Appendix B. Online Appendix: Extensions of the Model

In this Appendix, we explore some variations of the basic model. The purpose is twofold: first, to assess the robustness of our main results, and second, to understand if additional insights emerge.

Appendix B.1. Contracts Conditioned Only on the Purchase Decision

So far we have assumed that contracts can be conditioned on both the observed signal *and* the purchase decision. One can envision situations where the signal (information), even though observable, may not be contractible. We now discuss how optimal contracts change if the purchase decision is contractible but the signal realization is not.

The choice variables in problem (6)−(9) are now subject to additional restrictions, namely, $w_{h1} = w_{\ell 1}$, $w_{h0} = w_{\ell 0}$, and $V_h = V_\ell$. Denote $w_1 \equiv w_{h1} = w_{\ell 1}$, $w_0 \equiv w_{h0} = w_{\ell 0}$, and $V_0 \equiv V_h = V_\ell$. Then the problem becomes

$$
\begin{aligned}
U(V) = \max_{e, w_1, w_0, V_0, \{\sigma_i\}_{i\in\{h,\ell\}}} \; & E\{\pi(e)[\sigma_h(E_h(e) - w_1) + (1-\sigma_h)(-w_0 + \delta U(V_0))] \\
& +(1-\pi(e))[\sigma_\ell(E_\ell(e) - w_1) + (1-\sigma_\ell)(-w_0 + \delta U(V_0))]\} \\
\text{s.t. } & E\{-\psi(e) + \pi(e)[\sigma_h w_1 + (1-\sigma_h)(w_0 + \delta V_0)] \\
& \quad +(1-\pi(e))[\sigma_\ell w_1 + (1-\sigma_\ell)(w_0 + \delta V_0)]\} = V, \\
& \psi'(e) = \pi'(e)(\sigma_h - \sigma_\ell)(w_1 - w_0 - \delta V_0) \\
& \quad \text{for each realization of } w_1, w_0, V_0, \sigma_i,\ i \in \{h, \ell\}, \\
& e \geq 0,\ w_1 \geq 0,\ w_0 \geq 0,\ V_0 \geq 0,\ \sigma_i \in \Sigma \text{ for } i \in \{h, \ell\}.
\end{aligned}
$$

We now explore how the additional constraints change the model predictions. Recall our discussion in Section 4.2.4: in our original model, reducing σ_h might be beneficial to the principal because she can provide incentives in a cheaper way through the continuation value than through an immediate payment. But this argument does not work when contracts can only be conditioned on the purchase decision. Indeed, just by looking at the problem one realizes that when $\gamma > \hat{\gamma}$, reducing σ_h weakens incentives to exert effort. It is straightforward to show that if effort is positive and $\sigma_\ell = 0$, then it is optimal to set $\sigma_h = 1$. (Similarly, if $\sigma_h = 1$, then it is optimal to set $\sigma_\ell = 0$.) Furthermore, if effort is positive and $\gamma > \hat{\gamma}$, then $\sigma_h > \sigma_\ell$, and both σ_i's cannot be interior, implying that at the optimum $\sigma_h = 1$ and $\sigma_\ell = 0$.

Even in the absence of lotteries over the purchase decision, i.e., in the $\Sigma = \{0, 1\}$ case, there are differences between this variation and our original problem. For instance, recall from Claim 2 in Section 4.2.3 that when $\sigma_h = \sigma_\ell = 0$, the principal implements positive effort on the upward sloping part of the value function. This was feasible because contracts could distinguish between V_h and V_ℓ. But when these two values are restricted to be the same, effort is *necessarily* zero in the U_{00} problem.

Appendix B.2. The Principal Observes an Additional Signal

In our model, the agent's signal is the only source of information available to the principal. In certain environments, she may observe an additional (contractible) signal about an option upon exercising it. We now analyze the effects of this extra source of information.

Suppose that if the principal exercises an option, then she observes an additional signal $\varphi \in \{\varphi_\ell, \varphi_h\}$, whose informativeness is exogenous, given by $\Pr\{\varphi_h | y = i\} = \rho_i$, $i = \ell, h$, with $0 \leq \rho_\ell \leq \rho_h \leq 1$.[37] In the investor and financial expert example, the quality of an option can be the project type, with a project of type i succeeding with probability ρ_i. After investing in the project, the investor observes whether it succeeded or failed, and in addition to the purchase decision, she can condition the payment to the expert on this event.

For simplicity, we will provide a detailed analysis of the static case only, and then later comment on the additional insights that arise in the dynamic case.

Let w_{i1j} be the agent's payment if the principal buys after signal θ_i, and the additional signal realization is φ_j. As before, w_{i0} is the payment if the principal does not buy after signal θ_i. As in our original model, it is straightforward to show that in the static case $\sigma_i \geq 0$ if and only if $E_i(e) \geq 0$, and thus there are only three cases to consider: $e = 0$ and $\sigma_h = \sigma_\ell = 0$; $e = 0$ and $\sigma_h = \sigma_\ell = 1$; and $e > 0$ and $\sigma_h = 1$, $\sigma_\ell = 0$. Consider the third alternative, where the choices of w_{h0}, $w_{\ell1\ell}$, $w_{\ell1h}$ are irrelevant. The next result shows how incentives are structured in this case.

Claim 7 (Additional Signal). *Suppose the principal buys only after observing a high signal. Then there is a threshold $\check{\gamma}$ such that*

(i) If $\gamma > \check{\gamma}$, then $w_{h1h} > w_{\ell0} = w_{h1\ell} = 0$ at the optimal contract.

(ii) If $\gamma < \check{\gamma}$, then $w_{\ell0} > w_{h1h} = w_{h1\ell} = 0$ at the optimal contract.

Proof. Define $\pi_{hh}(e) \equiv \Pr\{\theta = \theta_h, \varphi = \varphi_h | e\} = \gamma(\alpha + \beta_h\eta(e))\rho_h + (1-\gamma)(\alpha - \beta_\ell\eta(e))\rho_\ell$ and $\pi_{\ell\ell}(e) \equiv \Pr\{\theta = \theta_\ell, \varphi = \varphi_\ell | e\} = \gamma(1-\alpha-\beta_h\eta(e))(1-\rho_h) + (1-\gamma)(1-\alpha+\beta_\ell\eta(e))(1-\rho_\ell)$. The principal's problem can be written as

$$
\begin{aligned}
\max_{e,\{w_{i1j}, w_{i0}, \sigma_i\}_{i,j\in\{h,\ell\}}} \quad & \pi(e)\sigma_h E_h(e) - \pi_{hh}(e)\sigma_h w_{h1h} - (\pi(e) - \pi_{hh}(e))\sigma_h w_{h1\ell} - \pi(e)(1-\sigma_h)w_{h0} \\
& +(1-\pi(e))\sigma_\ell E_\ell(e) - \pi_{\ell\ell}(e)\sigma_\ell w_{\ell1\ell} - (1 - \pi - \pi_{\ell\ell}(e))\sigma_\ell w_{\ell1h} - (1-\pi(e))(1-\sigma_\ell)w_{\ell0} \\
\text{s.t.} \quad & -\psi(e) + \pi_{hh}(e)\sigma_h w_{h1h} + (\pi(e) - \pi_{hh}(e))\sigma_h w_{h1\ell} + \pi(e)(1-\sigma_h)w_{h0} \\
& \qquad +\pi_{\ell\ell}(e)\sigma_\ell w_{\ell1\ell} + (1-\pi(e)-\pi_{\ell\ell}(e))\sigma_\ell w_{\ell1h} + (1-\pi(e))(1-\sigma_\ell)w_{\ell0} \geq 0, \\
& \psi'(e) = \pi'_{hh}(e)\sigma_h(w_{h1h} - w_{h1\ell}) + \pi'_{\ell\ell}(e)\sigma_\ell(w_{\ell1\ell} - w_{\ell1h}) \\
& \qquad +\pi'(e)[\sigma_h w_{h1\ell} + (1-\sigma_h)w_{h0} - \sigma_\ell w_{\ell1h} - (1-\sigma_\ell)w_{\ell0}], \\
& e \geq 0,\ w_{i1j} \geq 0,\ w_{i0} \geq 0,\ \sigma_i \in [0,1] \text{ for } i,j \in \{h,\ell\}.
\end{aligned}
$$

Suppose that the principal buys only after the high signal ($\sigma_h = 1$, $\sigma_\ell = 0$) so that the choices of w_{h0}, $w_{\ell1\ell}$, $w_{\ell1h}$ are irrelevant. Let μ denote the Lagrange multiplier on the incentive constraint (the

[37] A special case is observing the option quality without noise, i.e., when $\rho_h = 1$ and $\rho_\ell = 0$.

participation constraint is slack). The first-order conditions with respect to w_{h1h}, $w_{h1\ell}$, and $w_{\ell 0}$ are

$$\begin{aligned} -\pi_{hh}(e) + \mu\pi'_{hh}(e) &\leq 0, \qquad w_{h1h} \geq 0, \\ -(\pi(e) - \pi_{hh}(e)) + \mu(\pi'(e) - \pi'_{hh}(e)) &\leq 0, \qquad w_{h1\ell} \geq 0, \\ -(1 - \pi(e)) - \mu\pi'(e) &\leq 0, \qquad w_{\ell 0} \geq 0, \end{aligned} \tag{B.1}$$

all with complementary slackness. Equivalently,

$$\begin{aligned} -1 + \mu\frac{\pi'_{hh}(e)}{\pi_{hh}(e)} &\leq 0, \qquad w_{h1h} \geq 0, \\ -1 + \mu\frac{\pi'(e) - \pi'_{hh}(e)}{\pi(e) - \pi_{hh}(e)} &\leq 0, \qquad w_{h1\ell} \geq 0, \\ -1 - \mu\frac{\pi'(e)}{1 - \pi(e)} &\leq 0, \qquad w_{\ell 0} \geq 0, \end{aligned}$$

Notice that only those of the above inequalities can hold with equality—and thus the wages for only those outcomes will be positive— for which the corresponding likelihood ratio is the highest. That is, we need to compare the likelihood ratios $\pi'_{hh}/\pi_{hh} = \eta'[\gamma\beta_h\rho_h - (1-\gamma)\beta_\ell\rho_\ell]/[\gamma(\alpha+\beta_h\eta)\rho_h + (1-\gamma)(\alpha-\beta_\ell\eta)\rho_\ell]$, $(\pi' - \pi'_{hh})/(\pi - \pi_{hh}) = \eta'[\gamma\beta_h(1-\rho_h) - (1-\gamma)\beta_\ell(1-\rho_\ell)]/[\gamma(\alpha+\beta_h\eta)(1-\rho_h) + (1-\gamma)(\alpha-\beta_\ell\eta)(1-\rho_\ell)]$, and $-\pi'/(1-\pi) = \eta'[(1-\gamma)\beta_\ell - \gamma\beta_h]/[\gamma(1-\alpha-\beta_h\eta) + (1-\gamma)(1-\alpha+\beta_\ell\eta)]$. It is easy to show that $\pi'_{hh}/\pi_{hh} > (\pi' - \pi'_{hh})/(\pi - \pi_{hh})$ as $\rho_h > \rho_\ell$. As a result, $w_{h1\ell}$ is always zero (it is dominated by either w_{hh} or $w_{\ell 0}$). Moreover, tedious algebra reveals that $\pi'_{hh}/\pi_{hh} \gtrless -\pi'/(1-\pi)$ if and only if $\gamma \gtrless \check{\gamma}$, where $\check{\gamma}$ is the positive solution of a quadratic equation, and it is given by

$$\check{\gamma} = \frac{-\left[\frac{\left(1+\frac{\beta_h\rho_h}{\beta_\ell\rho_\ell}\right)}{\alpha\left(1+\frac{\beta_h}{\beta_\ell}\right)\left(\frac{\rho_h}{\rho_\ell}-1\right)} - 1\right] + \sqrt{\left[\frac{\left(1+\frac{\beta_h\rho_h}{\beta_\ell\rho_\ell}\right)}{\alpha\left(1+\frac{\beta_h}{\beta_\ell}\right)\left(\frac{\rho_h}{\rho_\ell}-1\right)} - 1\right]^2 + \frac{4}{\alpha\left(1+\frac{\beta_h}{\beta_\ell}\right)\left(\frac{\rho_h}{\rho_\ell}-1\right)}}}{2}.$$

Thus, $w_{h1h} > w_{\ell 0} = 0$ $\gamma > \check{\gamma}$ and $w_{\ell 0} > w_{h1h} = 0$ if $\gamma < \check{\gamma}$.[38] □

To convey the intuition of the claim in the simplest way, let $\beta_h = \beta_\ell$ and $\alpha = 1/2$. Then $\check{\gamma} = 1/(1+\sqrt{\rho_h/\rho_\ell})$. Notice that when the agent's signal is high and the principal exercises the option, she does not compensate the agent if her own signal turns out to be low. Whether the agent is compensated after the high signal (in which case the principal buys) followed by the principal's high signal or after the low signal (in which case she does not buy) depends on the prior. As long as $\rho_\ell > 0$, there is an interval of γ where $w_{\ell 0} > w_{h1h} = 0$ and thus the agent is rewarded for bad news. For this to be a feature of the optimal contract, it must be the case that the principal finds it optimal to implement a positive level of effort for those values of γ. Recall that she implements zero effort near both $\gamma = 0$ and $\gamma = 1$. But if ρ_ℓ is sufficiently high—i.e., if the principal's signal generates a false positive with a sufficiently high probability —, then $1/(1+\sqrt{\rho_h/\rho_\ell})$ exceeds the threshold of γ below which $e = 0$, and hence rewarding for bad news emerges in equilibrium.

[38] It is easy to verify that $\check{\gamma} \leq \hat{\gamma}$, with strict inequality if $\rho_h > \rho_\ell$.

When the principal can observe an additional signal, rewarding for bad news is less likely to obtain. To see this, notice that $\check{\gamma} \leq \hat{\gamma} = 1/2$, with strict inequality so long as the additional signal is informative, i.e., $\rho_\ell < \rho_h$. The more informative the principal's signal is, the lower $\check{\gamma}$ is, and it becomes zero if the signal is perfectly informative.

In our model, no effort can be induced if the agent can misreport his signal. The reason is that the agent's payoff depends only on his report. This is not the case in this extension, as the agent's payoff also depends on the additional (contractible) signal, which in turn is correlated with the option quality.

Suppose the agent can misreport his signal. When $\sigma_h = \sigma_\ell$, effort is optimally zero, and hence the agent's signal bears no information, so misreporting is irrelevant. Consider a more interesting case where the principal optimally sets $\sigma_h = 1$ and $\sigma_\ell = 0$. When the agent can misreport, in addition to the moral hazard problem due to the unobservable effort we also have an adverse selection problem due to the unobservable signal. Thus we need to worry about double deviations. However, one can show that whenever the agent plans to misreport a signal realization, he will find it optimal to exert zero effort.

Indeed, suppose first that the agent plans to always report the high signal regardless of the signal he actually sees. Then his payoff (keeping in mind that $\sigma_h = 1$) is $-\psi(e) + [\gamma\rho_h + (1-\gamma)\rho_\ell]w_{h1h} + [\gamma(1-\rho_h) + (1-\gamma)(1-\rho_\ell)]w_{h1\ell}$, which clearly is maximized when $e = 0$. Intuitively, if the agent is not using the acquired information, it is not worthwhile to acquire it in the first place. Similarly, the agent will exert zero effort if he always plans to report the low signal, with the resulting payoff of $w_{\ell 0}$ (as $\sigma_\ell = 0$). Suppose the agent intends to only misreport part of the time, e.g., report low after the low signal, but report high and low with equal probabilities after the high signal. For the randomization to be optimal, the agent must be indifferent ex post between the two reports. But then his payoff is the same when he always reports the low signal, and we already showed that that payoff is maximized when effort is zero.

The above analysis suggests that only the following two additional ('truth-telling') constraints must be imposed:

$$
\begin{aligned}
-\psi(e) + \pi_{hh}(e)w_{h1h} + (\pi(e) - \pi_{hh}(e))w_{h1\ell} + (1-\pi(e))w_{\ell 0}& \\
\geq [\gamma\rho_h + (1-\gamma)\rho_\ell]w_{h1h} + [\gamma(1-\rho_h) + (1-\gamma)(1-\rho_\ell)]w_{h1\ell}&, \\
-\psi(e) + \pi_{hh}(e)w_{h1h} + (\pi(e) - \pi_{hh}(e))w_{h1\ell} + (1-\pi(e))w_{\ell 0} \geq w_{\ell 0}&.
\end{aligned}
$$

Notice that one of the constraints would be violated if only one of the two wages—w_{h1h} or $w_{\ell 0}$—were positive, as in Claim 7. Indeed, it is easy to show that in the optimal contract both w_{h1h} and $w_{\ell 0}$ must be positive for all γ, while $w_{h1\ell}$ is still zero.

As mentioned, Gromb and Martimort (2007) analyze a model, which is similar to this extension. Compared to their setup, continuous effort allows us to clearly separate effects of moral hazard (unobservable effort) and adverse selection (possibility of misreporting a signal). The technical reason is that we have the incentive constraint for the effort choice that is separate from the truth-telling constraints. In their model, both truth-telling constraints bind in equilibrium, and the prior plays no essential role. In our model, the prior is crucial; the principal rewards the agent with w_{h1h} ($w_{\ell 0}$) when $\gamma > \check{\gamma}$ ($\gamma < \check{\gamma}$) to induce him to exert effort, and pays $w_{\ell 0}$ (w_{h1h}) in order to eliminate incentives to misreport.

So far we have analyzed the static version of this extension. The main insights of this case extend to the dynamic model, plus an additional property of the optimal contract emerges, that we now explain.

Consider what happens after the high signal if $\sigma_h \in (0,1)$. Recall from part (i) of Proposition 2 that in this case making no immediate payment after the high signal realization ($w_h = 0$) is optimal. The reason is that the marginal cost of paying today vs. in the future is higher (1 vs. $|U'(V_h)| \leq 1$), while the marginal benefit of the incentive provision is the same. With an additional signal after purchase this is no longer true. As mentioned, the reason is that buying allows the principal to condition current (but not future) payments on the extra signal and to pay only if this signal matches the agent's. This makes the incentive provision using current payments more effective than in the case without the additional signal. As a result, positive payments after purchase can be part of the optimal contract even though $\sigma_h \in (0,1)$.

To see this formally, consider for simplicity the case where the agent cannot misreport his signal, and compare the first-order conditions with respect to w_{h1h} and V_h given by (B.1) and (A.9), respectively. They can be rewritten as follows (assuming for simplicity that U is differentiable at V_h):

$$-1 + \mu \frac{\pi'_{hh}(e)}{\pi_{hh}(e)} \leq 0, \qquad w_{h1h} \geq 0, \tag{B.2}$$

$$U'(V_h) + \mu \frac{\pi'(e)}{\pi'(e)} \leq 0, \qquad V_h \geq 0, \tag{B.3}$$

with complementary slackness. Straightforward algebra yields $\pi'_{hh}(e)/\pi_{hh}(e) - \pi'(e)/\pi(e) = (\rho_h - \rho_\ell)\gamma(1-\gamma)\alpha(\beta_h + \beta_\ell) > 0$. Thus w_{h1h} is not dominated by V_h when $U'(V_h)$ is close enough to -1.