment problem from the firm's standpoint. Second, the conflict of interest between the firm and the worker arises both because they have different underlying preferences and because they interpret information differently. The presence of differences of opinion alters the incentives of the worker to acquire information and those of the firm to provide training and delegate authority.

Our main analysis focuses on the case in which the firm and the worker are symmetrically informed about the level of incongruity in the preferences of the firm and those of the worker. We show that in equilibrium firms that have a difference of opinion with the worker delegate authority if the uncertainty in the environment is sufficiently high or if the worker's information is sufficiently precise, whereas other firms find it optimal to retain authority. Further, firms that delegate authority also provide more training than those firms which retain authority. In addition to the positive correlation between training and delegation, comparative statics results yield testable predictions. Accordingly, the model predicts that firms should be more likely to delegate authority as the uncertainty in the environment increases and as the information the worker may acquire becomes more precise.

To analyze the commitment problem faced by the firm, we also consider the asymmetric-information case in which the firm is privately informed about its type wherein the type refers to whether the firm has a difference of opinion with the worker. In this case, we show that firms with a difference of opinion could use training provision to credibly signal to workers that they are not going to retract their authority. However, in equilibrium these firms do not find it optimal to do that; instead, we observe pooling equilibria in which either all firms delegate authority or they retain it.

In the empirical part of the paper, we test the model's predictions and provide evidence broadly consistent with the model. First, we find strong evidence that providing higher levels of training is associated with higher likelihood of delegation. Second, we find evidence that firms operating in turbulent markets are more likely to delegate than firms operating in mature markets. Similarly, the effect of operating in a growing market, as opposed to operating in a mature market, is positive, but it is estimated with less precision. These findings provide support for the model to the extent that the state of the market measures the uncertainty faced by the firm. Finally, we proxy for the precision of the worker's signal using the worker's age, tenure at the firm, highest educational attainment and the quality of the match between the worker's skills and his or her job. In general, the empirical results concerning these variables are consistent with the model. The effect of age on delegation is positive up to certain age, and then it flattens out. We observe an analogous pattern for tenure at firm. Concerning education, we find that holding a college degree or higher is associated with higher likelihood of being delegated authority. Last, we find strong evidence that being either qualified or overqualified (measured by the job-skills match quality) are positively related to the probability of delegation. Overall, to the extent that these variables measure the precision of the worker's information, our results show that delegation is more likely to be granted to the worker as the worker's signal become more precise.

A Appendix

This Appendix contains the proofs omitted in the text. To simplify the notation, we let $\frac{\partial^2 g(\psi,t)}{\partial\psi^2} = g_{\psi\psi} > 0$ and $\frac{\partial^2 g(\psi,t)}{\partial\psi\partial t} = g_{\psi t} < 0$ for any ψ and t. Before presenting the proofs, we first provide a result that will be useful to show some of the results.

Remark 1. *Let m and q denote the organizational mode and the principal type, respectively, where , $m = c, d$ and $q = L, H$. We have $\frac{\partial\psi_q^m(t)}{\partial t} = -\frac{g_{\psi t}}{g_{\psi\psi}} > 0$.*

Proof of Remark 1. *Fix m and q. We write the first-order condition that characterize the agent's optimal effort choice as follows:*

$$\Lambda(.) = \frac{\partial g\left(\psi_q^m, t_q^m\right)}{\partial\psi_q}, \tag{A.1}$$

where $\Lambda(.)$ is the marginal benefit of effort and it is a function of the model's parameters. Using the implicit function theorem, we write the optimal effort as a function of t, that is, $\psi_q^m = \psi_q^m(t)$. Differentiating (A.1) with respect to t gives $\frac{\partial\psi_q^m(t)}{\partial t} = -\frac{\partial^2 g\left(\psi_q^m,t\right)/\partial\psi_q\partial t}{\partial^2 g\left(\psi_q^m,t\right)/\partial\psi_q^2} = -\frac{g_{\psi t}}{g_{\psi\psi}} > 0$, since $g_{\psi t} < 0$ and $g_{\psi\psi} > 0$.

Proof of Lemma 1. *Suppose that the type H principal initially delegates authority to the agent who then generates a signal, i.e., $\Omega_{P_H} = s$. The principal's payoff is $-\widetilde{\sigma}^2 - B_s^2$ if she lets the agent exercise his authority and $-\widetilde{\sigma}^2 - R - \kappa$ if she takes the authority back from the agent and chooses her preferred action. Therefore, conditional on observing a signal, the type H principal finds it optimal to retract the agent's authority if and only if $R < R_H^*(\Omega_{P_H} = s)$, where $R_H^*(\Omega_{P_H} = s) = B_s^2 - \kappa$. Similarly, when the agent does not generate a signal, the same type of principal finds it optimal to revoke delegation decision if and only if $R < R_H^*(\Omega_{P_H} = \emptyset)$, where $R_H^*(\Omega_{P_H} = \emptyset) = B_\emptyset^2 - \kappa$.*

Next, consider the type L principal's decision to revoke authority. Suppose that the agent generates a signal, i.e., $\Omega_{P_L} = s$. The principal's payoff is $-\widetilde{\sigma}^2 - b^2$ if the agent chooses an action and $-\widetilde{\sigma}^2 - R - \kappa$ if the principal takes the agent's authority back and chooses the action to be implemented. Therefore, conditional on observing a signal, the type L principal finds it optimal to revoke her delegation decision if and only if $R < R_L^*(\Omega_{P_L} = s)$, where $R_L^*(\Omega_{P_L} = s) = b^2 - \kappa$. Applying a similar argument shows that $R_L^*(\Omega_{P_L} = \emptyset) = b^2 - \kappa$.

To show the last part of the result, it is sufficient to note that $B_\emptyset^2 > B_s^2 > b^2$.

Proof of Lemma 2. *Consider the type L principal's training decision. Writing (8) and (14) explicitly gives*

$$\left[\sigma_0^2 - \widetilde{\sigma}^2\right] \frac{\partial \psi_L^c(t)}{\partial t_L^c} = c'\left(t_L^c\right) \tag{A.2}$$

and

$$\left[\sigma_0^2 - \widetilde{\sigma}^2\right] \frac{\partial \psi_L^d(t)}{\partial t_L^d} = c'\left(t_L^d\right), \tag{A.3}$$

respectively.[34] *Given that* $\frac{\partial \psi_L^c(t)}{\partial t_L^c} = \frac{\partial \psi_L^d(t)}{\partial t_L^d} = -\frac{g_{\psi t}}{g_{\psi\psi}} > 0$ *(see Remark 1) and* $c(.)$ *is strictly convex, we obtain* $t_L^d = t_L^c > 0$. *Next, consider the type H principal's training decisions under both organizational modes. The first-order conditions (8) and (15) yield*

$$\left[\sigma_0^2 - \widetilde{\sigma}^2\right] \frac{\partial \psi_H^c(t)}{\partial t_H^c} = c'\left(t_H^c\right) \tag{A.4}$$

and

$$\left[\left[\sigma_0^2 - \widetilde{\sigma}^2\right] + \left[R + \kappa - B_s^2\right]\right] \frac{\partial \psi_H^d(t)}{\partial t_H^d} = c'\left(t_H^d\right), \tag{A.5}$$

respectively. It follows from (A.4) and (A.2) that $t_H^c = t_L^c$. *Finally, since* $[R + \kappa - B_s^2] > 0$, *we have* $t_H^d > t_H^c$. *This shows i).*

To show ii), we consider the agent's optimal effort choices. Suppose that the agent is hired by the type L principal. Evaluating (6) and (10) at $t = t_L^c$ *and* $t = t_L^d$, *respectively, yields that* $\frac{\partial g\left(\psi_L^c\left(t_L^c\right), t_L^c\right)}{\partial \psi_L} = \frac{\partial g\left(\psi_L^d\left(t_L^d\right), t_L^d\right)}{\partial \psi_L} = \left[\sigma_0^2 - \widetilde{\sigma}^2\right]$. *By the strict convexity of* $g(.)$, *we obtain* $\psi_L^d\left(t_L^d\right) = \psi_L^c\left(t_L^c\right)$. *Next, consider the agent's effort choice when he is hired by the type H principal. Evaluating (6) and (11) at* $t = t_H^c$ *and* $t = t_H^d$, *respectively, gives*

$$\left[\left[\sigma_0^2 - \widetilde{\sigma}^2\right] + \left[B_\emptyset^2 - B_s^2\right]\right] = \frac{\partial g\left(\psi_H^c\left(t_H^c\right), t_H^c\right)}{\partial \psi_H} \tag{A.6}$$

and

$$\left[\left[\sigma_0^2 - \widetilde{\sigma}^2\right] + B_\emptyset^2\right] = \frac{\partial g\left(\psi_H^d\left(t_H^d\right), t_H^d\right)}{\partial \psi_H}, \tag{A.7}$$

respectively. Since $B_\emptyset^2 > B_\emptyset^2 - B_s^2 > 0$ *and* $t_H^d > t_H^c$, *the strict convexity of* $g(.)$ *implies that* $\psi_H^d\left(t_H^d\right) > \psi_H^c\left(t_H^c\right)$. *Finally, it is easy to see that* $\psi_H^c\left(t_H^c\right) > \psi_L^d\left(t_L^d\right)$ *as* $B_\emptyset^2 > 0$.

Proof of Lemma 3. *Consider the type L principal's rents from delegation,* Δ_L^d. *As Lemma*

[34] Note that the strict convexity of the cost function ensures that the sufficiency condition is satisfied for (A.2) and (A.3) to characterize the optimal training levels t_L^c and t_L^d, respectively.

2 indicates, $t_L^d = t_L^c$ *and* $\psi_L^c(t_L^c) = \psi_L^d(t_L^d)$*. Substituting these into* Δ_L^d *yields that* $\Delta_L^d = -b^2 + \kappa$*. Assuming that* $\kappa < b^2$*, we obtain* $\Delta_L^d < 0$*.*

To show ii), we first substitute the optimal effort and training levels into $U_{P_H}^d\left(t_H^d; \psi_H^d\left(t_H^d\right)\right)$ *and* $U_{P_H}^c\left(t_H^c; \psi_H^c\left(t_H^c\right)\right)$ *to write* Δ_H^d *explicitly as:*

$$\Delta_H^d = \left[\psi_H^d\left(t_H^d\right) - \psi_H^c\left(t_H^c\right)\right]\Delta\sigma^2 + \psi_H^d\left(t_H^d\right)\left[R - \left[B_s^2 - \kappa\right]\right] - R - \left[c\left(t_H^d\right) - c\left(t_H^c\right)\right], \quad \text{(A.8)}$$

where $\Delta\sigma^2 = \sigma_0^2 - \widetilde{\sigma}^2$*.*

To derive the effect of σ_0^2 *on* Δ_H^d*, we first look at how choice variables are affected when* σ_0^2 *changes. To this end, we apply the implicit function theorem to (A.4) and (A.5) to obtain* $\frac{\partial t_H^c}{\partial \sigma_0^2} = -\frac{g_{\psi t}}{g_{\psi\psi}}\left[c''\left(t_H^c\right)\right]^{-1}\frac{\partial \Delta\sigma^2}{\partial \sigma_0^2}$ *and* $\frac{\partial t_H^d}{\partial \sigma_0^2} = -\frac{g_{\psi t}}{g_{\psi\psi}}\left[c''\left(t_H^d\right)\right]^{-1}\frac{\partial\left(\Delta\sigma^2 - B_s^2\right)}{\partial \sigma_0^2}$*. This implies that* $\frac{\partial t_H^d}{\partial \sigma_0^2} > \frac{\partial t_H^c}{\partial \sigma_0^2}$*, since* $\frac{\partial\left(\Delta\sigma^2 - B_s^2\right)}{\partial \sigma_0^2} > \frac{\partial \Delta\sigma^2}{\partial \sigma_0^2} > 0$ *and* $c''' \leq 0$*. Similarly, applying the implicit function theorem to (A.6) and (A.7) yields* $\frac{\partial \psi^c\left(t_H^c\right)}{\partial \sigma_0^2} = \left[g_{\psi\psi}\right]^{-1}\left\{\frac{\partial\left(\Delta\sigma^2 + B_\emptyset^2 - B_s^2\right)}{\partial \sigma_0^2} - g_{\psi t}\frac{\partial t_H^c}{\partial \sigma_0^2}\right\}$ *and* $\frac{\partial \psi_H^d\left(t_H^d\right)}{\partial \sigma_0^2} = \left[g_{\psi\psi}\right]^{-1}\left\{\frac{\partial\left(\Delta\sigma^2 + B_\emptyset^2\right)}{\partial \sigma_0^2} - g_{\psi t}\frac{\partial t_H^d}{\partial \sigma_0^2}\right\}$*, respectively. Note that* $\frac{\partial \psi_H^d\left(t_H^d\right)}{\partial \sigma_0^2} > \frac{\partial \psi^c\left(t_H^c\right)}{\partial \sigma_0^2} > 0$*, where the first inequality holds as long as* $g_{\psi t}$ *is sufficiently large in absolute value.*[35] *Considering the direct effect of a change in* σ_0^2 *on* Δ_H^d *gives* $\left[\psi_H^d\left(t_H^d\right) - \psi_H^c\left(t_H^c\right)\right]\frac{\partial \Delta\sigma^2}{\partial \sigma_0^2} - \psi_H^d\left(t_H^d\right)\frac{\partial B_s^2}{\partial \sigma_0^2} > 0$*, since* $\frac{\partial \Delta\sigma^2}{\partial \sigma_0^2} > 0$*,* $\frac{\partial B_s^2}{\partial \sigma_0^2} < 0$*, and* $\psi_H^d\left(t_H^d\right) > \psi_H^c\left(t_H^c\right)$*. That is, ignoring the partial effects of* σ_0^2 *on effort and training choices, the expected rents from delegation increase with* σ_0^2*. Next, we consider the partial effects of* σ_0^2*, i.e.,* $\left[\frac{\partial \psi_H^d\left(t_H^d\right)}{\partial \sigma_0^2} - \frac{\partial \psi_H^c\left(t_H^c\right)}{\partial \sigma_0^2}\right]\Delta\sigma^2 + \frac{\partial \psi_H^d\left(t_H^d\right)}{\partial \sigma_0^2}\left[R - \left[B_s^2 - \kappa\right]\right] - \left[c'\left(t_H^d\right)\frac{\partial t_H^d}{\partial \sigma_0^2} - c'\left(t_H^c\right)\frac{\partial t_H^c}{\partial \sigma_0^2}\right]$*. Note that the first two terms are positive, whereas the last term, the change in the training cost under delegation and under centralization, is negative since* $\frac{\partial t_H^d}{\partial \sigma_0^2} > \frac{\partial t_H^c}{\partial \sigma_0^2} > 0$*. As a sufficient (but not a necessary condition), we assume that the training cost does not rise very quickly, i.e.,* $c'(.)$ *is not very large. Hence,* $\frac{\partial \Delta_H^d}{\partial \sigma_0^2} > 0$*.*

To consider the effect of σ_s^2 *on* Δ_H^d*, we follow the same approach. Applying the implicit function theorem to (A.4) and (A.5) yields* $\frac{\partial t_H^c}{\partial \sigma_s^2} = -\frac{g_{\psi t}}{g_{\psi\psi}}\left[c''\left(t_H^c\right)\right]^{-1}\frac{\partial \Delta\sigma^2}{\partial \sigma_s^2}$ *and* $\frac{\partial t_H^d}{\partial \sigma_s^2} = -\frac{g_{\psi t}}{g_{\psi\psi}}\left[c''\left(t_H^d\right)\right]^{-1}\frac{\partial\left(\Delta\sigma^2 - B_s^2\right)}{\partial \sigma_s^2}$*. This implies that* $\frac{\partial t_H^d}{\partial \sigma_s^2} < \frac{\partial t_H^c}{\partial \sigma_s^2} < 0$*, since* $\frac{\partial\left(\Delta\sigma^2 - B_s^2\right)}{\partial \sigma_s^2} < \frac{\partial \Delta\sigma^2}{\partial \sigma_s^2} < 0$ *and* $c'''(.) \leq 0$*. Next, we apply the implicit function theorem to (A.6) and (A.7) to obtain* $\frac{\partial \psi_H^c\left(t_H^c\right)}{\partial \sigma_s^2} = \left[g_{\psi\psi}\right]^{-1}\left\{\frac{\partial\left(\Delta\sigma^2 + B_\emptyset^2 - B_s^2\right)}{\partial \sigma_s^2} - g_{\psi t}\frac{\partial t_H^c}{\partial \sigma_s^2}\right\}$ *and* $\frac{\partial \psi_H^d\left(t_H^d\right)}{\partial \sigma_s^2} = \left[g_{\psi\psi}\right]^{-1}\left\{\frac{\partial\left(\Delta\sigma^2 + B_\emptyset^2\right)}{\partial \sigma_s^2} - g_{\psi t}\frac{\partial t_H^d}{\partial \sigma_s^2}\right\}$*, respectively. When* $g_{\psi t}$ *is sufficiently large, we have* $\frac{\partial \psi_H^d\left(t_H^d\right)}{\partial \sigma_s^2} < \frac{\partial \psi_H^c\left(t_H^c\right)}{\partial \sigma_s^2} < 0$*. Consid-*

[35] More specifically, the assumption required to have the first inequality is $(g_{\psi t})^2 > g_{\psi\psi}c''(t)$ for any $t > 0$.

ering the direct effect of σ_s^2 *on* Δ_H^d *gives* $\left[\psi_H^d\left(t_H^d\right)-\psi_H^d\left(t_H^c\right)\right]\frac{\partial\Delta\sigma^2}{\partial\sigma_s^2}-\psi_H^d\left(t_H^d\right)\frac{\partial B_s^2}{\partial\sigma_s^2}<0$*, since* $\frac{\partial\Delta\sigma^2}{\partial\sigma_s^2}<0$, $\frac{\partial B_s^2}{\partial\sigma_s^2}>0$*, and* $\psi_H^d\left(t_H^d\right)>\psi_H^c\left(t_H^c\right)$*. We then consider the partial effects of* σ_s^2 *(i.e., the effect of* σ_s^2 *through optimal effort and training choices) and obtain* $\left[\frac{\partial\psi_H^d\left(t_H^d\right)}{\partial\sigma_s^2}-\frac{\partial\psi_H^c\left(t_H^c\right)}{\partial\sigma_s^2}\right]\Delta\sigma^2+\frac{\partial\psi_H^d\left(t_H^d\right)}{\partial\sigma_s^2}\left[R-\left[B_s^2-\kappa\right]\right]-\left[c'\left(t_H^d\right)\frac{\partial t_H^d}{\partial\sigma_s^2}-c'\left(t_H^c\right)\frac{\partial t_H^c}{\partial\sigma_s^2}\right]<0$*. Hence,* $\frac{\partial\Delta_H^d}{\partial\sigma_s^2}<0$.

Proof of Proposition 1. *Suppose that principal types are public information. As Lemma 3 shows, the type L principal does not earn any rents from delegation. Thus, she finds it optimal to retain authority and provide training* t_L^c*. In contrast, the type H principal earns positive rents from delegation for some values of* σ_0^2 *and* σ_s^2 *(see ii) in Lemma 3). In particular, because when* $\Delta_H^d>0$ *for* $(\sigma_0^2,\sigma_s^2)\in\Gamma$*, in those cases, the type H principal delegates authority to the agent and provides training* t_H^d*; otherwise, she retains authority and provides training* t_H^c*. The agent exerts effort* $\psi_H^d\left(t_H^d\right)$ *when he is granted authority by the type H principal and* $\psi_H^c\left(t_H^c\right)$ *when the same type of principal retains authority, whereas he exerts effort* $\psi_L^c\left(t_L^c\right)$ *when he is hired by the type L principal.*

Proof of Corollary 1. *See the Proof of Lemma 2.*

Proof of Proposition 2. *Suppose that the type L principal has an incentive to mimic the type H principal at the full-information equilibrium in which the former retains authority and provides training* t_L^c*, whereas the latter delegates authority and provides training* t_H^d *(see Proposition 1).*

Assume, for contradiction, that there exists a separating equilibrium in which the type H principal delegates authority and provides training $\widetilde{t}^d$*, whereas the type L principal retains authority and provides training* $\widetilde{t}^c$*. Given that* $\widetilde{t}^d>\widetilde{t}^c$*, Bayes' rule implies that* $\lambda\left(m=d,t\geq\widetilde{t}^d\right)=1$ *and* $\lambda\left(m=c,t=\widetilde{t}^c\right)=0$*. Sequential rationality of the players imply that the type L principal provides training* $\widetilde{t}^c=t_L^c$ *(because increasing training does not affect the agent's belief concerning her type and it is therefore not profitable to do so) and that the type H principal increases training to a level at which mimicry is not profitable for the type L principal. The latter condition implies that* $U_{P_L}^d\left(\widetilde{t}^d;\psi\left(\widetilde{t}^d,\lambda\left(d,\widetilde{t}^d\right)\right)\right)=U_{P_L}^c\left(t_L^c;\psi\left(t_L^c,\lambda\left(c,t_L^c\right)\right)\right)$*, i.e.,*

$$\left[\psi\left(\widetilde{t}^d,\lambda\left(d,\widetilde{t}^d\right)\right)-\psi\left(t_L^c,\lambda\left(c,t_L^c\right)\right)\right]\Delta\sigma^2-\left[b^2-\kappa\right]=c\left(\widetilde{t}^d\right)-c\left(t_L^c\right) \tag{A.9}$$

The right-hand side of (A.9) gives the cost of pretending to be of type H and it consists *of the additional training cost borne by the principal. The left-hand side of (A.9) gives the benefits from mimicry and it includes two elements. First, the agent exerts higher effort and this leads to higher expected reduction in the uncertainty about the state. Second, because the principal delegates authority to the agent and does not take it back, she is exposed to the agent's biased decision but avoids the cost of implementation. If* $\left(\left(d, \widetilde{t}^d\right), \left(c, \widetilde{t}^c\right)\right)$ *constitutes an equilibrium, the type H principal must have no incentive to deviate to any alternative strategy. In particular, she must have no incentive to deviate to play* $\left(c, \widetilde{t}^c\right)$*, i.e.,* $U^d_{P_H}\left(\widetilde{t}^d; \psi\left(\widetilde{t}^d, \lambda\left(d, \widetilde{t}^d\right)\right)\right) > U^c_{P_H}\left(t^c_L; \psi\left(t^c_L, \lambda\left(c, t^c_L\right)\right)\right)$*. This condition yields that*

$$\left[\psi\left(\widetilde{t}^d, \lambda\left(d, \widetilde{t}^d\right)\right) - \psi\left(t^c_L, \lambda\left(c, t^c_L\right)\right)\right]\Delta\sigma^2 + \psi\left(\widetilde{t}^d, \lambda\left(d, \widetilde{t}^d\right)\right)\left[R + \kappa - B^2_s\right] - R > c\left(\widetilde{t}^d\right) - c\left(t^c_L\right) \tag{A.10}$$

Substituting (A.9) into (A.10), and simplifying the terms gives

$$\psi\left(\widetilde{t}^d, \lambda\left(d, \widetilde{t}^d\right)\right)\left[R + \kappa - B^2_s\right] > R - \left[b^2 - \kappa\right] \tag{A.11}$$

Since $\left[R + \kappa - B^2_s\right] < R - \left[b^2 - \kappa\right]$ *and* $\psi\left(\widetilde{t}^d, \lambda\left(d, \widetilde{t}^d\right)\right) \leq 1$*, (A.11) never holds. This leads to a contradiction, that is, the type H principal has an incentive to deviate from* $\left(d, \widetilde{t}^d\right)$*. Hence, there exists no separating equilibrium in which* $\widetilde{t}^d > \widetilde{t}^c$*.*

To conclude the proof, we need to show the type L principal has no incentive to delegate authority and provide additional training to signal her type. Note that the type L principal would find it optimal to delegate only if she could convince the agent that she is of type H, thereby inducing him to exert higher effort. Since the agent does not increase his effort provision if he believes that the principal is of type L, there exists no separating equilibria in which the type L principal uses training provision to convey her type to the agent.

Proof of Proposition 3. *Suppose that both types of principals delegate authority and provide training* $\widetilde{t}^d$*. By Bayes' rule, the agent assesses that the principal who hired him is of type H with probability* $\lambda\left(d, \widetilde{t}^d\right) = \gamma$*. Anticipating each type of principal's decision about revoking delegation decision (as described in Lemma 1), the agent chooses a level of effort,* $\widetilde{\psi}$*, to maximize his expected payoff:*

$$\widetilde{\psi}\mathbb{E}\left[u_A(\alpha_A(s))\right] + \left(1 - \widetilde{\psi}\right)\left[\gamma\mathbb{E}\left[u_A(\alpha_{P_H}(\emptyset))\right] + (1 - \gamma)\mathbb{E}\left[u_A(\alpha_A(\emptyset))\right]\right] - g\left(\widetilde{\psi}, \widetilde{t}^d\right) \tag{A.12}$$

Taking the first-order condition with respect to $\widetilde{\psi}$ yields

$$\mathbb{E}\left[u_A(\alpha_A(s))\right] - \left[\gamma \mathbb{E}\left[u_A(\alpha_{P_H}(\emptyset))\right] + (1-\gamma)\mathbb{E}\left[u_A(\alpha_A(\emptyset))\right]\right] = \frac{\partial g\left(\widetilde{\psi}, \widetilde{t}^d\right)}{\partial \widetilde{\psi}}. \tag{A.13}$$

Writing (A.13) more explicitly gives

$$\left[\left[\sigma_0^2 - \widetilde{\sigma}^2\right] + \gamma B_\emptyset^2\right] = \frac{\partial g\left(\widetilde{\psi}, \widetilde{t}^d\right)}{\partial \widetilde{\psi}}. \tag{A.14}$$

It follows from (A.14) that the agent exerts higher effort as γ, her belief that he faces the type H principal, increases. Given the agent's optimal effort choice, characterized by (A.14), each type of principal decides how much training to provide. If training $\widetilde{t}^d$ constitutes the level of training provided in a pooling equilibrium, neither type of the principals must have an incentive to deviate to an alternative organizational mode and training provision. Consider first the type H principal's decision. As we show in the Proof of Proposition 2, providing a different level of training under delegation would lead the agent to believe that she is of type L. Therefore, this not a profitable deviation for the type H principal, since the agent lowers his effort provision if he conjectures that he faces the type L principal. The type H principal must have no incentive to deviate to centralization, i.e., condition $U_{P_H}^d\left(\widetilde{t}^d; \psi\left(\widetilde{t}^d, \lambda\left(\widetilde{t}^d, d\right)\right)\right) \geq U_{P_H}^c\left(t; \psi\left(t, \lambda\left(c, t\right)\right)\right)$ must hold for any t. Fixing t and writing this condition explicitly gives

$$\left[\psi\left(\widetilde{t}^d, \lambda\left(d, \widetilde{t}^d\right)\right) - \psi\left(t, \lambda\left(c, t\right)\right)\right]\Delta\sigma^2 + \psi\left(\widetilde{t}^d, \lambda\left(d, \widetilde{t}^d\right)\right)\left[R + \kappa - B_s^2\right] - R \geq c\left(\widetilde{t}^d\right) - c\left(t\right) \tag{A.15}$$

Similarly, the type L principal anticipates that providing a training level different from $\widetilde{t}^d$ would signal her type and consequently induces the agent to lower his effort provision. Thus, deviating to an alternative training provision under delegation is not profitable. In addition, the type L principal must have no incentive to deviate to centralization, i.e., $U_{P_L}^d\left(\widetilde{t}^d; \psi\left(\widetilde{t}^d, \lambda\left(d, \widetilde{t}^d\right)\right)\right) \geq U_{P_L}^c\left(t; \psi\left(t, \lambda\left(c, t\right)\right)\right)$ for any t. For a given t, this condition can be written as

$$\left[\psi\left(\widetilde{t}^d, \lambda\left(d, \widetilde{t}^d\right)\right) - \psi\left(t, \lambda\left(c, t\right)\right)\right]\Delta\sigma^2 + \left[\kappa - B_s^2\right] \geq c\left(\widetilde{t}^d\right) - c\left(t\right) \tag{A.16}$$

Note that condition (A.16) necessarily holds when condition (A.15) holds. Therefore, it is sufficient to focus on (A.15) to show the existence of a pooling equilibrium under delegation.

Observe that (A.15) holds if the pooling equilibrium coincided with the type H principal's organizational mode and training choices in the asymmetric-information equilibrium, that is, $\widetilde{t}^d = t_H^d$ *and* $\gamma = 1$. *Holding the training provision fixed at* t_H^d, *consider what happens to the agent's effort choice when* $\gamma = 1 - \epsilon$ *with* $\epsilon > 0$ *is small. As (A.14) shows, the agent's effort decreases slightly; thus, the left-hand side of (A.15) becomes smaller. If (A.15) still holds despite lower effort provision, it means that we obtain a pooling equilibrium under delegation. If not, note that the expression on the left-hand side of (A.15) is increasing in* σ_0^2 *and* κ, *and decreasing in* σ_s^2 *(see the Proof of Lemma 4). Hence, it is possible to find a combination of* σ_0^2, σ_s^2, *and* κ *such that (A.15) holds when* $\gamma < 1$. *In other words, the lowest value of* γ *above which pooling on delegation would be optimal for the type H principal is decreasing in* σ_0^2 *and* κ, *and increasing in* σ_s^2. *Finally, note that both types of principals must provide training* t_H^d *in this equilibrium. To see this, consider the type H principal's optimal training choice given by (A.5). Since it does not depend on the agent's assessment concerning the type of the principal (note the importance of* $\frac{\partial \psi_H^d(t)}{\partial t_H^d}$ *being equal to a constant), the principal finds it optimal to provide* t_H^d. *Recall that* $t_L^d < t_H^d$ *(see Lemma 2). This means that if the type L principal does not provide* t_H^d, *the agent correctly infers her type and lowers his effort provision accordingly. Clearly, this would not be profitable for the type L principal. Hence, both types of principals provide training* t_H^d.

To complete the proof, consider what happens if pooling under delegation is not an equilibrium (that is, γ *is too small given the values of* σ_0^2, σ_s^2, *and* κ*). Suppose that both types of principals centralize authority and provide training* $\widetilde{t}^c$. *Conditional on his training, the agent chooses level of effort,* $\widetilde{\psi}$, *to maximize his expected payoff:*

$$\sum_{q'=L,H} \Pr(q = q') \left[\widetilde{\psi}\mathbb{E}\left[u_A(\alpha_{P_q}(s))\right] + \left(1 - \widetilde{\psi}\right)\mathbb{E}\left[u_A(\alpha_{P_q}(\emptyset))\right]\right] - g\left(\widetilde{\psi}, \widetilde{t}^c\right), \qquad \text{(A.17)}$$

where $\Pr(q = H) = \lambda\left(c, \widetilde{t}^c\right) = \gamma$ *by Bayes' rule. Taking the first-order condition with respect to* $\widetilde{\psi}$ *yields*

$$\left[\left[\sigma_0^2 - \widetilde{\sigma}^2\right] + \gamma\left[B_\emptyset^2 - B_s^2\right]\right] = \frac{\partial g\left(\widetilde{\psi}, \widetilde{t}^c\right)}{\partial \widetilde{\psi}}. \qquad \text{(A.18)}$$

As Lemma 2 shows, both types of principals provide the same level of training under centralization (i.e., $t_L^c = t_H^c$ *). This means that they both centralize authority and provide training* $\widetilde{t}^c = t_L^c = t_H^c$, *while the agent exerts effort* $\widetilde{\psi}$, *characterized by (A.18).*

References

Acemoglu, D., Aghion, P., Lelarge, C., Van Reenen, J., and Zilibotti, F. (2007). Technology, information, and the decentralization of the firm. *The Quarterly Journal of Economics*, 122(4):1759–1799.

Acemoglu, D. and Pischke, J. S. (1998). Why do firms train? theory and evidence. *Quarterly Journal of Economics*, 113(1):79–119.

Acemoglu, D. and Pischke, J. S. (1999). The structure of wages and investment in general training. *Journal of Political Economy*, 107(3):539–572.

Aghion, P., Bloom, N., and Van Reenen, J. (2013). Incomplete contracts and the internal organization of firms. *The Journal of Law, Economics, & Organization*, 30(suppl_1):i37–i63.

Aghion, P., Dewatripont, M., and Stein, J. C. (2008). Academic freedom, private-sector focus, and the process of innovation. *The RAND Journal of Economics*, 39(3):617–635.

Aghion, P. and Tirole, J. (1997). Formal and real authority in organizations. *Journal of Political Economy*, 105(1):1–29.

Alonso, R. and Matouschek, N. (2007). Relational delegation. *The RAND Journal of Economics*, 38(4):1070–1089.

Baker, G., Gibbons, R., and Murphy, K. J. (1999). Informal authority in organizations. *Journal of Law, Economics, & Organization*, 15(1):56–73.

Bester, H. and Krähmer, D. (2008). Delegation and incentives. *The RAND Journal of Economics*, 39(3):664–682.

Bilanakos, C., Heywood, J. S., Sessions, J. G., and Theodoropoulos, N. (2017). Does delegation increase worker training? *Economic Inquiry*. doi:10.1111/ecin.12515.

Bloom, N., Sadun, R., and Van Reenen, J. (2012). The organization of firms across countries. *The Quarterly Journal of Economics*, 127(4):1663–1705.

Bolton, P. and Dewatripont, M. (2013). *Handbook of Organizational Economics*, chapter Authority in Organizations, pages 342–372. Princeton University Press.

Caroli, E. and Van Reenen, J. (2001). Skill-biased organizational change? evidence from a panel of british and french establishments. *The Quarterly Journal of Economics*, 116(4):1449–1492.

Chaplin, J., Mangla, J., Purdon, S., and Airey, C. (2005). The workplace employment relations survey (wers) 2004 technical report. *National Centre for oficial Research: London.*

Che, Y.-K. and Kartik, N. (2009). Opinions as incentives. *Journal of Political Economy*, 117(5):815–860.

DeGroot, M. H. (2004). *Optimal Statistical Decisions.* Hoboken, NJ: John Wiley & Sons.

Dessein, W. (2002). Authority and communication in organizations. *The Review of Economic Studies*, 69(4):811–838.

DeVaro, J. and Kurtulus, F. A. (2010). An empirical analysis of risk, incentives and the delegation of worker authority. *ILR Review*, 63(4):641–661.

DeVaro, J. and Prasad, S. (2015). The relationship between delegation and incentives across occupations: Evidence and theory. *The Journal of Industrial Economics*, 63(2):279–312.

Dustmann, C. and Schönberg, U. (2012). What makes firm-based vocational training schemes successful? the role of commitment. *American Economic Journal: Applied Economics*, 4(2):36–61.

Fuchs, W. (2015). Subjective evaluations: Discretionary bonuses and feedback credibility. *American Economic Journal: Microeconomics*, 7(1):99–108.

Fudenberg, D. and Tirole, J. (1991). *Game Theory.* MIT Press, Cambridge, MA.

Gibbons, R., Matouschek, N., and Roberts, J. (2013). *Handbook of Organizational Economics*, chapter Decisions in Organizations, pages 373–431. Princeton University Press.

Grossman, S. J. and Hart, O. D. (1986). The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of Political Economy*, 94(4):691–719.

Hart, O. and Holmstrom, B. (2010). A theory of firm scope. *The Quarterly Journal of Economics*, 125(2):483–513.

Laffont, J.-J. and Martimort, D. (2009). *The theory of incentives: The principal-agent model.* Princeton University Press, Princeston.

Lo, D., Dessein, W., Ghosh, M., and Lafontaine, F. (2016). Price delegation and performance pay: Evidence from industrial sales forces. *The Journal of Law, Economics, and Organization*, 32(3):508–544.

Marino, A. M., Matsusaka, J. G., and Zábojník, J. (2009). Disobedience and authority. *The Journal of Law, Economics, & Organization*, 26(3):427–459.

Prendergast, C. (2002). The tenuous trade-off between risk and incentives. *Journal of Political Economy*, 110(5):1071–1102.

Rantakari, H. (2012). Employee initiative and managerial control. *American Economic Journal: Microeconomics*, 4(3):171–211.

Shin, D. and Strausz, R. (2014). Delegation and dynamic incentives. *The RAND Journal of Economics*, 45(3):495–520.

Zabojnik, J. (2002). Centralized and decentralized decision making in organizations. *Journal of Labor Economics*, 20(1):1–22.

TABLE 1. Descriptive Statistics

Variable	Mean	St. dev.
Individual level variables		
Training		
1="None"	0.394	0.489
2="Less than 1 day"	0.109	0.311
3="1 to less than 2 days"	0.140	0.347
4="2 to less than 5 days"	0.200	0.400
5="5 to less than 10 days"	0.085	0.278
6="10 days or more"	0.073	0.261
Personal characteristics		
Female	0.472	0.499
Male	0.528	0.499
Age	39.054	12.444
Job characteristics		
Tenure	7.144	7.317
Permanent job	0.923	0.267
Temporary job	0.050	0.218
Fixed term job	0.027	0.163
Occupations		
Managers, senior administrators	0.102	0.303
Professional	0.092	0.289
Associate professional, technical	0.132	0.339
Clerical and secretarial	0.162	0.369
Craft and skilled service	0.084	0.277
Personal and protective service	0.059	0.235
Sales	0.111	0.315
Operative and assembly	0.108	0.310
Routine/unskilled	0.149	0.356
Academic qualifications		
GCSE, grades D-G	0.081	0.272
GCSE, grades A-C	0.173	0.378
A levels (1 A-level and above)	0.093	0.291
First degree (BSc, BA, Bed, HND, HNC)	0.125	0.331
Higher degree or equivalent (MSc, MA, MBA, PGCE, PhD)	0.040	0.195
No educational qualification or other educational qualification	0.488	0.500
Job-skill match		
Overqualified	0.543	0.498
Qualified	0.409	0.492
Underqualified	0.047	0.212
Workplace level variables		
Industry		
Manufacturing	0.156	0.363
Utilities (electricity, gas, water)	0.001	0.037
Construction	0.042	0.200
Wholesale and retail	0.229	0.420
Hotels and restaurants	0.065	0.246
Transportation and communication	0.056	0.230
Financial services	0.063	0.243
Other business services	0.171	0.376
Education	0.013	0.115
Health	0.141	0.348
Other community services	0.063	0.243
Region		

Scotland	0.088	0.284
North	0.058	0.235
North west	0.139	0.346
East Midlands	0.075	0.263
West Midlands	0.11	0.313
East Anglia	0.054	0.226
South east	0.258	0.437
South west	0.094	0.291
Wales	0.028	0.166
Yorkshire and Humberside	0.095	0.293
Workplace characteristics		
Number of employees in the workplace	55.930	136.197
Part of a larger organization	0.617	0.486
Single independent workplace not belonging	0.366	0.482
Sole UK establishment of a foreign organization	0.017	0.130
Workplace age - less than 5 years	0.091	0.287
Workplace age - 5 to 9 years	0.126	0.332
Workplace age - 10 to 14 years	0.137	0.344
Workplace age - 15 to 20 years	0.173	0.378
Workplace age - 21 to 24 years	0.061	0.239
Workplace age - 25 years plus	0.413	0.492
Occupational group percentages		
Managerial and senior administrative	0.125	0.097
Professional	0.071	0.164
Technical	0.071	0.151
Sales	0.181	0.304
Operative and assembly	0.109	0.221
Clerical and secretarial	0.146	0.207
Craft and skilled services	0.086	0.188
Protective and personal service	0.092	0.244
Routine/unskilled	0.117	0.232
Current state of the market		
Market is growing	0.508	0.500
Market is declining	0.105	0.306
Market is turbulent	0.143	0.350
Market is mature	0.244	0.429
Number of competitors		
No competitors	0.066	0.249
Few competitors (less than 5)	0.386	0.487
Many competitors	0.547	0.498
Observations	10,983	

Notes. This table displays the summary statistics for all variables used in the analysis. Individual characteristics are weighted using employment weights and establishment characteristics are weighted using establishment weights.

TABLE 2: The Effect of Training Propensity and Training Intensity on Delegation Dependent Variable: Delegation (0="no delegation", 1="delegation")

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Training	0.228***	0.157***	0.155***			
	(0.034)	(0.036)	(0.036)			
Training						
<1 day				0.053	0.060	0.063
				(0.055)	(0.055)	(0.055)
>=1 day to <2 days				0.230***	0.176***	0.171***
				(0.053)	(0.054)	(0.054)
>=2 days to <5 days				0.332***	0.235***	0.233***
				(0.048)	(0.050)	(0.050)
>=5 days to <10 days				0.292***	0.149**	0.150**
				(0.067)	(0.069)	(0.069)
>=10 days				0.219***	0.130**	0.125*
				(0.065)	(0.066)	(0.066)
Growing market	0.072	0.080*	0.061	0.071	0.079*	0.061
	(0.047)	(0.046)	(0.046)	(0.046)	(0.046)	(0.046)
Declining market	0.057	0.068	0.038	0.055	0.066	0.036
	(0.079)	(0.078)	(0.080)	(0.079)	(0.078)	(0.080)
Turbulent market	0.118**	0.126**	0.099*	0.119**	0.126**	0.100*
	(0.054)	(0.054)	(0.053)	(0.054)	(0.054)	(0.053)
GCSE grades D-G	-0.041	0.035	0.031	-0.041	0.033	0.030
	(0.061)	(0.063)	(0.063)	(0.061)	(0.063)	(0.063)
GCSE grades A-C	-0.021	-0.021	-0.022	-0.019	-0.022	-0.022
	(0.042)	(0.045)	(0.045)	(0.042)	(0.045)	(0.045)
A-levels	0.085	0.026	0.019	0.082	0.024	0.018
	(0.058)	(0.060)	(0.061)	(0.058)	(0.060)	(0.061)
First degree	0.223***	0.049	0.038	0.216***	0.047	0.036
	(0.059)	(0.062)	(0.062)	(0.060)	(0.063)	(0.062)
Higher degree	0.315***	0.094	0.085	0.309***	0.092	0.084
	(0.101)	(0.106)	(0.107)	(0.102)	(0.107)	(0.107)
Age	0.040***	0.027***	0.027***	0.039***	0.026***	0.026***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Age sq/100	-0.053***	-0.037***	-0.036***	-0.051***	-0.036***	-0.035***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Tenure	0.042***	0.036***	0.037***	0.042***	0.036***	0.037***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Tenure sq/100	-0.101*	-0.079	-0.085	-0.101*	-0.078	-0.084
	(0.054)	(0.055)	(0.055)	(0.054)	(0.055)	(0.055)
Overqualified	0.158**	0.221***	0.225***	0.158**	0.218***	0.221***
	(0.072)	(0.073)	(0.073)	(0.072)	(0.073)	(0.073)
Qualified	0.280***	0.318***	0.324***	0.272***	0.311***	0.317***
	(0.074)	(0.076)	(0.076)	(0.074)	(0.076)	(0.076)
Female	-0.142***	-0.161***	-0.154***	-0.135***	-0.158***	-0.151***
	(0.038)	(0.041)	(0.041)	(0.039)	(0.041)	(0.041)
Permanent job	-0.036	0.023	0.027	-0.037	0.021	0.024
	(0.096)	(0.103)	(0.104)	(0.095)	(0.102)	(0.103)
Temporary job	-0.192	-0.061	-0.050	-0.175	-0.057	-0.048
	(0.118)	(0.125)	(0.127)	(0.118)	(0.124)	(0.125)
Log no. of employees	-0.044***	-0.043***	-0.040**	-0.044***	-0.042***	-0.040**
	(0.015)	(0.015)	(0.017)	(0.015)	(0.015)	(0.017)
Part of a larger workplace	-0.239**	-0.158	-0.157	-0.250**	-0.166	-0.166
	(0.102)	(0.108)	(0.115)	(0.100)	(0.106)	(0.114)
Single independent workplace	-0.151	-0.033	-0.022	-0.160	-0.041	-0.030
	(0.107)	(0.113)	(0.121)	(0.105)	(0.112)	(0.120)

Workplace age < 5 years	0.182**	0.172*	0.175**	0.182**	0.170*	0.174**
	(0.090)	(0.092)	(0.086)	(0.091)	(0.092)	(0.085)
Workplace age >=5 to <=9	0.132**	0.131**	0.135**	0.130**	0.130**	0.134**
	(0.064)	(0.062)	(0.062)	(0.063)	(0.062)	(0.062)
Workplace age >=10 to <=14	0.009	0.051	0.065	0.012	0.052	0.065
	(0.061)	(0.063)	(0.064)	(0.061)	(0.063)	(0.064)
Workplace age >=15 to <=20	0.041	0.045	0.045	0.044	0.045	0.045
	(0.056)	(0.055)	(0.054)	(0.055)	(0.055)	(0.053)
Workplace age >=21 to <=24	0.122	0.137	0.144*	0.122	0.135	0.142*
	(0.085)	(0.086)	(0.085)	(0.086)	(0.086)	(0.085)
Few competitors (<5)	-0.126	-0.118	-0.161**	-0.122	-0.115	-0.158**
	(0.085)	(0.088)	(0.078)	(0.085)	(0.088)	(0.078)
Many competitors	-0.134	-0.139	-0.177**	-0.132	-0.138	-0.175**
	(0.083)	(0.086)	(0.076)	(0.083)	(0.086)	(0.076)
Manager		1.188***	1.192***		1.170***	1.175***
		(0.110)	(0.114)		(0.110)	(0.113)
Professional		0.668***	0.676***		0.657***	0.666***
		(0.084)	(0.090)		(0.084)	(0.090)
Technical		0.587***	0.596***		0.573***	0.583***
		(0.074)	(0.077)		(0.074)	(0.076)
Clerk		0.300***	0.332***		0.291***	0.324***
		(0.068)	(0.069)		(0.068)	(0.069)
Craft		0.080	0.105		0.074	0.101
		(0.072)	(0.077)		(0.073)	(0.077)
Service		0.120	0.142		0.113	0.137
		(0.085)	(0.091)		(0.085)	(0.092)
Sales		0.138*	0.136*		0.130*	0.130
		(0.075)	(0.081)		(0.075)	(0.081)
Operative		-0.064	-0.093		-0.069	-0.097
		(0.065)	(0.071)		(0.065)	(0.071)
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes	Yes	Yes
Workforce composition	No	No	Yes	No	No	Yes
Constant	0.462*	0.307	0.376	0.490*	0.337	0.408
	(0.268)	(0.277)	(0.290)	(0.265)	(0.273)	(0.286)
Observations	10,983					

Notes: Cell entries are probit coefficients with standard errors clustered at the workplace level. The omitted categories are: male, no training, no academic or other educational qualification, underqualified, routine/unskilled occupation, fixed term job, percentage of routine/unskilled staff, mature market, no competitors, sole UK workplace of a foreign organization, workplace older than 25 years of age, other community services sector, Yorkshire and Humberside". Levels of significance: *** p<0.01, ** p<0.05, * p<0.1.

TABLE 3: The Effect of Training Propensity and Training Intensity on Delegation Dependent Variable: Delegation (0="None", 1="A little", 2="Some", 3="A lot")

VARIABLES	(1)	(2)	(3)	(4)
Training	0.086***			
	(0.024)			
Training				
<1 day		-0.056	-0.047	-0.044
		(0.037)	(0.037)	(0.037)
>=1 day to <2 days		0.136***	0.083**	0.083**
		(0.034)	(0.034)	(0.034)
>=2 days to <5 days		0.237***	0.151***	0.151***
		(0.032)	(0.033)	(0.033)
>=5 days to <10 days		0.238***	0.113***	0.116***
		(0.040)	(0.041)	(0.041)
>=10 days		0.206***	0.118***	0.120***
		(0.045)	(0.045)	(0.044)
Growing market	0.059*	0.063*	0.069**	0.059*
	(0.034)	(0.034)	(0.034)	(0.034)
Declining market	0.036	0.045	0.054	0.036
	(0.056)	(0.057)	(0.055)	(0.056)
Turbulent market	0.037	0.042	0.047	0.039
	(0.038)	(0.038)	(0.037)	(0.038)
GCSE grades D-G	0.013	-0.046	0.016	0.012
	(0.043)	(0.042)	(0.043)	(0.043)
GCSE grades A-C	-0.121***	-0.116***	-0.118***	-0.119***
	(0.029)	(0.029)	(0.029)	(0.029)
A-levels	-0.020	0.027	-0.022	-0.020
	(0.039)	(0.039)	(0.039)	(0.039)
First degree	-0.027	0.090***	-0.033	-0.026
	(0.034)	(0.033)	(0.034)	(0.034)
Higher degree	0.096*	0.200***	0.084	0.099*
	(0.054)	(0.054)	(0.054)	(0.054)
Age	0.034***	0.047***	0.033***	0.034***
	(0.006)	(0.006)	(0.006)	(0.006)
Age sq/100	-0.037***	-0.053***	-0.036***	-0.037***
	(0.008)	(0.008)	(0.008)	(0.008)
Tenure	0.043***	0.046***	0.042***	0.043***
	(0.008)	(0.008)	(0.008)	(0.008)
Tenure sq/100	-0.124***	-0.133***	-0.118***	-0.125***
	(0.037)	(0.036)	(0.036)	(0.036)
Overqualified	0.286***	0.238***	0.284***	0.287***
	(0.045)	(0.046)	(0.045)	(0.045)
Qualified	0.329***	0.287***	0.319***	0.325***
	(0.047)	(0.047)	(0.047)	(0.047)
Female	-0.107***	-0.114***	-0.104***	-0.102***
	(0.027)	(0.027)	(0.027)	(0.027)
Permanent job	0.027	-0.006	0.029	0.024
	(0.069)	(0.066)	(0.068)	(0.069)
Temporary job	-0.095	-0.156*	-0.079	-0.087
	(0.083)	(0.080)	(0.083)	(0.083)
Log no. of employees	-0.040***	-0.045***	-0.045***	-0.041***
	(0.011)	(0.010)	(0.010)	(0.011)
Part of a larger workplace	0.072	0.017	0.075	0.062
	(0.057)	(0.055)	(0.056)	(0.055)
Single independent workplace	0.178***	0.096	0.182***	0.168***
	(0.062)	(0.059)	(0.061)	(0.061)
Workplace age < 5 years	0.171***	0.183***	0.167***	0.171***

	(0.057)	(0.062)	(0.059)	(0.057)
Workplace age >=5 to <=9	0.123***	0.141***	0.124***	0.123***
	(0.039)	(0.041)	(0.040)	(0.040)
Workplace age >=10 to <=14	0.091**	0.077*	0.098**	0.092**
	(0.045)	(0.046)	(0.046)	(0.046)
Workplace age >=15 to <=20	0.063	0.068*	0.069*	0.066*
	(0.039)	(0.040)	(0.038)	(0.039)
Workplace age >=21 to <=24	0.189***	0.183***	0.192***	0.189***
	(0.058)	(0.058)	(0.059)	(0.059)
Few competitors (< 5)	-0.066	-0.049	-0.039	-0.063
	(0.053)	(0.054)	(0.055)	(0.053)
Many competitors	-0.035	-0.008	-0.010	-0.032
	(0.052)	(0.053)	(0.054)	(0.052)
Manager	0.886***		0.836***	0.859***
	(0.054)		(0.053)	(0.055)
Professional	0.345***		0.288***	0.328***
	(0.053)		(0.051)	(0.053)
Technical	0.478***		0.427***	0.459***
	(0.051)		(0.049)	(0.051)
Clerk	0.182***		0.134***	0.170***
	(0.050)		(0.049)	(0.050)
Craft	0.031		0.001	0.021
	(0.057)		(0.054)	(0.057)
Service	0.104		0.087	0.089
	(0.069)		(0.064)	(0.069)
Sales	0.000		-0.025	-0.003
	(0.058)		(0.054)	(0.058)
Operative	-0.106**		-0.088*	-0.114**
	(0.053)		(0.049)	(0.053)
Industry dummies	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes
Workforce composition	Yes	No	No	Yes
Ordered probit cutoffs				
Cutoff1	0.065	0.072	0.070	0.042
	(0.190)	(0.175)	(0.179)	(0.189)
Cutoff2	0.659***	0.649***	0.663***	0.636***
	(0.190)	(0.175)	(0.179)	(0.189)
Cutoff3	1.707***	1.669***	1.711***	1.685***
	(0.191)	(0.176)	(0.179)	(0.190)
Observations	10,983			

Notes: Cell entries are ordered probit coefficients with standard errors clustered at the workplace level. The omitted categories are: male, no training, no academic or other educational qualification, underqualified, routine/unskilled occupation, fixed term job, percentage of routine/unskilled staff, mature market, no competitors, sole UK workplace of a foreign organization, workplace older than 25 years of age, other community services sector, Yorkshire and Humberside”. Levels of significance: *** p<0.01, ** p<0.05, * p<0.1.