Risk Limits as Optimal Contracts

DIMITRI VAYANOS
London School of Economics, CEPR and NBER

15 September 2018 *

PRELIMINARY AND INCOMPLETE

Abstract

I develop a model of optimal contracting between investors and asset managers that combines moral hazard and adverse selection. Managers incur a private cost to acquire information, and differ in their preferences and the private information they may acquire. The optimal contract keeps the risk chosen by managers within bounds, even when the optimal level of risk given managers' private information exceeds the bounds. Risk limits arise because managers may not acquire information and gamble for a high fee. The model assumes symmetric return distributions. The case of asymmetric distributions is treated in Buffa, Vayanos, and Woolley (2019).

* `d.vayanos@lse.ac.uk`. I thank Peter DeMarzo and Paul Woolley for helpful discussions.

1 Introduction

To be written.

2 Model

There are two periods 0 and 1. The riskless rate is zero. A risky asset pays D in period 1 and trades at price S in period 0. We assume that D takes the values $S+d$ or $S-d$, where $d>0$. The prior probabilities of these outcomes are π_0 and $1-\pi_0$, respectively.

An investor can invest in the risky asset either directly or by employing an asset manager. The manager can observe an informative signal about the asset payoff by incurring a private cost. The investor has negative exponential utility over consumption in period 1, with coefficient of absolute risk aversion ρ. The manager can be risk-averse or risk-neutral. A risk-averse manager has negative exponential utility over period 1 consumption, with coefficient of absolute risk aversion $\bar{\rho}$, and can observe the signal by incurring a cost K. A risk-neutral manager has linear utility over period 1 consumption, and cannot observe the signal at any cost ($K=\infty$). The probability that the manager is risk-neutral is λ. Both the risk-averse and the risk-neutral manager have an outside option of zero. Our assumed heterogeneity across managers captures the idea that investors are concerned that managers may take excessive risk relative to their information.

The investor's wealth in period 0 is equal to zero. This assumption is without loss of generality because the investor has negative exponential utility. Denoting by z the number of shares of the risky asset that the investor holds in period 0, the investor's wealth in period 1 is

$$W = z(D-S), \tag{2.1}$$

the capital gains between periods 0 and 1. This wealth is gross of any fee that the investor pays the manager in period 1.

If the investor employs the manager, then they agree on a contract in period 0. The contract specifies a fee $f(W,D)$ that can depend on the investor's wealth W and the asset payoff D in period 1. Given the fee, the manager chooses whether or not to observe the signal and what position z in the risky asset to choose for the investor. If the investor invests directly, then he pays no fee to the manager and chooses a position z on the risky asset without observing the signal.

We allow the fee $f(W,D)$ to be a general function of W and D subject to two restrictions. The first restriction is that the fee must be non-negative, i.e., the manager has limited liability. The second restriction, to which we refer as monotonicity, applies only to values of W that can be obtained in equilibrium. The fee levels corresponding to any two such values $W_1 > W_2$ must satisfy $f(W_1) - f(W_2) \geq \epsilon(W_1 - W_2)$, where ϵ is a positive constant that does not depend on (W_1, W_2). The constant ϵ can be arbitrarily small, and in fact we focus on the limit of the optimal fee when ϵ goes to zero. In that limit, the monotonicity restriction is only that the fee is non-decreasing in W across values of W that can be obtained in equilibrium.

A non-decreasing fee is economically appealing because it ensures that the manager does not engage in (unmodeled) activities that reduce W, e.g., costly round-trip transactions, so to raise her fee. If, in addition, these activities yield a slight benefit to the manager, then the fee must be strictly increasing. We impose the lower bound ϵ on the fee's slope to rule out that the investor can separate the risk-averse and risk-neutral manager types by exploiting their indifference over z. We elaborate on this point in Section 3.

The investor maximizes his expected utility over the following decisions: whether or not to employ the manager; if he employs the manager, whether or not to induce her to observe the signal, and what position z in the risky asset to induce her to choose; if he invests directly, what position z in the risky asset to choose. The investor must take into account the manager's incentive-compatibility (IC) constraints on whether or not to observe the signal and what position z to choose. These constraints depend on the fee. The investor only needs to determine the fee levels for the values of W that can be obtained under one of the induced positions z. The investor can ensure that all non-induced positions are always dominated for the manager by setting the corresponding fee levels to zero: this ensures that only the induced positions need to be considered in the manager's (IC) constraints.

Four versions of the static contracting model can be studied, in order of increasing complexity. In the first version, the prior probability π_0 is equal to one-half, the fee $f(W, D)$ can depend only on W, and there are only two values of the signal, which furthermore yield posterior probabilities that are symmetric around one-half. The second version allows for a continuum of signal values. The third version allows π_0 to differ from one-half, and signal values not to be symmetric around π_0. The fourth version allows the fee $f(W, D)$ to depend on D. The first and second versions are studied in Sections 3 and 4, respectively. The third version is studied in Buffa, Vayanos, and Woolley (2019). The fourth version is work in progress.

3 Symmetry and Two Signal Values

We denote the posterior probability of $S + d$ under the two values of the signal by $\bar{\pi} > \frac{1}{2}$ and $1 - \bar{\pi}$. By symmetry, the two signal values are equally likely. We refer to the risk-averse manager with posterior $\pi \in \{1 - \bar{\pi}, \bar{\pi}\}$ for $S + d$ as the risk-averse type π. When not making reference to a specific posterior, e.g., before the signal is observed, we refer to the risk-averse manager as the risk-averse type. We likewise to the risk-neutral manager as the risk-neutral type. We sketch the main steps of the solution of the investor's optimization problem below, and provide formal proofs in the Appendix (proof of Proposition 3.1).

We can simplify the investor's optimization problem using six observations. A first observation is that if the investor invests directly, then his optimal position is $z = 0$. This is because the expected capital gains from the risky asset are zero under the prior probability $\pi_0 = \frac{1}{2}$. A second observation is that if the investor prefers to employ the manager, then he must be inducing the risk-averse type to observe the signal and choose a non-zero position z. This is because if the manager does not observe the signal, then she has the same information as the investor and hence does not add value. The manager also does not add value if she chooses a zero position, which is what the investor would choose if he were to invest directly. Based on these two observations, we can restrict attention to two outcomes: either the investor employs the manager and the risk-averse type observes the signal and chooses a non-zero position, or the investor invests directly and chooses a zero position. We next characterize the first outcome, assuming that the investor prefers it to the second.

A third observation follows from the symmetry of the problem: if a position $z(\bar{\pi})$ maximizes the utility of the risk-averse type $\bar{\pi}$, then the position $-z(\bar{\pi})$ maximizes the utility of the risk-averse type $1 - \bar{\pi}$. This is because the fee depends only on the investor's wealth, which has the same distribution under $(1 - \bar{\pi}, -z)$ as under $(\bar{\pi}, z)$: it takes the values zd and $-zd$ with probabilities $\bar{\pi}$ and $1-\bar{\pi}$, respectively. A fourth observation is that only a long position $z(\bar{\pi}) \geq 0$ can maximize the utility of the risk-averse type $\bar{\pi}$ (and hence only a short position $-z(\bar{\pi}) \leq 0$ can maximize the utility of the risk-averse type $1-\bar{\pi}$). Indeed, wealth under a long position $z \geq 0$ takes the same values, zd and $-zd$, as under the short position $-z$ but the probabilities are inverted: probabilities are $\bar{\pi}$ and $1-\bar{\pi}$, respectively, under the long position, and $1-\bar{\pi}$ and $\bar{\pi}$, respectively, under the short position. Hence, under the short position, probability weight is shifted towards the negative outcome $-zd$ under which the fee is smaller. The optimal long position $z(\bar{\pi})$ is strictly positive because when the investor prefers to employ the manager the fee must be inducing non-zero positions.

A fifth observation is that if a long position $\hat{z} \geq 0$ maximizes the utility of the risk-neutral type, then the short position $-\hat{z}$ also maximizes her utility. This is because the fee depends only on the investor's wealth, which has the same distribution under $(\frac{1}{2}, -z)$ as under $(\frac{1}{2}, z)$: zd and $-zd$ with equal probabilities. Hence, we can assume without loss of generality that the risk-neutral type chooses $\hat{z}$ and $-\hat{z}$ with equal probabilities. A sixth and final observation is that because the risk-neutral type is uninformed, the investor wants her to choose a position $\hat{z}$ that is as close to zero as possible. The investor may be unable to induce a zero position, but can at least ensure that

$\hat{z}$ does not exceed the position $z(\bar{\pi})$ of the risk-averse type $\bar{\pi}$ by setting the fee $f(W)$ to zero for $W > z(\bar{\pi})d$.

Based on the latter four observations, there are two fee levels that are relevant for the risk-averse type and at most two additional fee levels that are relevant for the risk-neutral type. These are: the fee $f(z(\bar{\pi})d)$ when the risk-averse type observes the signal, chooses $z = z(\bar{\pi}) > 0$ when she has posterior $\bar{\pi}$ for $S + d$ and $z = -z(\bar{\pi})$ when she has posterior $1 - \bar{\pi}$, and the payoff realization is favorable (d when $z = z(\bar{\pi})$ and $-d$ when $z = -z(\bar{\pi})$); the fee $f(-z(\bar{\pi})d)$ when the risk-averse type observes the signal, makes the above choices, and the payoff realization is unfavorable; the fee $f(\hat{z}d)$ when the risk-neutral type chooses $z = \hat{z} \in [0, z(\bar{\pi})]$ or $z = \hat{z}$, and the payoff realization is favorable (d when $z = \hat{z}$ and $-d$ when $z = -\hat{z}$); and the fee $f(-\hat{z}d)$ when the risk-neutral type chooses $z = \hat{z}$ or $z = \hat{z}$, and the payoff realization is unfavorable.

The fee levels $(f(z(\bar{\pi})d), f(-z(\bar{\pi})d), f(\hat{z}d), f(-\hat{z}d))$ must satisfy a number of constraints. Non-negativity and monotonicity of the fee imply

$$f(z(\bar{\pi})d) - 2\epsilon z(\bar{\pi})d \geq f(\hat{z}d) - \epsilon(z(\bar{\pi}) + \hat{z})d \geq f(-\hat{z}d) - \epsilon(z(\bar{\pi}) - \hat{z})d \geq f(-z(\bar{\pi})d) \geq 0. \qquad (3.1)$$

Additional constraints follow from incentive compatibility (IC). A first (IC) constraint concerns the risk-averse type's choice whether or not to observe the signal. Incurring the cost K and observing the signal is preferable to not observing the signal if

$$\begin{aligned} &-\left[\bar{\pi}e^{-\bar{\rho}[f(z(\bar{\pi})d)-K]} + (1-\bar{\pi})e^{-\bar{\rho}[f(-z(\bar{\pi})d)-K]}\right] \\ &\geq \max\left\{-\frac{1}{2}\left[e^{-\bar{\rho}f(z(\bar{\pi})d)} + e^{-\bar{\rho}f(-z(\bar{\pi})d)}\right], -\frac{1}{2}\left[e^{-\bar{\rho}f(\hat{z}d)} + e^{-\bar{\rho}f(-\hat{z}d)}\right]\right\}. \end{aligned} \qquad (3.2)$$

The left-hand side of (3.2) exceeds the first term in the maximum in the right-hand side under the following conditions:

$$e^{-\bar{\rho}K} - 2(1-\bar{\pi}) > 0, \qquad (3.3)$$

$$e^{\bar{\rho}[f(z(\bar{\pi})d)-f(-z(\bar{\pi})d)]} \geq \frac{2\bar{\pi} - e^{-\bar{\rho}K}}{e^{-\bar{\rho}K} - 2(1-\bar{\pi})}. \qquad (3.4)$$

Equation (3.3) requires that the cost K of observing the signal is not too large. Equation (3.4) requires that given such a value of K, the investor pays the risk-averse type sufficiently more when she earns a high return ($W_1 = z(\bar{\pi})d$) than when she earns a low return ($W_1 = -z(\bar{\pi})d$). A sufficient difference $f(z(\bar{\pi})d) - f(-z(\bar{\pi})d)$ between the two fee levels is required to induce the risk-averse type to observe the signal. The minimum required difference is increasing in the observation cost K.

The two remaining (IC) constraints concern the choice of position z by the risk-averse and risk-neutral types. We consider the (IC) constraints concerning long positions only. These imply the (IC) constraints concerning short positions only because of the symmetry of the problem, while the (IC) constraints concerning both long and short positions are implied by the monotonicity of the fee. We hence focus on the risk-averse type $\bar{\pi}$ and on the long position $\hat{z}$ chosen by the risk-neutral type. For expositional ease, we refer from now on to $\hat{z}$ as the risk-neutral type's position, although that type chooses $\hat{z}$ and $-\hat{z}$ with equal probabilities. The risk-averse type $\bar{\pi}$ prefers $z(\bar{\pi})$ to $\hat{z}$ if

$$-\left[\bar{\pi}e^{-\bar{\rho}f(z(\bar{\pi})d)} + (1-\bar{\pi})e^{-\bar{\rho}f(-z(\bar{\pi})d)}\right] \geq -\left[\bar{\pi}e^{-\bar{\rho}f(\hat{z}d)} + (1-\bar{\pi})e^{-\bar{\rho}f(-\hat{z}d)}\right]. \tag{3.5}$$

Conversely, the risk-neutral type prefers $\hat{z}$ to $z(\bar{\pi})$ if

$$\frac{1}{2}\left[f(\hat{z}d) + f(-\hat{z}d)\right] \geq \frac{1}{2}\left[f(z(\bar{\pi})d) + f(-z(\bar{\pi})d)\right]. \tag{3.6}$$

We denote by

$$\Delta(\bar{\pi}) \equiv f(z(\bar{\pi})d) - f(-z(\bar{\pi})d) > 0$$

the difference in fee levels across favorable and unfavorable payoff realizations for the risk-averse type $\bar{\pi}$. We denote by

$$\Gamma(\bar{\pi}) \equiv \frac{1}{2}\left[f(z(\bar{\pi})d) + f(-z(\bar{\pi})d)\right]$$

the arithmetic average of the fee levels for the risk-averse type $\bar{\pi}$. We denote by

$$\hat{\Delta} \equiv f(\hat{z}d) - f(-\hat{z}d) \geq 0,$$
$$\hat{\Gamma} \equiv \frac{1}{2}\left[f(\hat{z}d) + f(-\hat{z}d)\right],$$

the corresponding quantities for the risk-neutral type. Because the asset payoff can take only two values, the fee levels $(f(zd), f(-zd))$ corresponding to a position z can be described fully by the average fee and the fee difference.

Using the constraints (3.1)-(3.2), we can determine the optimal fee. This requires determining the positions $(z(\bar{\pi}), \hat{z})$ and fee levels $(f(z(\bar{\pi})d), f(-z(\bar{\pi})d), f(\hat{z}d), f(-\hat{z}d))$. The first and key step is to determine $\hat{z}$ and $(f(\hat{z}d), f(-\hat{z}d))$.

According to the (IC) constraint (3.6), the average fee $\hat{\Gamma}$ of the risk-neutral type must be at least as large as the average fee $\Gamma(\bar{\pi})$ of the risk-averse type $\bar{\pi}$. If (3.6) held as a strict inequality, then the investor could raise his utility by reducing payments to the risk-neutral type. Hence, the average fees $\Gamma(\bar{\pi})$ and $\hat{\Gamma}$ must be equal. Equality of average fees implies that the fees paid to the risk-averse type $\bar{\pi}$ and to the risk-neutral type can differ only in $\Delta(\bar{\pi})$ and $\hat{\Delta}$.

Since $z(\bar{\pi}) \geq \hat{z}$ and the monotonicity of the fee implies

$$\min\{f(z(\bar{\pi})d) - f(\hat{z}d), f(-\hat{z}d) - f(-z(\bar{\pi})d)\} \geq \epsilon(z(\bar{\pi}) - \hat{z})d, \tag{3.7}$$

the fee difference $\hat{\Delta}$ for the risk-neutral type does not exceed its counterpart $\Delta(\bar{\pi})$ for the risk-averse type $\bar{\pi}$. If, in addition, $z(\bar{\pi}) > \hat{z}$, then (3.7) implies that the fee difference for the risk-averse type $\bar{\pi}$ exceeds that for the risk-neutral type. Hence, the investor can separate the risk-averse type $\bar{\pi}$ and the risk-neutral type in terms of their positions (i.e., $\hat{z} < z(\bar{\pi})$) only if he can induce the risk-averse type to accept a riskier fee (i.e., larger fee difference) than the risk-neutral type. Separation is possible for some parameter values because the risk-averse type observes the signal and is hence better informed than the risk-neutral type. For other parameter values, however, risk-aversion precludes separation. We next determine the conditions for separation versus pooling.

Multiplying both sides of (3.5) by $e^{\bar{\rho}\Gamma(\bar{\pi})} = e^{\bar{\rho}\hat{\Gamma}}$, we can write that equation as

$$\begin{aligned} & -\left(\bar{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})} + (1-\bar{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\right) \geq -\left(\bar{\pi}e^{-\frac{\bar{\rho}}{2}\hat{\Delta}} + (1-\bar{\pi})e^{\frac{\bar{\rho}}{2}\hat{\Delta}}\right) \\ & \Leftrightarrow G_{\bar{\pi}}(\Delta(\bar{\pi})) \geq G_{\bar{\pi}}(\hat{\Delta}), \end{aligned} \tag{3.8}$$

where the function $G_{\bar{\pi}}(\Delta)$ is defined by

$$G_{\bar{\pi}}(\Delta) \equiv -\left(\bar{\pi}e^{-\frac{\bar{\rho}}{2}\Delta} + (1-\bar{\pi})e^{\frac{\bar{\rho}}{2}\Delta}\right). \tag{3.9}$$

The function $G_{\bar{\pi}}(\Delta)$ gives the utility of the risk-averse type $\bar{\pi}$ when she receives a payment $\frac{\Delta}{2}$ under the favorable payoff realization and $-\frac{\Delta}{2}$ under the unfavorable realization. This function is hump-shaped: increasing for $\Delta \in [0, \Delta^*(\bar{\pi}))$ and decreasing for $\Delta \in (\Delta^*(\bar{\pi}), \infty)$, where $\Delta^*(\bar{\pi}) > 0$ is defined by

$$e^{\bar{\rho}\Delta^*(\bar{\pi})} \equiv \frac{\bar{\pi}}{1-\bar{\pi}}. \tag{3.10}$$

Increasing Δ has two countervailing effects on the risk-averse type's utility. On the one hand, the expected payment conditional on that type's information increases. This is because the payment under the favorable payoff realization increases, and the risk-averse type has better-than-equal odds of identifying that realization by observing the signal. On the other hand, the variance of the payment increases, and this reduces the risk-averse type's utility. The first effect dominates when Δ is small because the risk-averse type is approximately risk-neutral and hence does not care about variance. The second effect dominates for large Δ.

Separation is possible when the lower bound on $\Delta(\bar{\pi}) \equiv f(z(\bar{\pi})d) - f(-z(\bar{\pi})d)$, given in (3.4), is smaller than $\Delta^*(\bar{\pi})$. From (3.4) and (3.10), this condition is

$$\frac{2\bar{\pi} - e^{-\bar{\rho}K}}{e^{-\bar{\rho}K} - 2(1-\bar{\pi})} < \frac{\bar{\pi}}{1-\bar{\pi}} \Leftrightarrow e^{-\bar{\rho}K} > 4(1-\bar{\pi})\bar{\pi}. \tag{3.11}$$

Equation (3.11) ensures that the function $G_{\bar{\pi}}(\Delta)$ is increasing for all values of Δ between zero and the lower bound in (3.4). Hence, the risk-averse type $\bar{\pi}$ prefers the position $z(\bar{\pi})$ with the fee difference $\Delta(\bar{\pi})$, to any position $\hat{z} < z(\bar{\pi})$ with smaller fee difference $\hat{\Delta}$. The investor exploits his ability to separate by setting $\hat{z}$ close to the value $\frac{\hat{\Delta}}{2d}$, under which his net-of-fee risk exposure is zero, and $f(\hat{z}d)$ close to $f(z(\bar{\pi})d)$, with closeness becoming equality in the limit where ϵ goes to zero. The intuition in the case of $\hat{z}$ is that because the risk-neutral type is uninformed, the investor wants to have a net-of-fee risk exposure that is as close to zero as possible when facing that type. The intuition in the case of $f(\hat{z}d)$ is that raising the difference $\hat{\Delta}$ towards $\Delta(\bar{\pi})$ makes it less attractive for the risk-averse type to not observe the signal. Indeed, without the signal, the risk-averse type faces equal odds for the two payoff realizations, and hence her utility decreases in fee difference.

Separation becomes impossible if

$$2(1-\bar{\pi}) < e^{-\bar{\rho}K} < 4(1-\bar{\pi})\bar{\pi}, \tag{3.12}$$

i.e., (3.11) does not hold and (3.3) holds. Indeed, since $G_{\bar{\pi}}(\Delta)$ is decreasing for $\Delta > \Delta^*(\bar{\pi})$, and since $\Delta(\bar{\pi})$ lies in that region because it exceeds the lower bound in (3.4), the risk-averse type $\bar{\pi}$ prefers any $\Delta \in [\Delta^*(\bar{\pi}), \Delta(\bar{\pi}))$ to $\Delta(\bar{\pi})$. Hence, the fee difference $\hat{\Delta}$ for the risk-neutral type cannot lie in $[\Delta^*(\bar{\pi}), \Delta(\bar{\pi}))$. In must instead lie in the increasing part of $G_{\pi}(\Delta)$ and not exceed the level $\Delta_1(\bar{\pi}) < \Delta^*(\bar{\pi})$ defined by $G_{\bar{\pi}}(\Delta_1(\bar{\pi})) = G_{\bar{\pi}}(\Delta(\bar{\pi}))$. Equation (3.2) then implies

$$\begin{aligned}
&- e^{\bar{\rho}K}\left(\bar{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})} + (1-\bar{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\right)e^{-\bar{\rho}\Gamma(\bar{\pi})} \geq -\frac{1}{2}\left(e^{-\frac{\bar{\rho}}{2}\hat{\Delta}} + e^{\frac{\bar{\rho}}{2}\hat{\Delta}}\right)e^{-\bar{\rho}\hat{\Gamma}} \\
&\Leftrightarrow -e^{\bar{\rho}K}\left(\bar{\pi}e^{-\frac{\bar{\rho}}{2}\Delta_1(\bar{\pi})} + (1-\bar{\pi})e^{\frac{\bar{\rho}}{2}\Delta_1(\bar{\pi})}\right) \geq -\frac{1}{2}\left(e^{-\frac{\bar{\rho}}{2}\hat{\Delta}} + e^{\frac{\bar{\rho}}{2}\hat{\Delta}}\right) \\
&\Rightarrow -e^{\bar{\rho}K}\left(\bar{\pi}e^{-\frac{\bar{\rho}}{2}\Delta_1(\bar{\pi})} + (1-\bar{\pi})e^{\frac{\bar{\rho}}{2}\Delta_1(\bar{\pi})}\right) \geq -\frac{1}{2}\left(e^{-\frac{\bar{\rho}}{2}\Delta_1(\bar{\pi})} + e^{\frac{\bar{\rho}}{2}\Delta_1(\bar{\pi})}\right),
\end{aligned} \tag{3.13}$$

where the first step follows from the definitions of $(\Delta(\bar{\pi}), \Gamma(\bar{\pi}), \hat{\Delta}, \hat{\Gamma})$, the second step follows from $\Gamma(\bar{\pi}) = \hat{\Gamma}$ and $G_{\bar{\pi}}(\Delta_1(\bar{\pi})) = G_{\bar{\pi}}(\Delta(\bar{\pi}))$, and the third step follows from $\hat{\Delta} \leq \Delta_1(\bar{\pi})$. Equation (3.13) implies that the lower bound in (3.8) applies not only to $\Delta(\bar{\pi})$ but also to $\Delta_1(\bar{\pi})$. This would mean that $\Delta_1(\bar{\pi}) > \Delta^*(\bar{\pi})$, a contradiction. Because the investor cannot separate, he must set $\hat{z} = z(\bar{\pi})$ and $f(\hat{z}d) = f(z(\bar{\pi})d)$.

Completing the determination of the optimal fee is straightforward. The investor sets $f(-z(\bar{\pi})d)$

to zero and $\Delta(\bar{\pi}) \equiv f(z(\bar{\pi})d) - f(-z(\bar{\pi})d)$ equal to the lower bound in (3.4) (assuming that ϵ is small enough so that the constraint $f(z(\bar{\pi})d) - f(-z(\bar{\pi})d) \geq 2z(\bar{\pi})d$ is met). When separation is possible, the investor sets $z(\bar{\pi})$ close to the value $\frac{1}{2\rho d}\log\left(\frac{\bar{\pi}}{1-\bar{\pi}}\right) + \frac{\Delta(\bar{\pi})}{2d}$, with closeness becoming equality in the limit where ϵ goes to zero. This value yields the optimal net-of-fee risk exposure for posterior probability $\bar{\pi}$ for $S + d$. When separation is impossible, the investor sets $z(\bar{\pi})$ close to the value that yields the optimal net-of-fee risk exposure for a posterior probability that lies between $\bar{\pi}$ and $\frac{1}{2}$. This probability reflects the fact that the same position $z(\bar{\pi})$ is chosen by the informed risk-averse type $\bar{\pi}$ and the uninformed risk-neutral type. Pooling causes the investor to discount the risk-averse type's signal and to limit her position below the one in the absence of the risk-neutral type. The investor's decision can be interpreted as a risk limit, although it is a limit to which the manager adheres voluntarily given the fee structure.

Proposition 3.1 computes the limits of the optimal positions $(z(\bar{\pi}), \hat{z})$ and fee levels $(f(z(\bar{\pi})d)$, $f(-z(\bar{\pi})d), f(\hat{z}d), f(-\hat{z}d))$ when ϵ goes to zero. The proposition additionally extends the fee function $f(W)$ to values of W other than $(z(\bar{\pi})d, -z(\bar{\pi})d, \hat{z}d, -\hat{z}d)$ (while ensuring that the positions $(z(\bar{\pi}), \hat{z})$ are optimal when choice is over all values of z) and characterizes when the investor employs the manager. One way to extend $f(W)$ is to set it to zero for all values of W different from $(z(\bar{\pi})d, -z(\bar{\pi})d, \hat{z}d, -\hat{z}d)$. Proposition 3.1 extends $f(W)$ as a non-decreasing function.

Proposition 3.1. *When ϵ goes to zero, the investor's and manager's decisions converge to the following limits:*

- **[Separation]** *When $e^{-\bar{\rho}K} > 4(1-\bar{\pi})\bar{\pi}$, the investor employs the manager if*

$$e^{\frac{\rho\Delta(\bar{\pi})}{2}}\left[(1-\lambda)\sqrt{\bar{\pi}(1-\bar{\pi})} + \frac{\lambda}{2}\right] < \frac{1}{2}, \tag{3.14}$$

  where $\Delta(\bar{\pi}) = \frac{1}{\bar{\rho}}\log\left(\frac{2\bar{\pi}-e^{-\bar{\rho}K}}{e^{-\bar{\rho}K}-2(1-\bar{\pi})}\right) > 0$. If (3.14) is satisfied, then the risk-averse type $\bar{\pi}$ chooses $z(\bar{\pi}) = \frac{1}{2\rho d}\log\left(\frac{\bar{\pi}}{1-\bar{\pi}}\right) + \frac{\Delta(\bar{\pi})}{2d}$ and the risk-neutral type chooses $\hat{z} = \frac{\Delta(\bar{\pi})}{2d}$. An optimal fee for the investor is $f(W) = 0$ for $W < \frac{\Delta(\bar{\pi})}{2}$ and $f(W) = \Delta(\bar{\pi})$ for $W \geq \frac{\Delta(\bar{\pi})}{2}$.

- **[Pooling]** *When $2(1-\bar{\pi}) < e^{-\bar{\rho}K} < 4(1-\bar{\pi})\bar{\pi}$, the investor employs the manager if*

$$e^{\frac{\rho\Delta(\bar{\pi})}{2}}\sqrt{\left[(1-\lambda)\bar{\pi} + \frac{\lambda}{2}\right]\left[1-(1-\lambda)\bar{\pi} - \frac{\lambda}{2}\right]} < \frac{1}{2}, \tag{3.15}$$

  where $\Delta(\bar{\pi})$ is as in the previous case. If (3.15) is satisfied, then the risk-averse type $\bar{\pi}$ and the risk-neutral type choose $z(\bar{\pi}) = \frac{1}{2\rho d}\log\left(\frac{(1-\lambda)\bar{\pi}+\frac{\lambda}{2}}{1-(1-\lambda)\bar{\pi}-\frac{\lambda}{2}}\right) + \frac{\Delta(\bar{\pi})}{2d}$. An optimal fee for the investor is $f(W) = 0$ for $W < z(\bar{\pi})d$ and $f(W) = \Delta(\bar{\pi})$ for $W \geq z(\bar{\pi})d$.

In both cases of Proposition 3.1, the fee is a step function that starts from zero and ends at $\Delta(\bar{\pi}) > 0$. The location of the step is different in the two cases. When separation is possible (first case), the step occurs at $\frac{\Delta(\bar{\pi})}{2}$, which is the position chosen by the risk-neutral type. The risk-averse type $\bar{\pi}$ chooses a position that exceeds $\frac{\Delta(\bar{\pi})}{2}$. Both the risk-averse type $\bar{\pi}$ and the risk-neutral type are indifferent between their position and the position of the other type.[1] When separation is impossible (second case), the step occurs at $z(\bar{\pi})d$, which is the common position chosen by the two types. The pooling position $z(\bar{\pi})$ exceeds the position of the risk-neutral type under separation but is smaller than the position of the risk-averse type.

Imposing the lower bound $\epsilon > 0$ on the fee's slope and then taking the limit when ϵ goes to zero rules out separation for some parameter values. Indeed, when $\epsilon = 0$, the investor can employ the fee derived in the first case of Proposition 3.1 and induce the positions under separation even when the pooling condition $e^{-\bar{\rho}K} < 4(1-\bar{\pi})\bar{\pi}$ holds. This outcome, however, relies heavily on the manager's indifference across different values of z. Imposing $\epsilon > 0$ breaks this indifference and makes it impossible for the investor to achieve the ranking of positions under separation ($\hat{z} < z(\bar{\pi})$) when $e^{-\bar{\rho}K} < 4(1-\bar{\pi})\bar{\pi}$.

An increase in the cost K of observing the signal favors pooling. This is because the investor must pay a higher fee $f(z(\bar{\pi})d)$ to the risk-averse type conditional on high performance, to ensure that she observes the signal. That type is hence exposed to more risk, and the position of the risk-neutral type, which is less risky, may become more attractive to her. A decrease in signal informativeness, as measured by the posterior probability $\bar{\pi}$, also favors pooling. This is because the risk-averse type becomes less informed and hence less willing to take risk.

We finally consider two limit cases to which we return in subsequent sections. The first is when the risk-aversion coefficient $\bar{\rho}$ of the manager becomes large, while the risk-aversion coefficient ρ of the investor and the product $\bar{\rho}K$ are held constant. To motivate this case, suppose that the investor is an aggregate of a large number N of individual investors, each with risk-aversion coefficient ρ, and that the manager has also risk-aversion coefficient ρ. Define one share of the risky asset to pay off $(S+d)N$ or $(S-d)N$ instead of $S+d$ or $S-d$. The risk-aversion coefficient of the investor group is $\frac{\rho}{N}$. Redefine the numeraire so that one new unit is N old units. The payoff of one share then becomes $S+d$ or $S-d$, and the cost of observing the signal becomes $\frac{K}{N}$. Moreover, absolute risk-aversion coefficients are multiplied by N, so the risk aversion of the investor group becomes ρ and that of the manager becomes $N\rho \equiv \bar{\rho}$. Raising N hence raises the risk-aversion coefficient $\bar{\rho}$ of the manager, while the risk-aversion coefficient of the investor group remains equal to ρ. Moreover, the product of $\bar{\rho}$ times the cost of observing the signal remains constant, and so does the payoff of

[1] They are also indifferent between their position and any other position z that exceeds $\hat{z}$ and differs from $z(\bar{\pi})$. Such indifference, however, can be broken by setting the fee $f(W)$ to zero for all values of W other than $(z(\bar{\pi})d, \hat{z}d)$.

one share of the risky asset.

Proposition 3.1 implies that in the limit, $\Delta(\bar{\pi})$ is of order $\frac{1}{\bar{\rho}}$. Hence, the fee becomes negligible relative to the investor's wealth. The fee's contribution $\frac{\Delta(\bar{\pi})}{2d}$ to the optimal positions $\hat{z}$ and $z(\bar{\pi})$ also becomes negligible: the investor's risk exposure gross and net of the fee is approximately the same. Because the fee becomes negligible and the manager brings valuable information, the investor always employs the manager, i.e., (3.14) and (3.15) are always satisfied in the limit.

The second case adds small uncertainty to the first, and is particularly useful when integrating the model in a continuous-time setting. We take d to be small, the probability $\bar{\pi}$ of $S+d$ to be $\frac{1}{2}+\bar{\mu}d$, K to be $\frac{kd}{\rho}$, and $\bar{\rho}$ to be $N\rho$, where N is large and $(\bar{\mu}, k, \rho)$ are held constant. Proposition 3.1 implies that in the limit, $\Delta(\bar{\pi})$ is of order $\frac{1}{N}$ and hence small. Although the fee is small, the investor does not always employ the manager because the signal has small informational content. Equations (3.14) and (3.15) imply that the investor always employs the manager if $\frac{1}{N}$ is of order smaller than d^2, never employs her if $\frac{1}{N}$ is of order larger than d^2, and employs her depending on $(\bar{\mu}, k, \lambda)$ if $\frac{1}{N}$ and d^2 are of the same order. We do not provide detailed proofs for these results, but do so for their counterparts in the case of a continuum of signal values in the next section.

4 Symmetry and a Continuum of Signal Values

We denote the posterior probability of $S+d$ by π. Since there is a continuum of signal values, π takes values in an interval according to a continuous density $h(\pi)$. Symmetry implies that the interval is of the form $[1-\bar{\pi}, \bar{\pi}]$ for $\bar{\pi} > \frac{1}{2}$ and the density satisfies $h(\pi) = h(1-\pi)$.

The observations made in the case of two signal values generalize as follows: (i) if the investor invests directly, then his optimal position is $z = 0$; (ii) if the investor prefers to employ the manager, then he must be inducing the risk-averse type to observe the signal and choose a position that is non-zero for at least some signal values; (iii) if the investor optimally induces the risk-averse type π to choose a position $z(\pi)$, then he optimally induces the risk-averse type $1-\pi$ to choose $-z(\pi)$; (iv) only a long position $z(\pi) \geq 0$ can maximize the utility of a risk-averse type $\pi > \frac{1}{2}$; (v) if the investor optimally induces the risk-neutral type to choose a long position $\hat{z} \geq 0$, then he is indifferent between that choice and inducing $-\hat{z}$; and (vi) the position $\hat{z}$ does not exceed the maximum position of the risk-averse type.

Based on the above observations, the relevant fee levels are $f(z(\pi)d)$ and $f(-z(\pi)d)$ for $\pi \in [\frac{1}{2}, \bar{\pi}]$, $f(\hat{z}d)$ and $f(-\hat{z}d)$. These fee levels satisfy non-negativity, monotonicity and incentive compatibility (IC) constraints.

The (IC) constraint that concerns the risk-averse type's choice whether or not to observe the signal is

$$-2\int_{\frac{1}{2}}^{\bar{\pi}}\left[\pi e^{-\bar{\rho}[f(z(\pi)d)-K]}+(1-\pi)e^{-\bar{\rho}[f(-z(\pi)d)-K]}\right]h(\pi)d\pi\geq -e^{-\bar{\rho}f(0)}. \tag{4.1}$$

The left-hand side of (4.1) is the utility when observing the signal. The coefficient two is present because symmetry implies that the integral of the utility from $\frac{1}{2}$ to $\bar{\pi}$ is equal to that from $1-\bar{\pi}$ to $\frac{1}{2}$. The right-hand side of (4.1) is the utility when not observing the signal and employing the prior probability $\pi_0=\frac{1}{2}$. The optimal position in that case is zero, same as when observing the signal and having posterior $\pi=\frac{1}{2}$.

Consider next the (IC) constraints that concern the choice of position z by the risk-averse and risk-neutral types. As in the case of two signal values, we consider the (IC) constraints concerning long positions only. A risk-averse type $\pi\geq\frac{1}{2}$ prefers the position $z(\pi)$ to positions chosen by other risk-averse types if

$$-\left[\pi e^{-\bar{\rho}f(z(\pi)d)}+(1-\pi)e^{-\bar{\rho}f(-z(\pi)d)}\right]\geq\max_{\hat{\pi}\in[\frac{1}{2},\bar{\pi}]}\left\{-\left[\pi e^{-\bar{\rho}f(z(\hat{\pi})d)}+(1-\pi)e^{-\bar{\rho}f(-z(\hat{\pi})d)}\right]\right\}, \tag{4.2}$$

and prefers $z(\pi)$ to the position $\hat{z}$ of the risk-neutral type if

$$-\left[\pi e^{-\bar{\rho}f(z(\pi)d)}+(1-\pi)e^{-\bar{\rho}f(-z(\pi)d)}\right]\geq-\left[\pi e^{-\bar{\rho}f(\hat{z}d)}+(1-\pi)e^{-\bar{\rho}f(-\hat{z}d)}\right]. \tag{4.3}$$

The risk-neutral type prefers the position $\hat{z}$ to positions chosen by risk-averse types if

$$\frac{1}{2}\left[f(\hat{z}d)+f(-\hat{z}d)\right]\geq\max_{\pi\in[\frac{1}{2},\bar{\pi}]}\frac{1}{2}\left[f(z(\pi)d)+f(-z(\pi)d)\right]. \tag{4.4}$$

Generalizing the notation introduced in the case of two signal values, we denote by

$$\begin{aligned}\Delta(\pi)&\equiv f(z(\pi)d)-f(-z(\pi)d)\geq 0,\\ \Gamma(\pi)&\equiv\frac{1}{2}\left[f(z(\pi)d)+f(-z(\pi)d)\right],\end{aligned}$$

the fee difference and average fee, respectively, for the risk-averse type π.

The (IC) constraint (4.2) implies that the (IC) constraint (4.3) is met if there exists $\pi\in[\frac{1}{2},\bar{\pi}]$ such that $\hat{z}=z(\pi)$. The (IC) constraint (4.4) can be written as

$$\hat{\Gamma}\geq\max_{\pi\in[\frac{1}{2},\bar{\pi}]}\Gamma(\pi). \tag{4.5}$$

Lemma 4.1 derives necessary and sufficient conditions for the (IC) constraint (4.2) to hold. We denote by $U(\pi)$ the utility of the risk-averse type π:

$$U(\pi) \equiv -\left[\pi e^{-\bar{\rho}f(z(\pi)d)} + (1-\pi)e^{-\bar{\rho}f(-z(\pi)d)}\right]$$
$$= -\left[\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}\right] e^{-\bar{\rho}\Gamma(\pi)}. \tag{4.6}$$

Lemma 4.1. *The (IC) constraint (4.2) holds if and only if for all $\pi \in [\frac{1}{2}, \bar{\pi}]$ the following conditions hold:*

(i) $\Delta(\pi)$ is non-decreasing.

(ii) If $\Delta(\pi)$ is continuous at π, then $U(\pi)$ is differentiable at π and

$$U'(\pi) = \left[e^{\frac{\bar{\rho}}{2}\Delta(\pi)} - e^{-\frac{\bar{\rho}}{2}\Delta(\pi)}\right] e^{-\bar{\rho}\Gamma(\pi)}. \tag{4.7}$$

If instead $\Delta(\pi)$ is discontinuous at π, then $U(\pi)$ has left- and right-derivatives at π, which are given by substituting the left- and right-limits of $(\Delta(\pi), \Gamma(\pi))$, respectively, into (4.7).

Condition (i) of Lemma 4.1 is a sorting condition. When the risk-averse type is more optimistic that the state $S + d$ will occur, she must receive a higher fee in that state relative to the state $S - d$. Condition (ii) of Lemma 4.1 is an envelope condition. If the functions $(\Delta(\pi), \Gamma(\pi))$ are differentiable with respect to π, then (4.7) can be derived by differentiating $U(\pi)$, given by (4.6), and combining with the first-order condition implied by (4.2). Condition (ii) does not require, however, differentiability or even continuity of $(\Delta(\pi), \Gamma(\pi))$.

Combining (4.6) and (4.7), we find

$$\frac{U'(\pi)}{U(\pi)} = -\frac{e^{\frac{\bar{\rho}}{2}\Delta(\pi)} - e^{-\frac{\bar{\rho}}{2}\Delta(\pi)}}{\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}}, \tag{4.8}$$

which integrates to

$$U(\pi) = U(\bar{\pi}) \exp\left[\int_{\pi}^{\bar{\pi}} \frac{e^{\frac{\bar{\rho}}{2}\Delta(\pi')} - e^{-\frac{\bar{\rho}}{2}\Delta(\pi')}}{\pi' e^{-\frac{\bar{\rho}}{2}\Delta(\pi')} + (1-\pi')e^{\frac{\bar{\rho}}{2}\Delta(\pi')}} d\pi'\right]. \tag{4.9}$$

Using (4.9), we can write the (IC) constraint (4.1) in terms of the function $\Delta(\pi)$.

Lemma 4.2. *The (IC) constraint (4.1) is equivalent to*

$$2e^{\bar{\rho}K} \int_{\frac{1}{2}}^{\bar{\pi}} \exp\left[-\int_{\frac{1}{2}}^{\pi} \frac{e^{\frac{\bar{\rho}}{2}\Delta(\pi')} - e^{-\frac{\bar{\rho}}{2}\Delta(\pi')}}{\pi' e^{-\frac{\bar{\rho}}{2}\Delta(\pi')} + (1-\pi')e^{\frac{\bar{\rho}}{2}\Delta(\pi')}} d\pi'\right] h(\pi) d\pi \leq 1, \tag{4.10}$$

and yields the following bounds on K and $\Delta(\bar{\pi})$:

$$e^{-\bar{\rho}K} - 2\left(1 - 2\int_{\frac{1}{2}}^{\bar{\pi}} \pi h(\pi) d\pi\right) > 0, \tag{4.11}$$

$$e^{\bar{\rho}\Delta(\bar{\pi})} \geq \frac{4\int_{\frac{1}{2}}^{\bar{\pi}} \pi h(\pi) d\pi - e^{-\bar{\rho}K}}{e^{-\bar{\rho}K} - 2\left(1 - 2\int_{\frac{1}{2}}^{\bar{\pi}} \pi h(\pi) d\pi\right)}. \tag{4.12}$$

Equation (4.11) requires that the cost K of observing the signal is not too large. Given such a value of K, (4.10) can be interpreted as a lower bound involving the fee difference $\Delta(\pi)$. The intuition is the same as in the case of two signal values: to induce the risk-averse type to observe the signal, the investor must pay her sufficiently more when she earns a high return than when she earns a low return. The lower bound in (4.10) concerns the entire function $\Delta(\pi)$. For example, (4.10) can hold even when $\Delta(\pi)$ is close to zero for π close to and larger than $\frac{1}{2}$, provided that at the same time $\Delta(\pi)$ is sufficiently large for larger values of π. Using the monotonicity of $\Delta(\pi)$ shown in Lemma 4.1, we can sharpen the lower bound implied by (4.10) to one pertaining only to the scalar $\Delta(\bar{\pi})$. The latter lower bound, given in (4.12), generalizes to a continuum of signal values the lower bound in (3.4) obtained for two signal values. Indeed, the case of two signal values can be derived from the continuum case by taking the density $h(\pi)$ to be a Dirac mass $\frac{1}{2}$ at each of the points $1 - \bar{\pi}$ and $\bar{\pi}$. Substituting that density in (4.12), we find (3.4).

We next examine whether the investor can separate the risk-averse type $\bar{\pi}$ and the risk-neutral type in terms of their positions. Under separation, the position $z(\bar{\pi})$ of the risk-averse type $\bar{\pi}$ exceeds the position $\hat{z}$ of the risk-neutral type, and hence the investor can benefit from the risk-averse type's most informative signal. Lemma 4.3 shows that separation is impossible when the lower bound in (4.12) exceeds $\Delta^*(\bar{\pi})$, for $\Delta^*(\bar{\pi})$ defined by (3.10).

Lemma 4.3. *When the pooling condition*

$$2\left(1 - 2\int_{\frac{1}{2}}^{\bar{\pi}} \pi h(\pi) d\pi\right) < e^{-\bar{\rho}K} < 2\left(\bar{\pi} + 2\int_{\frac{1}{2}}^{\bar{\pi}} \pi h(\pi) d\pi - 4\bar{\pi}\int_{\frac{1}{2}}^{\bar{\pi}} \pi h(\pi) d\pi\right) \tag{4.13}$$

holds, $\hat{z}$ cannot be smaller than $z(\bar{\pi})$.

A sketch of the proof of Lemma 4.3 is as follows. If the position $z(\bar{\pi})$ of the risk-averse type $\bar{\pi}$ exceeds the position $\hat{z}$ of the risk-neutral type, then the fee differences must rank in the same manner, i.e., $\Delta(\bar{\pi}) > \hat{\Delta}$, because the monotonicity of the fee implies

$$\Delta(\bar{\pi}) \equiv f(z(\bar{\pi})d) - f(-z(\bar{\pi})d) \geq f(\hat{z}d) - f(-\hat{z}d) + 2\epsilon(z(\bar{\pi}) - \hat{z})d > f(\hat{z}d) - f(-\hat{z}d) \equiv \hat{\Delta}.$$

Since the function $G_{\bar{\pi}}(\Delta)$, which characterizes how the utility of the risk-averse type $\bar{\pi}$ depends on the fee difference Δ holding the average fee constant, is decreasing for $\Delta > \Delta^*(\bar{\pi})$, and since $\Delta(\bar{\pi})$ lies in that region because it exceeds the lower bound in (4.12), the risk-averse type $\bar{\pi}$ prefers any $\Delta \in [\Delta^*(\bar{\pi}), \Delta(\bar{\pi}))$ to $\Delta(\bar{\pi})$. Since, in addition, (4.5) implies $\hat{\Gamma} \geq \Gamma(\bar{\pi})$, the risk-averse type $\bar{\pi}$ can be induced to accept $(\Delta(\bar{\pi}), \Gamma(\bar{\pi}))$ over $(\hat{\Delta}, \hat{\Gamma})$ only if $\hat{\Delta}$ is sufficiently smaller than $\Delta^*(\bar{\pi})$.

Consider next the risk-averse type $\hat{\pi} \in [\frac{1}{2}, \bar{\pi})$ such by $\Delta^*(\hat{\pi}) = \hat{\Delta}$. Since the optimal fee difference for that type, holding the average fee Γ constant, is $\Delta^*(\hat{\pi}) = \hat{\Delta}$, and since (4.5) implies $\hat{\Gamma} \geq \Gamma(\hat{\pi})$, that type can be induced to accept $(\Delta(\hat{\pi}), \Gamma(\hat{\pi}))$ over $(\hat{\Delta}, \hat{\Gamma})$ only if $\Delta(\hat{\pi}) = \Delta^*(\hat{\pi}) = \hat{\Delta}$ and $\Gamma(\hat{\pi}) = \hat{\Gamma} \geq \max_{\pi \in [\frac{1}{2}, \bar{\pi}]} \Gamma(\pi)$. In the proof of Lemma 4.3, however, we show that under $\Delta(\hat{\pi}) = \Delta^*(\hat{\pi})$ and $\Gamma(\hat{\pi}) \geq \max_{\pi \in [\frac{1}{2}, \bar{\pi}]} \Gamma(\pi)$, the (IC) constraint (4.10) can hold only if $\Delta^*(\bar{\pi})$ exceeds the lower bound in (4.12).

When the pooling condition (4.13) holds, the position $\hat{z}$ of the risk-neutral type coincides with that of an interval $[\pi^*, \bar{\pi}]$ of risk-averse types, for $\pi^* \in (\frac{1}{2}, \bar{\pi})$. Thus, the investor pools the risk-neutral type with multiple risk-averse types. To explain the intuition for this result, we start by observing that the position $z(\pi)$ of the risk-averse type π is non-decreasing in π: if $z(\pi)$ were decreasing, then $\Delta(\pi)$ would also have to be decreasing because of the monotonicity of the fee, but this is ruled out by Lemma 4.1.

Since $\hat{z}$ does not exceed $\max_{\pi \in (\frac{1}{2}, \bar{\pi}]} z(\pi) = z(\bar{\pi})$ and is not smaller than $z(\bar{\pi})$, it has to equal $z(\bar{\pi})$. Thus, in choosing $z(\bar{\pi})$, the investor determines the position of both the risk-averse type $\bar{\pi}$ and the risk-neutral type. If $z(\pi) < z(\bar{\pi})$ for all $\pi \in (\frac{1}{2}, \bar{\pi})$, then $z(\bar{\pi})$ is equal to the position of only those two types, and the investor determines the position of types $\pi \in (\frac{1}{2}, \bar{\pi})$ through a separate choice of $z(\pi)$. The investor sets $z(\bar{\pi})$ and $z(\pi)$ for $\pi \in (\frac{1}{2}, \bar{\pi})$ close to the values that optimize his net-of-fee risk exposure for posterior probabilities that reflect the types that he is facing (with closeness becoming equality in the limit where ϵ goes to zero). The investor is subject to the constraint that $z(\pi)$ must be non-decreasing.

In the case of $z(\bar{\pi})$, the unconstrained optimal net-of-fee risk exposure is zero. This is because the investor is effectively facing only the risk-neutral type: that type has probability λ, while the risk-averse type $\bar{\pi}$ has infinitesimal probability given the continuous density $h(\pi)$. In the case of $z(\pi)$ for $\pi \in (\frac{1}{2}, \bar{\pi})$, the unconstrained optimal risk exposure is positive, and is bounded away from zero when π is close to $\bar{\pi}$. Since the unconstrained choices yield a drop of $z(\pi)$ at $\bar{\pi}$, but $z(\pi)$ must be non-decreasing, there must be pooling for π close to $\bar{\pi}$. The lower bound π^* of the pooling interval is such that the unconstrained choice for risk-averse type π^* is the same as for the types in the pooling interval, i.e., the risk-averse types $\pi \in [\pi^*, \bar{\pi}]$ and the risk-neutral type. This condition yields a unique $\pi^* \in (\frac{1}{2}, \bar{\pi})$.

Proposition 4.1 determines the pooling interval $[\pi^*, \bar{\pi}]$, the optimal positions $(z(\pi), \hat{z})$, and the condition under which the investor employs the manager, all in the limit when ϵ goes to zero. The optimal positions are expressed in terms of the fee difference $\Delta(\pi)$. The employment condition is expressed in terms of $\Delta(\pi)$ and the average fee $\Gamma(\pi)$. Proposition 4.1 determines properties of $(\Delta(\pi), \Gamma(\pi))$ but does not characterize these functions fully. Proposition 4.2 provides a fuller characterization in the case where the manager's risk-aversion coefficient $\bar{\rho}$ is large and uncertainty is small.

Proposition 4.1. *Suppose that the pooling condition (4.13) holds. When ϵ goes to zero, the limits of the investor's and manager's decisions have the following properties:*

- *The investor employs the manager if*

$$(1-\lambda)\left[\int_{\frac{1}{2}}^{\pi^*} e^{\rho\Gamma(\pi)}\sqrt{\pi(1-\pi)}h(\pi)d\pi + e^{\frac{\rho\Delta(\bar{\pi})}{2}}\sqrt{\pi^*(1-\pi^*)}\frac{\int_{\pi^*}^{\bar{\pi}}\left(\pi-\frac{1}{2}\right)h(\pi)d\pi}{\pi^*-\frac{1}{2}}\right] < \frac{1}{4}, \quad (4.14)$$

  where $\pi^* \in (\frac{1}{2}, \bar{\pi})$ is the unique solution of

$$2(1-\lambda)\int_{\pi^*}^{\bar{\pi}}(\pi-\pi^*)h(\pi)d\pi + \lambda\left(\frac{1}{2}-\pi^*\right) = 0. \quad (4.15)$$

- *If (4.14) is satisfied, then the risk-averse types $\pi \in [\pi^*, \bar{\pi}]$ and the risk-neutral type choose*

$$z(\bar{\pi}) = \frac{1}{2\rho d}\log\left(\frac{\pi^*}{1-\pi^*}\right) + \frac{\Delta(\bar{\pi})}{2d}, \quad (4.16)$$

  and have fee difference $\Delta(\bar{\pi})$ and average fee $\Gamma(\bar{\pi}) = \frac{\Delta(\bar{\pi})}{2}$. The risk-averse types $\pi \in [\frac{1}{2}, \pi^*)$ choose

$$z(\pi) = \frac{1}{2\rho d}\log\left(\frac{\pi}{1-\pi}\right) + \frac{\Delta(\pi)}{2d}. \quad (4.17)$$

  For $\pi \in (\frac{1}{2}, \bar{\pi}]$, the fee difference $\Delta(\pi)$ is continuous and satisfies $\Delta(\pi) > \Delta^*(\pi)$. Moreover, $\Delta(\frac{1}{2}) = 0$.

Equation (4.15) determines the pooling interval $[\pi^*, \bar{\pi}]$. The length of that interval increases in the probability λ of the risk-neutral type. Equations (4.16) and (4.17) determine the investor's net-of-fee risk exposure $z(\pi) - \frac{\Delta(\pi)}{2d}$. Net-of-fee exposure is continuous in π, zero for $\pi = \frac{1}{2}$, increasing for $\pi \in [\frac{1}{2}, \pi^*]$, and constant for $\pi \in [\pi^*, \bar{\pi}]$. For $\pi \in [\frac{1}{2}, \pi^*]$, the investor sets net-of-fee exposure to its unconstrained optimum conditionally on facing the risk-averse type π only. For $\pi \in [\pi^*, \bar{\pi}]$, he sets it to its unconstrained optimum conditionally on facing risk-averse types $\pi \in [\pi^*, \bar{\pi}]$ and the risk-neutral type.

The fee difference $\Delta(\pi)$ is constant in the pooling interval $[\pi^*, \bar{\pi}]$. Proposition 4.1 establishes additionally that $\Delta(\pi)$ is continuous over $(\frac{1}{2}, \bar{\pi}]$ and exceeds the level $\Delta^*(\pi)$ that the risk-averse type π would choose if she could vary fee difference holding average fee constant. The risk-averse type π would face the latter choice problem if the fee were affine $f(W) = \gamma(W - \bar{W})$. Hence, Proposition 4.1 implies that an affine fee is not optimal.

The intuition why the investor over-exposes the manager to risk is as follows. Over-exposure for the risk-averse type $\bar{\pi}$ is necessary to provide the manager with ex-ante incentives to observe the signal, given that the signal's cost exceeds the threshold implied by the pooling condition (4.13). If, in addition, a risk-averse type $\pi < \bar{\pi}$ were under-exposed, then by raising her exposure (while keeping it strictly below that of type $\bar{\pi}$) the investor could benefit in two ways. He would improve ex-ante incentives to observe the signal, relaxing the (IC) constraint (4.10). He could also lower the average fee $\Gamma(\pi)$ paid to type π, and by extension to all types $[\frac{1}{2}, \pi)$ through the (IC) constraint (4.2), because he would be bringing type π closer to her optimal level of risk.

Continuity of $\Delta(\pi)$ follows from the over-exposure result. Consider a point $\hat{\pi}$ at which $\Delta(\pi)$ increases discontinuously. The investor could save on fees by lowering the exposure of types who are close to and larger than $\hat{\pi}$, and for whom over-exposure is significant, and by raising the exposure of types who are close to and smaller than $\hat{\pi}$, and for whom over-exposure is discontinuously less significant.

We next return to the two limit cases considered in the previous section. The first is when the risk-aversion coefficient $\bar{\rho}$ of the manager becomes large, while the risk-aversion coefficient ρ of the investor and the product $\bar{\rho}K$ are held constant. In that limit, the fee difference $\Delta(\pi)$ is of order $\frac{1}{\bar{\rho}}$, as this suffices to make the (IC) constraint (4.10) hold as an equality when K is also of order $\frac{1}{\bar{\rho}}$. Hence, $\Delta(\pi)$ becomes negligible, and the position $z(\pi)$ characterizes the investor's risk exposure not only gross but also net of the fee. Because the fee becomes negligible and the manager brings valuable information, the investor always employs the manager.

The second limit case is derived from the first by taking asset payoff uncertainty to be small. The probability π of $S+d$ is $\frac{1}{2}+\mu d$, K is $\frac{kd}{\bar{\rho}}$, and $\bar{\rho}$ is $N\rho$, where N is large, d is small, and (μ, k, ρ) are held constant. The variable μ takes values in the interval $[-\bar{\mu}, \bar{\mu}]$ with density $\hat{h}(\mu)$. Since $h(\pi)$ is symmetric around $\frac{1}{2}$, $\hat{h}(\mu)$ is symmetric around zero. Proposition 4.2 shows that if $\hat{h}(\mu)$ is non-decreasing in $(0, \bar{\mu}]$, then the fee difference $\Delta(\pi)$ is constant in $(\frac{1}{2}, \bar{\pi}]$ for sufficiently small d and large N. The proposition additionally computes the constant value of $\Delta(\pi)$ and shows that it is of order $\frac{1}{N}$; extends the fee function $f(W)$ to all values of W; and provides a fuller characterization of when the investor employs the manager.

Proposition 4.2. *Suppose that the pooling condition (4.13) holds. Suppose also that* $\pi = \frac{1}{2} + \mu d$, $K = \frac{kd}{\bar{\rho}}$, *and* $\bar{\rho} = N\rho$, *with* d *small and* N *large, and that the density* $\hat{h}(\mu)$ *is non-decreasing in* $[0, \bar{\mu}]$. *When* ϵ *goes to zero, the limits of the investor's and manager's decisions have the following properties additional to those derived in Proposition 4.1:*

- *The investor always employs the manager if* $\frac{1}{N}$ *is of order smaller than* d^2, *and never employs her if* $\frac{1}{N}$ *is of order larger than* d^2. *If* $\frac{1}{N}$ *and* d^2 *are of the same order, then the investor employs the manager when*

$$\mathcal{Z}(1-\lambda)\left[\int_0^{\mu^*} \mu^2 \hat{h}(\mu)d\mu + \mu^* \int_{\mu^*}^{\bar{\mu}} \mu \hat{h}(\mu)d\mu\right] > \frac{1}{8}\log\left(\frac{4\int_0^{\bar{\mu}} \hat{h}(\mu)d\mu + k}{4\int_0^{\bar{\mu}} \hat{h}(\mu)d\mu - k}\right), \tag{4.18}$$

where $\mathcal{Z}$ *denotes the limit of* Nd^2, *and* $\mu^* \in (0, \bar{\mu})$ *is the unique solution of*

$$2(1-\lambda)\int_{\mu^*}^{\bar{\mu}} (\mu - \mu^*)\hat{h}(\mu)d\mu - \lambda\mu^* = 0. \tag{4.19}$$

- *If the investor employs the manager, then* $(\Delta(\pi), \Gamma(\pi))$ *are constant in* $(\frac{1}{2}, \bar{\pi}]$ *and given by*

$$\Delta(\pi) = \frac{1}{N\rho}\log\left(\frac{4\int_0^{\bar{\mu}} \hat{h}(\mu)d\mu + k}{4\int_0^{\bar{\mu}} \hat{h}(\mu)d\mu - k}\right) + o\left(\frac{1}{N}\right) \tag{4.20}$$

and $\Gamma(\pi) = \frac{\Delta(\pi)}{2}$, *where* $o\left(\frac{1}{N}\right)$ *denotes terms of order smaller than* $\frac{1}{N}$. *Moreover,* $\Gamma(\frac{1}{2})$ *is given by*

$$e^{-\bar{\rho}\Gamma(\frac{1}{2})} = \frac{e^{-\bar{\rho}\Delta(\bar{\pi})} + 1}{2}. \tag{4.21}$$

- *An optimal fee for the investor is* $f(W) = 0$ *for* $W < -\frac{\Delta(\bar{\pi})}{2}$, $f(W) = \Gamma(\frac{1}{2})$ *for* $W \in [-\frac{\Delta(\bar{\pi})}{2}, -\frac{\Delta(\bar{\pi})}{2}]$, *and* $f(W) = \Delta(\bar{\pi})$ *for* $W > \frac{\Delta(\bar{\pi})}{2}$.

5 General case

To be written.

Appendix

Proof of Proposition 3.1. If the investor invests directly, then his utility is

$$-\frac{1}{2}\left[e^{-\rho z d}+e^{\rho z d}\right]. \tag{A.1}$$

Maximizing (A.1) with respect to z, we find $z=0$.

Suppose next that the investor prefers to employ the manager. As noted in Section 3, he induces the risk-averse type to observe the signal and to choose a non-zero position. As also noted in Section 3, he induces the risk-averse type $\bar{\pi}$ to choose a long position $z(\bar{\pi})>0$ and the risk-averse type $1-\bar{\pi}$ to choose the opposite short position $-z(\bar{\pi})$. Moreover, he induces the risk-neutral type to choose a long position $\hat{z}\geq 0$ and its opposite short position $-\hat{z}$ with equal probabilities.

Given positions $(z(\bar{\pi}),\hat{z})$ and fee levels $(f(z(\bar{\pi})d), f(-z(\bar{\pi})d), f(\hat{z}d), f(-\hat{z}d))$, the investor's utility is

$$U=-(1-\lambda)\left[\bar{\pi}e^{-\rho[z(\bar{\pi})d-f(z(\bar{\pi})d)]}+(1-\bar{\pi})e^{-\rho[-z(\bar{\pi})d-f(-z(\bar{\pi})d)]}\right]-\frac{\lambda}{2}\left[e^{-\rho[\hat{z}d-f(\hat{z}d)]}+e^{-\rho[-\hat{z}d-f(-\hat{z}d)]}\right]. \tag{A.2}$$

The investor maximizes (A.2) over $(z(\bar{\pi}),\hat{z})$ and $(f(z(\bar{\pi})d), f(-z(\bar{\pi})d), f(\hat{z}d), f(-\hat{z}d))$, subject to $z(\pi)\geq 0$, $\hat{z}\geq 0$, the (IC) constraints (3.2), (3.5) and (3.6), and the non-negativity and monotonicity of the fee. We refer to this optimization problem as $(\mathcal{P})$.

We determine the optimal values of $(z(\bar{\pi}),\hat{z})$ and $(f(z(\bar{\pi})d), f(-z(\bar{\pi})d), f(\hat{z}d), f(-\hat{z}d))$ by showing a number of properties that these must satisfy. A first property is $\hat{z}\leq z(\bar{\pi})$. Suppose, by contradiction, that $\hat{z}>z(\bar{\pi})$, and consider the investor's alternative choice of positions $(z(\bar{\pi}),z(\bar{\pi}))$ and fee levels $(g(z(\bar{\pi})d), g(-z(\bar{\pi})d), g(z(\bar{\pi})d), g(-z(\bar{\pi})d))$ defined by

$$g(-z(\bar{\pi})d)=f(-\hat{z}d),$$
$$g(z(\bar{\pi})d)=f(-\hat{z}d)+f(z(\bar{\pi})d)-f(-z(\bar{\pi})d).$$

Under $(z(\bar{\pi}),z(\bar{\pi}))$ and $(g(z(\bar{\pi})d), g(-z(\bar{\pi})d), g(z(\bar{\pi})d), g(-z(\bar{\pi})d))$, the risk-averse type $\bar{\pi}$ and the risk-neutral type hold the same position $z(\bar{\pi})$. Moreover, the fee levels $g(z(\bar{\pi})d)$ and $g(-z(\bar{\pi})d)$ that both types receive are smaller than $f(z(\bar{\pi})d)$ and $f(-z(\bar{\pi})d)$, respectively, because $f(-z(\bar{\pi})d)\geq f(-\hat{z}d)+\epsilon(\hat{z}-z(\bar{\pi}))d>f(-\hat{z}d)$. The fee levels $(g(z(\bar{\pi})d), g(-z(\bar{\pi})d), g(z(\bar{\pi})d), g(-z(\bar{\pi})d))$ satisfy the (IC) constraints and non-negativity and monotonicity. The (IC) constraints (3.5) and (3.6) are satisfied because fee levels are identical for the risk-averse and the risk-neutral type. The (IC) constraint (3.2) and monotonicity are satisfied because $g(z(\bar{\pi})d)-g(-z(\bar{\pi})d)=f(z(\bar{\pi})d)-$

$f(-z(\bar{\pi})d)$. Non-negativity is satisfied because $f(z(\bar{\pi})d) - f(-z(\bar{\pi})d) > 0$. Since payments to the risk-averse type do not increase under the investor's alternative choice, (A.2) implies that the investor is better off under the original choice only if

$$\begin{aligned} & -\left[e^{-\rho[\hat{z}d - f(\hat{z}d)]} + e^{-\rho[-\hat{z}d - f(-\hat{z}d)]}\right] \geq -\left[e^{-\rho[z(\bar{\pi})d - g(z(\bar{\pi})d)]} + e^{-\rho[-z(\bar{\pi})d - g(-z(\bar{\pi})d)]}\right] \\ & \Leftrightarrow e^{\rho[-\hat{z}d + f(\hat{z}d) - f(-\hat{z}d)]} + e^{\rho\hat{z}d} \leq e^{\rho[-z(\bar{\pi})d + f(z(\bar{\pi})d) - f(-z(\bar{\pi})d)]} + e^{\rho z(\bar{\pi})d}, \end{aligned} \tag{A.3}$$

where the second step follows from the first by multiplying both sides by $e^{-\rho f(-\hat{z}d)}$ and using the definitions of $g(z(\bar{\pi})d)$ and $g(-z(\bar{\pi})d)$. Since

$$f(\hat{z}d) - f(-\hat{z}d) \geq f(z(\bar{\pi})d) - f(-z(\bar{\pi})d) + 2\epsilon(\hat{z} - z(\bar{\pi}))d > f(z(\bar{\pi})d) - f(-z(\bar{\pi})d),$$

(A.3) can hold only if

$$\begin{aligned} & e^{\rho[-\hat{z}d + f(z(\bar{\pi})d) - f(-z(\bar{\pi})d)]} + e^{\rho\hat{z}d} \leq e^{\rho[-z(\bar{\pi})d + f(z(\bar{\pi})d) - f(-z(\bar{\pi})d)]} + e^{\rho z(\bar{\pi})d} \\ & \Leftrightarrow e^{\rho z(\bar{\pi})d}\left[e^{\rho(\hat{z} - z(\bar{\pi}))d} - 1\right] \leq e^{\rho[-z(\bar{\pi})d + f(z(\bar{\pi})d) - f(-z(\bar{\pi})d)]}\left[1 - e^{-\rho(\hat{z} - z(\bar{\pi}))d}\right] \\ & \Leftrightarrow e^{\rho z(\bar{\pi})d} e^{\rho(\hat{z} - z(\bar{\pi}))d} \leq e^{\rho[-z(\bar{\pi})d + f(z(\bar{\pi})d) - f(-z(\bar{\pi})d)]}. \end{aligned} \tag{A.4}$$

Since $\hat{z} > z(\bar{\pi})$, (A.4) can hold only if

$$z(\bar{\pi})d < -z(\bar{\pi})d + f(z(\bar{\pi})d) - f(-z(\bar{\pi})d) \Leftrightarrow z(\bar{\pi})d - f(z(\bar{\pi})d) < -z(\bar{\pi})d - f(-z(\bar{\pi})d) < 0.$$

Substituting into (A.2), we find

$$U < -(1-\lambda) - \frac{\lambda}{2}\left[e^{-\rho[\hat{z}d - f(\hat{z}d)]} + e^{-\rho[-\hat{z}d - f(-\hat{z}d)]}\right] \leq -(1-\lambda) - \lambda e^{\rho\frac{f(\hat{z}d) + f(-\hat{z}d)}{2}} \leq -1,$$

where the second step follows from the convexity of the exponential function and the third step because the fee is non-negative. That $U < -1$ is a contradiction because the investor prefers to employ the manager and receives utility $U = -1$ if he invests directly.

A second property is that the (IC) constraint (3.6) must hold as an equality. Suppose, by contradiction, that (3.6) holds as a strict inequality, in which case $\hat{z} < z(\bar{\pi})$. A (slight) reduction in $f(-\hat{z}d)$ can then raise the investor's utility while preserving the (IC), non-negativity, and monotonicity constraints. The (IC) constraint (3.6) is preserved because it holds strictly, and the (IC) constraints (3.2) and (3.5) are preserved because the right-hand side of each is increasing in $f(-\hat{z}d)$. Since (3.6) holds as a strict inequality,

$$f(-\hat{z}d) - f(-z(\bar{\pi})d) > f(z(\bar{\pi})d) - f(\hat{z}d) \geq \epsilon(z(\bar{\pi}) - \hat{z})d,$$

where the second step follows from the monotonicity constraint for fee levels $(f(z(\bar{\pi})d), f(\hat{z}d))$. Since $f(-\hat{z}d) - f(-z(\bar{\pi})d)$ exceeds $\epsilon(z(\bar{\pi}) - \hat{z})$, reducing $f(-\hat{z}d)$ preserves the monotonicity constraint. It also preserves the non-negativity constraint because $f(-z(\bar{\pi})d) \geq 0$.

A third property is $f(-z(\bar{\pi})d) = 0$. Indeed, if $f(-z(\bar{\pi})d) > 0$, then an equal (slight) reduction in all fee levels $(f(z(\bar{\pi})d), f(-z(\bar{\pi})d), f(\hat{z}d), f(-\hat{z}d))$ raises the investor's utility while preserving the (IC), non-negativity, and monotonicity constraints.

Suppose next that the separating condition $e^{-\bar{\rho}K} > 4(1-\bar{\pi})\bar{\pi}$ is satisfied. Consider the relaxed version of the optimization problem $(\mathcal{P})$ without the (IC) constraint (3.5). Since (3.6) holds as an equality,

$$f(z(\bar{\pi})d) - f(\hat{z}d) = f(-\hat{z}d) - f(-z(\bar{\pi})d). \tag{A.5}$$

If the two sides of (A.5) exceed $\epsilon(z(\bar{\pi}) - \hat{z})d$, then the investor can raise his utility through the following changes: reduce $f(-\hat{z}d)$ by a small $\phi > 0$; raise $f(\hat{z}d)$ by ϕ; raise $\hat{z}$ by $\frac{\phi}{d}$; and lower $f(z(\bar{\pi})d)$ so that the (IC) constraint (3.2) is preserved. These changes leave the component

$$-\frac{\lambda}{2}\left[e^{-\rho[\hat{z}d - f(\hat{z}d)]} + e^{-\rho[-\hat{z}d - f(-\hat{z}d)]}\right]$$

of utility corresponding to the risk-neutral type unchanged, but raise the component

$$-(1-\lambda)\left[\bar{\pi}e^{-\rho[z(\bar{\pi})d - f(z(\bar{\pi})d)]} + (1-\bar{\pi})e^{-\rho[-z(\bar{\pi})d - f(-z(\bar{\pi})d)]}\right]$$

corresponding to the risk-averse type $\bar{\pi}$. The changes preserve the (IC) constraint (3.6) because the left-hand side remains the same while the right-hand side decreases. They preserve the (IC) constraint (3.2) because the second term in the maximum in the right-hand side decreases while still exceeding the first term for small ϕ. They preserve the non-negativity and monotonicity constraints because the two sides of (A.5) remain larger than $\epsilon(z(\bar{\pi}) - \hat{z})d$. Therefore,

$$\begin{aligned} f(\hat{z}d) &= f(z(\bar{\pi})d) - \epsilon(z(\bar{\pi}) - \hat{z})d, \\ f(-\hat{z}d) &= f(-z(\bar{\pi})d) + \epsilon(z(\bar{\pi}) - \hat{z})d. \end{aligned}$$

Using these equations and $f(-z(\bar{\pi})d) = 0$, we can write (A.2) as

$$U = -(1-\lambda)\left[\bar{\pi}e^{-\rho[z(\bar{\pi})d - f(z(\bar{\pi})d)]} + (1-\bar{\pi})e^{\rho z(\bar{\pi})d}\right] - \frac{\lambda}{2}\left[e^{-\rho[\hat{z}d - f(z(\bar{\pi})d) + \epsilon(z(\bar{\pi}) - \hat{z})d]} + e^{-\rho[-\hat{z}d - \epsilon(z(\bar{\pi}) - \hat{z})d]}\right]. \tag{A.6}$$

Equation (A.6) must be maximized over $(z(\bar{\pi}), \hat{z}, f(z(\bar{\pi})d))$ subject to $0 \leq \hat{z} \leq z(\bar{\pi})$ and the (IC) constraint (3.2), which takes the form

$$-\left[\bar{\pi}e^{-\bar{\rho}[f(z(\bar{\pi})d)-K]} + (1-\bar{\pi})e^{\bar{\rho}K}\right] \geq -\frac{1}{2}\left[e^{-\bar{\rho}[f(z(\bar{\pi})d)-\epsilon(z(\bar{\pi})-\hat{z})d]} + e^{-\bar{\rho}\epsilon(z(\bar{\pi})-\hat{z})d}\right] \tag{A.7}$$

because the second term in the maximum in the right-hand side exceeds the first term. We refer to this optimization problem as $(\mathcal{P}_s)$. Since (3.2) implies the lower bound in (3.4), the monotonicity constraint for fee levels $(f(z(\bar{\pi})d), f(-z(\bar{\pi})d))$ is satisfied if ϵ is small enough so that the lower bound exceeds $2\epsilon z(\bar{\pi})d$. For $\epsilon = 0$, $(\mathcal{P}_s)$ reduces to maximizing

$$U = -(1-\lambda)\left[\bar{\pi}e^{-\rho[z(\bar{\pi})d-f(z(\bar{\pi})d)]} + (1-\bar{\pi})e^{\rho z(\bar{\pi})d}\right] - \frac{\lambda}{2}\left[e^{-\rho[\hat{z}d-f(z(\bar{\pi})d)]} + e^{\rho\hat{z}d}\right] \tag{A.8}$$

over $(z(\bar{\pi}), \hat{z}, f(z(\bar{\pi})d))$ subject to $0 \leq \hat{z} \leq z(\bar{\pi})$ and

$$-\left[\bar{\pi}e^{-\bar{\rho}[f(z(\bar{\pi})d)-K]} + (1-\bar{\pi})e^{\bar{\rho}K}\right] \geq -\frac{1}{2}\left[e^{-\bar{\rho}f(z(\bar{\pi})d)} + 1\right]. \tag{A.9}$$

Equation (A.9) must hold with equality, otherwise the investor could raise his utility by reducing $f(z(\bar{\pi})d)$. This implies that $f(z(\bar{\pi})d) = f(z(\bar{\pi})d) - f(-z(\bar{\pi})d)$ attains the lower bound in (3.4), and is hence equal to $\Delta(\bar{\pi})$. Maximization of (A.8) over $(z(\bar{\pi}), \hat{z})$ yields the values in the proposition. The optimal values of $(z(\bar{\pi}), \hat{z}, f(z(\bar{\pi})d))$ for small $\epsilon > 0$ are close to their counterparts for $\epsilon = 0$ because of continuity. Since the optimal $f(z(\bar{\pi})d)$ for $\epsilon = 0$ attains the lower bound in (3.4), which is smaller than $\Delta^*(\bar{\pi})$ because of the separating condition, continuity implies that the optimal $f(z(\bar{\pi})d)$ for small $\epsilon > 0$ is also smaller than $\Delta^*(\bar{\pi})$. Hence, for small $\epsilon > 0$, $f(z(\bar{\pi})d)$ and $f(\hat{z}d)$ are smaller than $\Delta^*(\bar{\pi})$. Equation (3.5) is thus satisfied, and the solution to the relaxed problem is also the solution to the original problem. Substituting the $\epsilon = 0$ optimal values of $(z(\bar{\pi}), \hat{z}, f(z(\bar{\pi})d))$ into (A.8), we find that the condition $U \geq -1$ for the investor to employ the manager reduces to (3.14). The optimal fee in the proposition includes the $\epsilon = 0$ optimal values of $(f(z(\bar{\pi})d), f(-z(\bar{\pi})d), f(\hat{z}d), f(-\hat{z}d))$. Since $f(W)$ is equal to zero for $W < \hat{z}d$ and to a positive value for $W \geq \hat{z}d$, both the risk-averse type $\bar{\pi}$ and the risk-neutral type are indifferent between any position $z \geq \hat{z}$ and prefer it to any position $z \in [0, \hat{z})$. Therefore, the positions $z(\bar{\pi})$ and $\hat{z}$ are (weakly) optimal for the risk-averse type $\bar{\pi}$ and the risk-neutral type, respectively.

Suppose finally that the pooling condition $2(1-\bar{\pi}) < e^{-\bar{\rho}K} < 4(1-\bar{\pi})\bar{\pi}$ is satisfied. If $\hat{z} < z(\bar{\pi})$, then we can use

$$f(z(\bar{\pi})d) - f(-z(\bar{\pi})d) \geq f(\hat{z}d) - f(-\hat{z}d) + 2\epsilon(z(\bar{\pi}) - \hat{z})d > f(\hat{z}d) - f(-\hat{z}d)$$

to obtain a contradiction following the argument in Section 3. Hence, $\hat{z} = z(\bar{\pi})$. Using this property and $f(-z(\bar{\pi})d) = 0$, we can write (A.2) as

$$U = -\left[\left((1-\lambda)\bar{\pi} + \frac{\lambda}{2}\right) e^{\rho[z(\bar{\pi})d - f(z(\bar{\pi})d)]} + \left((1-\lambda)(1-\bar{\pi}) + \frac{\lambda}{2}\right) e^{\rho z(\bar{\pi})d}\right]. \tag{A.10}$$

Equation (A.10) must be maximized over $(z(\bar{\pi}), f(z(\bar{\pi})d))$ subject to $z(\bar{\pi}) \geq 0$ and the (IC) constraint (3.2), which takes the form (A.9). We refer to this optimization problem as $(\mathcal{P}_p)$. As in the case of separation, the monotonicity constraint for fee levels $(f(z(\bar{\pi})d), f(-z(\bar{\pi})d))$ is satisfied if ϵ is small enough so that the lower bound in (3.4) exceeds $2\epsilon z(\bar{\pi})d$. Equation (A.9) must hold with equality, otherwise the investor could raise his utility by reducing $f(z(\bar{\pi})d)$. This implies that $f(z(\bar{\pi})d) = f(z(\bar{\pi})d) - f(-z(\bar{\pi})d)$ attains the lower bound in (3.4), and is hence equal to $\Delta(\bar{\pi})$. Maximization of (A.10) over $z(\bar{\pi})$ yields the value in the proposition. Substituting the optimal values of $(z(\bar{\pi}), f(z(\bar{\pi})d))$ into (A.10), we find that the condition $U \geq -1$ for the investor to employ the manager reduces to (3.15). The optimal fee in the proposition includes the optimal values of $(f(z(\bar{\pi})d), f(-z(\bar{\pi})d), f(\hat{z}d), f(-\hat{z}d))$. Since $f(W)$ is equal to zero for $W < z(\bar{\pi})d$ and to a positive value for $W \geq z(\bar{\pi})d$, both the risk-averse type $\bar{\pi}$ and the risk-neutral type are indifferent between any position $z \geq z(\bar{\pi})$ and prefer it to any position $z \in [0, z(\bar{\pi}))$. Therefore, the position $z(\bar{\pi})$ is (weakly) optimal for both types. □

Proof of Lemma 4.1. We first show that the (IC) constraint (4.2) implies that $\Delta(\pi)$ is non-decreasing. The risk-averse type π prefers $(\Delta(\pi), \Gamma(\pi))$ to $(\Delta(\pi'), \Gamma(\pi'))$ if

$$-\left[\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}\right] e^{-\bar{\rho}\Gamma(\pi)} \geq -\left[\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi')} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi')}\right] e^{-\bar{\rho}\Gamma(\pi')}. \tag{A.11}$$

Conversely, the risk-averse type π' prefers $(\Delta(\pi'), \Gamma(\pi'))$ to $(\Delta(\pi), \Gamma(\pi))$ if

$$-\left[\pi' e^{-\frac{\bar{\rho}}{2}\Delta(\pi')} + (1-\pi')e^{\frac{\bar{\rho}}{2}\Delta(\pi')}\right] e^{-\bar{\rho}\Gamma(\pi')} \geq -\left[\pi' e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} + (1-\pi')e^{\frac{\bar{\rho}}{2}\Delta(\pi)}\right] e^{-\bar{\rho}\Gamma(\pi)}. \tag{A.12}$$

Multiplying each side of (A.11) by the corresponding side of (A.12) (after multiplying both equations by minus one, to make their sides positive), we find

$$\begin{aligned}
&\left[\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}\right]\left[\pi' e^{-\frac{\bar{\rho}}{2}\Delta(\pi')} + (1-\pi')e^{\frac{\bar{\rho}}{2}\Delta(\pi')}\right] \\
&\leq \left[\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi')} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi')}\right]\left[\pi' e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} + (1-\pi')e^{\frac{\bar{\rho}}{2}\Delta(\pi)}\right] \\
&\Leftrightarrow (\pi - \pi')\left[e^{\frac{\bar{\rho}}{2}(\Delta(\pi)-\Delta(\pi'))} - e^{-\frac{\bar{\rho}}{2}(\Delta(\pi)-\Delta(\pi'))}\right] \geq 0.
\end{aligned} \tag{A.13}$$

Equation (A.13) implies that if $\pi > \pi'$ then $\Delta(\pi) \geq \Delta(\pi')$. Hence, $\Delta(\pi)$ is non-decreasing.

We next show that the (IC) constraint (4.2) implies that $U(\pi)$ has the differentiability properties stated in the lemma. Consider first a point π at which $\Delta(\pi)$ is continuous. Equations (A.11) and

(A.12) imply

$$\frac{\left[\pi' e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} + (1-\pi')e^{\frac{\bar{\rho}}{2}\Delta(\pi)}\right] e^{-\bar{\rho}\Gamma(\pi)}}{\pi' e^{-\frac{\bar{\rho}}{2}\Delta(\pi')} + (1-\pi')e^{\frac{\bar{\rho}}{2}\Delta(\pi')}} \geq e^{-\bar{\rho}\Gamma(\pi')} \geq \frac{\left[\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}\right] e^{-\bar{\rho}\Gamma(\pi)}}{\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi')} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi')}}. \tag{A.14}$$

Since $\Delta(\pi)$ is continuous at π, both fractions in (A.14) converge to $e^{-\bar{\rho}\Gamma(\pi)}$ when π' goes to π. Equation (A.14) then implies that $e^{-\bar{\rho}\Gamma(\pi')}$ converges to the same limit. Hence, $\Gamma(\pi)$ is continuous at π. Using the definition of $U(\pi)$ from (4.6), we find

$$\frac{U(\pi)-U(\pi')}{\pi-\pi'} = \frac{\left[\pi' e^{-\frac{\bar{\rho}}{2}\Delta(\pi')} + (1-\pi')e^{\frac{\bar{\rho}}{2}\Delta(\pi')}\right] e^{-\bar{\rho}\Gamma(\pi')} - \left[\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}\right] e^{-\bar{\rho}\Gamma(\pi)}}{\pi-\pi'}. \tag{A.15}$$

Combining (A.15) with (A.11), we find

$$\begin{aligned}\frac{U(\pi)-U(\pi')}{\pi-\pi'} &\geq \frac{\left[\pi' e^{-\frac{\bar{\rho}}{2}\Delta(\pi')} + (1-\pi')e^{\frac{\bar{\rho}}{2}\Delta(\pi')}\right] e^{-\bar{\rho}\Gamma(\pi')} - \left[\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi')} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi')}\right] e^{-\bar{\rho}\Gamma(\pi')}}{\pi-\pi'} \\ &= \left[e^{\frac{\bar{\rho}}{2}\Delta(\pi')} - e^{-\frac{\bar{\rho}}{2}\Delta(\pi')}\right] e^{-\bar{\rho}\Gamma(\pi')}.\end{aligned} \tag{A.16}$$

Combining (A.15) with (A.12), we find

$$\begin{aligned}\frac{U(\pi)-U(\pi')}{\pi-\pi'} &\leq \frac{\left[\pi' e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} + (1-\pi')e^{\frac{\bar{\rho}}{2}\Delta(\pi)}\right] e^{-\bar{\rho}\Gamma(\pi)} - \left[\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}\right] e^{-\bar{\rho}\Gamma(\pi)}}{\pi-\pi'} \\ &= \left[e^{\frac{\bar{\rho}}{2}\Delta(\pi)} - e^{-\frac{\bar{\rho}}{2}\Delta(\pi)}\right] e^{-\bar{\rho}\Gamma(\pi)}.\end{aligned} \tag{A.17}$$

Since $(\Delta(\pi), \Gamma(\pi))$ are continuous at π, the right-hand side of (A.16) converges to

$$\left[e^{\frac{\bar{\rho}}{2}\Delta(\pi)} - e^{-\frac{\bar{\rho}}{2}\Delta(\pi)}\right] e^{-\bar{\rho}\Gamma(\pi)}$$

when π' goes to π. Equations (A.16) and (A.17) then imply that $\frac{U(\pi)-U(\pi')}{\pi-\pi'}$ converges to the same limit. Hence, $U(\pi)$ is differentiable at π, with $U'(\pi)$ given by (4.7).

Consider next a point π at which $U(\pi)$ is discontinuous. Since $\Delta(\pi)$ is non-decreasing, $\Delta(\pi)$ has left- and right-limits at π, which we denote by $\Delta(\pi^-)$ and $\Delta(\pi^+)$, respectively. Equation (A.14) written for $\pi' < \pi$ implies that $\Gamma(\pi)$ has a left-limit $\Gamma(\pi^-)$ at π, given by

$$e^{-\bar{\rho}\Gamma(\pi^-)} = \frac{\left[\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}\right] e^{-\bar{\rho}\Gamma(\pi)}}{\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi^-)} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi^-)}}. \tag{A.18}$$

Consider next $\pi' < \pi'' < \pi$ and the (IC) constraint that the risk-averse type π' prefers $(\Delta(\pi'), \Gamma(\pi'))$

to $(\Delta(\pi''), \Gamma(\pi''))$. Taking the limit of that equation when π'' goes to π, we find

$$-\left[\pi' e^{-\frac{\bar{\rho}}{2}\Delta(\pi')} + (1-\pi')e^{\frac{\bar{\rho}}{2}\Delta(\pi')}\right] e^{-\bar{\rho}\Gamma(\pi')} \geq -\left[\pi' e^{-\frac{\bar{\rho}}{2}\Delta(\pi^-)} + (1-\pi')e^{\frac{\bar{\rho}}{2}\Delta(\pi^-)}\right] e^{-\bar{\rho}\Gamma(\pi^-)}. \quad \text{(A.19)}$$

Combining (A.15) with (A.18) and (A.19), we obtain the following counterpart of (A.17):

$$\begin{aligned}\frac{U(\pi) - U(\pi')}{\pi - \pi'} &\leq \frac{\left[\pi' e^{-\frac{\bar{\rho}}{2}\Delta(\pi^-)} + (1-\pi')e^{\frac{\bar{\rho}}{2}\Delta(\pi^-)}\right] e^{-\bar{\rho}\Gamma(\pi^-)} - \left[\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi^-)} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi^-)}\right] e^{-\bar{\rho}\Gamma(\pi^-)}}{\pi - \pi'} \\ &= \left[e^{\frac{\bar{\rho}}{2}\Delta(\pi^-)} - e^{-\frac{\bar{\rho}}{2}\Delta(\pi^-)}\right] e^{-\bar{\rho}\Gamma(\pi^-)}. \quad \text{(A.20)}\end{aligned}$$

Since $(\Delta(\pi), \Gamma(\pi))$ have left-limits at π, the right-hand side of (A.16) converges to

$$\left[e^{\frac{\bar{\rho}}{2}\Delta(\pi^-)} - e^{-\frac{\bar{\rho}}{2}\Delta(\pi^-)}\right] e^{-\bar{\rho}\Gamma(\pi^-)}$$

when π' goes to π from the left. Equations (A.16) and (A.20) then imply that $\frac{U(\pi)-U(\pi')}{\pi-\pi'}$ converges to the same limit. Hence, $U(\pi)$ has a left-derivative at π, with $U'(\pi)$ given by that limit. The argument for the right-derivative is identical.

We finally show that properties (i) and (ii) in the lemma imply the (IC) constraint (4.2), i.e., (A.11) holds for all (π, π'). For this result and subsequent proofs we use (4.9), which follows by integrating the ordinary differential equation (ODE) (4.8). The proof that (4.8) integrates to (4.9) must account for possible points of discontinuity of $\Delta(\pi)$, which are at most countable because $\Delta(\pi)$ is monotone. The function

$$H(\Delta, \pi) \equiv \frac{e^{\frac{\bar{\rho}}{2}\Delta} - e^{-\frac{\bar{\rho}}{2}\Delta}}{\pi e^{-\frac{\bar{\rho}}{2}\Delta} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta}}$$

is increasing in Δ. It is also increasing in π for $\Delta > 0$ and independent of π for $\Delta = 0$. Since $\Delta(\pi)$ is non-negative and non-decreasing for $\pi \in [\frac{1}{2}, \bar{\pi}]$, the function $\pi \to H(\Delta(\pi), \pi)$ is non-decreasing. Hence, it is measurable, and the integral $\int_{\pi}^{\bar{\pi}} H(\Delta(\pi'), \pi')d\pi'$ in (4.9) is well-defined. Since $U(\pi)$ and $\int_{\pi}^{\bar{\pi}} H(\Delta(\pi'), \pi')d\pi'$ have left- and right-derivatives, the function $K(\pi) \equiv U(\pi) \exp\left[-\int_{\pi}^{\bar{\pi}} H(\Delta(\pi'), \pi')d\pi'\right]$ has also left- and right-derivatives. Moreover, property (ii) in the lemma implies that the left- and right-derivatives of $K(\pi)$ are zero for all π. Hence, $K(\pi) = U(\bar{\pi})$, which implies (4.9).

Combining (4.6) and (4.9), we find

$$e^{-\bar{\rho}\Gamma(\pi)} = -\frac{U(\bar{\pi}) \exp\left[\int_{\pi}^{\bar{\pi}} H(\Delta(\pi'), \pi')d\pi'\right]}{\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}}. \quad \text{(A.21)}$$

Substituting $e^{-\bar{\rho}\Gamma(\pi)}$ and $e^{-\bar{\rho}\Gamma(\pi')}$ from (A.21) into (A.11), we find that (A.11) is equivalent to

$$\begin{aligned} & U(\bar{\pi}) \exp\left[\int_{\pi}^{\bar{\pi}} H(\Delta(\pi''),\pi'')d\pi''\right] \\ & \geq U(\bar{\pi}) \exp\left[\int_{\pi'}^{\bar{\pi}} H(\Delta(\pi''),\pi'')d\pi''\right] \frac{\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi')} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi')}}{\pi' e^{-\frac{\bar{\rho}}{2}\Delta(\pi')} + (1-\pi')e^{\frac{\bar{\rho}}{2}\Delta(\pi')}} \\ & \Leftrightarrow 1 \leq \exp\left[-\int_{\pi}^{\pi'} H(\Delta(\pi''),\pi'')d\pi''\right] \frac{\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi')} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi')}}{\pi' e^{-\frac{\bar{\rho}}{2}\Delta(\pi')} + (1-\pi')e^{\frac{\bar{\rho}}{2}\Delta(\pi')}}, \end{aligned} \tag{A.22}$$

where the second step follows by dividing both sides by

$$U(\bar{\pi}) \exp\left[\int_{\pi}^{\bar{\pi}} H(\Delta(\pi''),\pi'')d\pi''\right],$$

which is negative. Since $H(\Delta,\pi)$ is increasing in Δ, and $\Delta(\pi)$ is non-decreasing,

$$\begin{aligned} \exp\left[-\int_{\pi}^{\pi'} H(\Delta(\pi''),\pi'')d\pi''\right] & \geq \exp\left[-\int_{\pi}^{\pi'} H(\Delta(\pi'),\pi'')d\pi''\right] \\ & = \exp\left[\left[\log\left(\pi'' e^{-\frac{\bar{\rho}}{2}\Delta(\pi')} + (1-\pi'')e^{\frac{\bar{\rho}}{2}\Delta(\pi')}\right)\right]_{\pi}^{\pi'}\right] \\ & = \frac{\pi' e^{-\frac{\bar{\rho}}{2}\Delta(\pi')} + (1-\pi')e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi}')}}{\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi')} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi')}}. \end{aligned} \tag{A.23}$$

Equation (A.23) implies that (A.22) holds for all (π,π'). □

Proof of Lemma 4.2. Substituting $U(\pi)$ from (4.6) into (4.1), after writing the latter equation as

$$2e^{\bar{\rho}K} \int_{\frac{1}{2}}^{\bar{\pi}} U(\pi)h(\pi)d\pi \geq U\left(\frac{1}{2}\right), \tag{A.24}$$

we find

$$2e^{\bar{\rho}K} \int_{\frac{1}{2}}^{\bar{\pi}} U(\bar{\pi}) \exp\left[\int_{\pi}^{\bar{\pi}} H(\Delta(\pi'),\pi')d\pi'\right] h(\pi)d\pi \geq U(\bar{\pi}) \exp\left[\int_{\frac{1}{2}}^{\bar{\pi}} H(\Delta(\pi'),\pi')d\pi'\right].$$

Dividing both sides by

$$U(\bar{\pi}) \exp\left[\int_{\frac{1}{2}}^{\bar{\pi}} H(\Delta(\pi'),\pi')d\pi'\right],$$

which is negative, we find (4.10).

To derive the lower bound on $\Delta(\bar{\pi})$, we note that since $H(\Delta,\pi)$ is increasing in Δ, and $\Delta(\pi)$ is non-decreasing,

$$\begin{aligned}\exp\left[-\int_{\frac{1}{2}}^{\pi} H(\Delta(\pi'),\pi')d\pi'\right] &\geq \exp\left[-\int_{\frac{1}{2}}^{\pi} H(\Delta(\bar{\pi}),\pi')d\pi'\right] \\ &= \exp\left[\left[\log\left(\pi' e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})} + (1-\pi')e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\right)\right]_{\frac{1}{2}}^{\pi}\right] \\ &= \frac{\pi e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}}{\frac{1}{2}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})} + \frac{1}{2}e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}},\end{aligned} \tag{A.25}$$

and hence

$$\int_{\frac{1}{2}}^{\bar{\pi}} \exp\left[-\int_{\frac{1}{2}}^{\pi} H(\Delta(\pi'),\pi')d\pi'\right] h(\pi)d\pi \geq \frac{e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\int_{\frac{1}{2}}^{\bar{\pi}} \pi h(\pi)d\pi + e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\int_{\frac{1}{2}}^{\bar{\pi}}(1-\pi)h(\pi)d\pi}{\frac{1}{2}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})} + \frac{1}{2}e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}}. \tag{A.26}$$

Equations (4.10) and (A.26) imply

$$\begin{aligned}&2e^{\bar{\rho}K}\frac{e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\int_{\frac{1}{2}}^{\bar{\pi}} \pi h(\pi)d\pi + e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\int_{\frac{1}{2}}^{\bar{\pi}}(1-\pi)h(\pi)d\pi}{\frac{1}{2}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})} + \frac{1}{2}e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}} \leq 1 \\ &\Leftrightarrow 4\int_{\frac{1}{2}}^{\bar{\pi}} \pi h(\pi)d\pi - e^{-\bar{\rho}K} \leq e^{\bar{\rho}\Delta(\bar{\pi})}\left[e^{-\bar{\rho}K} - 2\left(1 - 2\int_{\frac{1}{2}}^{\bar{\pi}} \pi h(\pi)d\pi\right)\right],\end{aligned} \tag{A.27}$$

where the last step follows by rearranging and noting that symmetry implies $\int_{\frac{1}{2}}^{\bar{\pi}} h(\pi)d\pi = \frac{1}{2}$. Since $\int_{\frac{1}{2}}^{\bar{\pi}} \pi h(\pi)d\pi > \frac{1}{4}$ (because π takes values in $(\frac{1}{2},\bar{\pi}]$ with positive probability), (A.27) implies (4.11). Dividing both sides of (A.27) by $e^{-\bar{\rho}K} - 2\left(1 - 2\int_{\frac{1}{2}}^{\bar{\pi}} \pi h(\pi)d\pi\right) > 0$, we find (4.12). □

Proof of Lemma 4.3. We first show that the left inequality in (4.13) is equivalent to the lower bound in (4.12) exceeding $\Delta^*(\bar{\pi})$. Equation (3.10) and Lemma 4.2 imply that the latter condition is equivalent to

$$\frac{4\int_{\frac{1}{2}}^{\bar{\pi}} \pi h(\pi)d\pi - e^{-\bar{\rho}K}}{e^{-\bar{\rho}K} - 2\left(1 - 2\int_{\frac{1}{2}}^{\bar{\pi}} \pi h(\pi)d\pi\right)} > \frac{\bar{\pi}}{1-\bar{\pi}}. \tag{A.28}$$

Rearranging (A.28), we find the left inequality in (4.13). Equation (4.13) implies that $\Delta(\bar{\pi})$, which exceeds the lower bound in (4.12), also exceeds $\Delta^*(\bar{\pi})$.

We next show that $\hat{z} \geq z(\bar{\pi})$. Suppose, by contradiction, that $\hat{z} < z(\bar{\pi})$. Using $(\Delta(\pi), \Gamma(\pi), \hat{\Delta}, \hat{\Gamma})$, we can write (4.3) as

$$-\left[\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}\right]e^{-\rho\Gamma(\pi)} \geq -\left[\pi e^{-\frac{\bar{\rho}}{2}\hat{\Delta}} + (1-\pi)e^{\frac{\bar{\rho}}{2}\hat{\Delta}}\right]e^{-\rho\hat{\Gamma}}. \tag{A.29}$$

Generalizing the notation introduced in the case of two signal values, we set

$$
\begin{aligned}
G_\pi(\Delta) &\equiv -\left(\pi e^{-\frac{\bar\rho}{2}\Delta} + (1-\pi)e^{\frac{\bar\rho}{2}\Delta}\right), \\
e^{\bar\rho\Delta^*(\pi)} &\equiv \frac{\pi}{1-\pi}.
\end{aligned}
$$

Using $G_\pi(\Delta)$, we can write (A.29) as

$$
G_\pi(\Delta(\pi))e^{-\rho\Gamma(\pi)} \ge G_\pi(\hat\Delta)e^{-\rho\hat\Gamma}. \tag{A.30}
$$

Since $G_{\bar\pi}(\Delta)$ is decreasing for Δ between $\Delta^*(\bar\pi)$ and the lower bound in (4.12), and since (4.5) implies $\hat\Gamma \ge \Gamma(\bar\pi)$, (A.30) written for $\pi = \bar\pi$ implies $\hat\Delta < \Delta^*(\bar\pi)$. Consider $\hat\pi \in [\frac{1}{2}, \bar\pi)$ such that $\Delta^*(\hat\pi) = \hat\Delta$. Since $\Delta^*(\hat\pi) = \hat\Delta$ maximizes $G_{\hat\pi}(\Delta)$, and since (4.5) implies $\hat\Gamma \ge \Gamma(\hat\pi)$, (A.30) written for $\pi = \hat\pi$ implies $\Delta(\hat\pi) = \Delta^*(\hat\pi) = \hat\Delta$ and $\Gamma(\hat\pi) = \hat\Gamma$. Hence, (4.5) implies $\Gamma(\hat\pi) = \max_{\pi\in[\frac{1}{2},\bar\pi]}\Gamma(\pi)$. Combining $\Gamma(\hat\pi) = \max_{\pi\in[\frac{1}{2},\bar\pi]}\Gamma(\pi)$ and (A.21), and noting that $U(\bar\pi) < 0$, we find

$$
\begin{aligned}
&\frac{\exp\left[\int_\pi^{\bar\pi} H(\Delta(\pi'),\pi')d\pi'\right]}{\pi e^{-\frac{\bar\rho}{2}\Delta(\pi)} + (1-\pi)e^{\frac{\bar\rho}{2}\Delta(\pi)}} \ge \frac{\exp\left[\int_{\hat\pi}^{\bar\pi} H(\Delta(\pi'),\pi')d\pi'\right]}{\hat\pi e^{-\frac{\bar\rho}{2}\Delta(\hat\pi)} + (1-\hat\pi)e^{\frac{\bar\rho}{2}\Delta(\hat\pi)}} \\
&\Leftrightarrow F(\pi) \equiv \exp\left[-\int_{\hat\pi}^{\pi} H(\Delta(\pi'),\pi')d\pi'\right] - \frac{\pi e^{-\frac{\bar\rho}{2}\Delta(\pi)} + (1-\pi)e^{\frac{\bar\rho}{2}\Delta(\pi)}}{\hat\pi e^{-\frac{\bar\rho}{2}\Delta(\hat\pi)} + (1-\hat\pi)e^{\frac{\bar\rho}{2}\Delta(\hat\pi)}} \ge 0.
\end{aligned} \tag{A.31}
$$

To derive the contradiction, we will show that (4.10) is violated for any $\Delta(\pi)$ that is defined over $[\frac{1}{2}, \bar\pi]$, is non-decreasing, and satisfies $F(\pi) \ge 0$ and $\Delta(\hat\pi) = \Delta^*(\hat\pi)$.

Consider the problem of minimizing

$$
\int_{\hat\pi}^{\bar\pi} \exp\left[-\int_{\hat\pi}^{\pi} H(\Delta(\pi'),\pi')d\pi'\right] h(\pi)d\pi \tag{A.32}
$$

over $\Delta(\pi)$ that is defined over $[\hat\pi, \bar\pi]$, is left-continuous with right-limits, and satisfies $F(\pi) \ge 0$ and $\Delta(\hat\pi) = \Delta^*(\hat\pi)$. The Lagrangian for this problem is

$$
\begin{aligned}
&\int_{\hat\pi}^{\bar\pi} \exp\left[-\int_{\hat\pi}^{\pi} H(\Delta(\pi'),\pi')d\pi'\right] h(\pi)d\pi - \int_{\hat\pi}^{\bar\pi} F(\pi)\mu(\pi)d\pi \\
&= \int_{\hat\pi}^{\bar\pi} \exp\left[-\int_{\hat\pi}^{\pi} H(\Delta(\pi'),\pi')d\pi'\right] [h(\pi) - \mu(\pi)]d\pi + \int_{\hat\pi}^{\bar\pi} \frac{\pi e^{-\frac{\bar\rho}{2}\Delta(\pi)} + (1-\pi)e^{\frac{\bar\rho}{2}\Delta(\pi)}}{\hat\pi e^{-\frac{\bar\rho}{2}\Delta(\hat\pi)} + (1-\hat\pi)e^{\frac{\bar\rho}{2}\Delta(\hat\pi)}}\mu(\pi)d\pi,
\end{aligned}
$$

and yields the first-order condition

$$
\begin{aligned}
&- H_\Delta(\Delta(\pi),\pi)\int_\pi^{\bar\pi}\exp\left[-\int_{\hat\pi}^{\pi'} H(\Delta(\pi''),\pi'')d\pi''\right][h(\pi')-\mu(\pi')]d\pi' \\
&+\frac{\bar\rho\mu(\pi)}{2}\frac{(1-\pi)e^{\frac{\bar\rho}{2}\Delta(\pi)}-\pi e^{-\frac{\bar\rho}{2}\Delta(\pi)}}{\hat\pi e^{-\frac{\bar\rho}{2}\Delta(\hat\pi)}+(1-\hat\pi)e^{\frac{\bar\rho}{2}\Delta(\hat\pi)}}=0.
\end{aligned}
\tag{A.33}
$$

If $F(\pi)>0$ for π in an open interval (π_1,π_2), then for π in that interval $\mu(\pi)=0$ and (A.33) becomes

$$
\int_\pi^{\bar\pi}\exp\left[-\int_{\hat\pi}^{\pi'} H(\Delta(\pi''),\pi'')d\pi''\right][h(\pi')-\mu(\pi')]d\pi'=0. \tag{A.34}
$$

Differentiating (A.35) in (π_1,π_2), and using $\mu(\pi)=0$, $h(\pi)>0$, (A.31) and $F(\pi)>0$, we find

$$
\exp\left[-\int_{\hat\pi}^{\pi} H(\Delta(\pi'),\pi')d\pi'\right]=0>\frac{\pi e^{-\frac{\bar\rho}{2}\Delta(\pi)}+(1-\pi)e^{\frac{\bar\rho}{2}\Delta(\pi)}}{\hat\pi e^{-\frac{\bar\rho}{2}\Delta(\hat\pi)}+(1-\hat\pi)e^{\frac{\bar\rho}{2}\Delta(\hat\pi)}},
$$

a contradiction. Hence, the function $\Delta(\pi)$ that minimizes (A.32) satisfies $F(\pi)=0$ in a set that is dense in $[\hat\pi,\bar\pi]$. Since $\Delta(\pi)$ is left-continuous and $F(\hat\pi)=0$, this set coincides with $[\hat\pi,\bar\pi]$. Suppose next, by contradiction, that $\Delta(\pi)$ is discontinuous at a point π_1. Since $\Delta(\pi)$ is left-continuous, $\pi_1\neq\bar\pi$. Since, in addition, $F(\pi_1)=F(\pi_1^+)=0$ implies $G_\pi(\Delta(\pi_1))=G_\pi(\Delta(\pi_1^+))$, and since $\Delta(\hat\pi)=\Delta^*(\hat\pi)$, $\pi_1\neq\hat\pi$. Hence, $\pi_1\in(\hat\pi,\bar\pi)$. Combining (A.33) for π_1 and for $\pi_1'>\pi_1$, we find

$$
\begin{aligned}
&\int_{\pi_1}^{\pi_1'}\exp\left[-\int_{\hat\pi}^{\pi'} H(\Delta(\pi''),\pi'')d\pi''\right][h(\pi')-\mu(\pi')]d\pi' \\
&+\frac{\mu(\pi_1)G'_{\pi_1}(\Delta(\pi_1))}{H_\Delta(\Delta(\pi_1),\pi_1)G_{\hat\pi}(\Delta(\hat\pi))}-\frac{\mu(\pi_1')G'_{\pi_1'}(\Delta(\pi_1'))}{H_\Delta(\Delta(\pi_1'),\pi_1')G_{\hat\pi}(\Delta(\hat\pi))}=0.
\end{aligned}
\tag{A.35}
$$

If $\Delta(\pi_1)<\Delta(\pi_1^+)$, in which case $\Delta(\pi_1)<\Delta^*(\pi_1)<\Delta(\pi_1^+)$, $G'_{\pi_1}(\Delta(\pi_1))>0$ and $G'_{\pi_1}(\Delta(\pi_1^+))<0$, then the second term in (A.35) is non-negative because $\mu(\pi_1)\geq 0$, and the third term is non-negative because $\mu(\pi_1')\geq 0$. If $\mu(\pi_1^+)>0$, then the third term is positive in the limit when π_1' goes to π_1. Since the first term goes to zero in that limit, (A.35) is violated, a contradiction. If instead $\mu(\pi_1^+)=0$, then the first term is positive close to the limit because $h(\pi_1)>0$, and hence (A.35) is again violated, a contradiction. Therefore, $\Delta(\pi_1)>\Delta^*(\pi_1)>\Delta(\pi_1^+)$. If $\Delta(\pi)$ is discontinuous at an additional point π_2', then the same reasoning implies $\Delta(\pi_2')>\Delta^*(\pi_2')>\Delta(\pi_2'^+)$. Assume without loss of generality that $\pi_2'>\pi_1$, and take the infimum π_2 of the discontinuity points π_2' that exceed π_1. The infimum π_2 must strictly exceed π_1 because otherwise taking the limit in $\Delta(\pi_2')>\Delta^*(\pi_2')$ yields $\Delta(\pi_1^+)\geq\Delta^*(\pi_1)$. Hence, $\Delta(\pi)$ is continuous in the non-empty interval

(π_1,π_2). Using (A.21) to differentiate $F(\pi)=0$, we find

$$\begin{aligned}
&H(\Delta(\pi),\pi)\exp\left[-\int_{\hat{\pi}}^{\pi}H(\Delta(\pi'),\pi')d\pi'\right]+\frac{e^{\frac{\bar{\rho}}{2}\Delta(\pi)}-e^{-\frac{\bar{\rho}}{2}\Delta(\pi)}}{\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\hat{\pi})}+(1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\hat{\pi})}}\\
&-\frac{\bar{\rho}\Delta'(\pi)}{2}\frac{(1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}-\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)}}{\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\hat{\pi})}+(1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\hat{\pi})}}=0\\
&\Leftrightarrow\frac{\bar{\rho}\Delta'(\pi)}{2}\frac{(1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}-\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)}}{\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\hat{\pi})}+(1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\hat{\pi})}}=0,
\end{aligned}\tag{A.36}$$

where the second step follows by using $F(\pi)=0$ and the definition of $H(\Delta,\pi)$. Equation (A.36) with the initial condition $\Delta(\pi_1^+)<\Delta^*(\pi_1)$ implies $\Delta(\pi)=0$ for $\pi\in(\pi_1,\pi_2)$, and hence $\Delta(\pi_1^+)=\Delta(\pi_2)$. This is a contradiction. Indeed, if π_2 is a discontinuity point, then $\Delta(\pi_2)>\Delta^*(\pi_2)>\Delta^*(\pi_1)>\Delta(\pi_1^+)$. If instead, π_2 is a continuity point and hence a limit of discontinuity points, then taking the limit in $\Delta(\pi_2')>\Delta^*(\pi_2')$ yields $\Delta(\pi_2^+)\geq\Delta^*(\pi_2)$, which implies $\Delta(\pi_2)=\Delta(\pi_2^+)\geq\Delta^*(\pi_2)>\Delta^*(\pi_1)>\Delta(\pi_1^+)$. Therefore, $\Delta(\pi)$ can have at most one discontinuity point π_1. Equation (A.36) with the initial condition $\Delta(\hat{\pi})=\Delta^*(\hat{\pi})$ implies, however, that $\Delta(\pi)$ is of the form $\Delta(\pi)=\Delta^*(\pi)$ for $\pi\in[\hat{\pi},\pi_3]$ and $\Delta(\pi)=\Delta(\pi_3)$ for $\pi\in(\pi_3,\pi_1]$, where $\pi_3\in[\hat{\pi},\pi_1]$. This implies $\Delta(\pi_1)\leq\Delta^*(\pi_1)$, a contradiction. Therefore, $\Delta(\pi)$ is continuous over the entire interval $[\hat{\pi},\bar{\pi}]$. Equation (A.36) with the initial condition $\Delta(\hat{\pi})=\Delta^*(\hat{\pi})$ implies that $\Delta(\pi)$ is of the form $\Delta(\pi)=\Delta^*(\pi)$ for $\pi\in[\hat{\pi},\pi_3]$ and $\Delta(\pi)=\Delta(\pi_3)$ for $\pi\in(\pi_3,\bar{\pi}]$, where $\pi_3\in[\hat{\pi},\bar{\pi}]$. The solution that minimizes (A.32) corresponds to $\pi_3=\bar{\pi}$. That solution, which is non-decreasing, also minimizes (A.32) over the set of non-decreasing functions that satisfy $F(\pi)\geq 0$ and $\Delta(\hat{\pi})=\Delta^*(\hat{\pi})$. Indeed, a non-decreasing function has a countable set of discontinuity points, and left- and right-limits at those points. By setting its value at the discontinuity points to its left-limit, we can transform it into a left-continuous function with right-limits. Since this operation is performed at a countable set of points, the resulting function yields an identical value for (A.32).

Using our solution of the minimization problem, we can show that (4.10) is violated for any $\Delta(\pi)$ that is defined over $[\frac{1}{2},\bar{\pi}]$, is non-decreasing, and satisfies $F(\pi)\geq 0$ and $\Delta(\hat{\pi})=\Delta^*(\hat{\pi})$. We write (4.10) as

$$\begin{aligned}
&2e^{\bar{\rho}K}\left\{\int_{\frac{1}{2}}^{\hat{\pi}}\exp\left[-\int_{\frac{1}{2}}^{\pi}H(\Delta(\pi'),\pi')d\pi'\right]h(\pi)d\pi\right.\\
&\left.+\exp\left[-\int_{\frac{1}{2}}^{\hat{\pi}}H(\Delta(\pi'),\pi')d\pi'\right]\int_{\hat{\pi}}^{\bar{\pi}}\exp\left[-\int_{\hat{\pi}}^{\pi}H(\Delta(\pi'),\pi')d\pi'\right]h(\pi)d\pi\right\}\leq 1.
\end{aligned}\tag{A.37}$$

Since (A.32) is minimized for $\Delta(\pi) = \Delta^*(\pi)$,

$$\begin{aligned}\int_{\hat{\pi}}^{\bar{\pi}} \exp\left[-\int_{\hat{\pi}}^{\pi} H(\Delta(\pi'), \pi')d\pi'\right] h(\pi)d\pi \geq & \int_{\hat{\pi}}^{\bar{\pi}} \exp\left[-\int_{\hat{\pi}}^{\pi} H(\Delta^*(\pi'), \pi')d\pi'\right] h(\pi)d\pi \\ \geq & \int_{\hat{\pi}}^{\bar{\pi}} \exp\left[-\int_{\hat{\pi}}^{\pi} H(\Delta^*(\bar{\pi}), \pi')d\pi'\right] h(\pi)d\pi, \end{aligned} \tag{A.38}$$

where the second step follows because $\Delta^*(\pi)$ is increasing and $H(\Delta, \pi)$ is increasing in Δ. Since $\Delta(\pi)$ is non-decreasing, $\Delta(\hat{\pi}) = \Delta^*(\hat{\pi})$, and $H(\Delta, \pi)$ is increasing in Δ,

$$\exp\left[-\int_{\frac{1}{2}}^{\pi} H(\Delta(\pi'), \pi')d\pi'\right] \geq \exp\left[-\int_{\frac{1}{2}}^{\pi} H(\Delta^*(\hat{\pi}), \pi')d\pi'\right] \geq \exp\left[-\int_{\frac{1}{2}}^{\pi} H(\Delta^*(\bar{\pi}), \pi')d\pi'\right] \tag{A.39}$$

for $\pi \in [\frac{1}{2}, \hat{\pi}]$, where the second step follows because $\Delta^*(\pi)$ is increasing and $H(\Delta, \pi)$ is increasing in Δ. Substituting (A.38) and (A.39) into (A.37), we find

$$\begin{aligned} & 2e^{\bar{\rho}K} \int_{\frac{1}{2}}^{\bar{\pi}} \exp\left[-\int_{\frac{1}{2}}^{\pi} H(\Delta^*(\bar{\pi}), \pi')d\pi'\right] h(\pi)d\pi \leq 1 \\ \Leftrightarrow\ & 2e^{\bar{\rho}K} \int_{\frac{1}{2}}^{\bar{\pi}} \frac{\pi e^{-\frac{\bar{\rho}}{2}\Delta^*(\bar{\pi})} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta^*(\bar{\pi})}}{\frac{1}{2}e^{-\frac{\bar{\rho}}{2}\Delta^*(\bar{\pi})} + \frac{1}{2}e^{\frac{\bar{\rho}}{2}\Delta^*(\bar{\pi})}} h(\pi)d\pi \leq 1 \\ \Leftrightarrow\ & 2e^{\bar{\rho}K} \frac{e^{-\frac{\bar{\rho}}{2}\Delta^*(\bar{\pi})} \int_{\frac{1}{2}}^{\bar{\pi}} \pi h(\pi)d\pi + e^{\frac{\bar{\rho}}{2}\Delta^*(\bar{\pi})} \int_{\frac{1}{2}}^{\bar{\pi}} (1-\pi)h(\pi)d\pi}{\frac{1}{2}e^{-\frac{\bar{\rho}}{2}\Delta^*(\bar{\pi})} + \frac{1}{2}e^{\frac{\bar{\rho}}{2}\Delta^*(\bar{\pi})}} \leq 1, \end{aligned} \tag{A.40}$$

where the second step follows from (A.25). Solving (A.27) for $e^{\bar{\rho}\Delta^*(\bar{\pi})}$, we find that $\Delta^*(\bar{\pi})$ exceeds the lower bound in (4.12). This contradicts (4.13). □

Proof of Proposition 4.1. If the investor invests directly, then he sets $z = 0$ and has utility -1. If the investor prefers to employ the manager, then he chooses positions $(z(\pi), \hat{z})$ and fee levels $(f(z(\pi)d), f(-z(\pi)d), f(\hat{z}d), f(-\hat{z}d))$ for $\pi \in [\frac{1}{2}, \bar{\pi}]$ to maximize the utility

$$\begin{aligned} U = & -2(1-\lambda) \int_{\frac{1}{2}}^{\bar{\pi}} \left[\pi e^{-\rho[z(\pi)d - f(z(\pi)d)]} + (1-\pi)e^{-\rho[-z(\pi)d - f(-z(\pi)d)]}\right] h(\pi)d\pi \\ & - \frac{\lambda}{2}\left[e^{-\rho[\hat{z}d - f(\hat{z}d)]} + e^{-\rho[-\hat{z}d - f(-\hat{z}d)]}\right]. \end{aligned} \tag{A.41}$$

The investor is subject to $z(\pi) \geq 0$, $\hat{z} \geq 0$, the (IC) constraints (4.1), (4.2), (4.3) and (4.4), and the non-negativity and monotonicity of the fee. We refer to this optimization problem as $(\mathcal{P})$.

When the pooling condition (4.13) holds, the problem $(\mathcal{P})$ is equivalent to maximizing (A.41) subject to the following constraints: (i) (4.10), (ii) $\Delta(\pi)$ is non-decreasing, (iii) $U(\pi)$ is given by

(4.9), (iv) $z(\pi)$ and $\hat{z}$ are non-negative, (v) $(\hat{z}, \hat{\Delta}, \hat{\Gamma}) = (z(\bar{\pi}), \Delta(\bar{\pi}), \Gamma(\bar{\pi}))$, (vi) $f(-z(\bar{\pi})d) = 0$, and (vii) fee monotonicity. Indeed, Lemma 4.1 shows that the (IC) constraint (4.2) is equivalent to (ii) and (iii), and Lemma 4.2 shows that the (IC) constraint (4.1) is equivalent to (i). The argument in the second paragraph after the statement of Lemma 4.3 shows that (ii) and (vii) imply that $z(\pi)$ is non-decreasing. A similar argument as in the proof of Proposition 3.1 shows that $\hat{z}$ cannot exceed $z(\bar{\pi}) = \max_{\pi \in [\frac{1}{2}, \bar{\pi}]} z(\pi)$. Combining this result with Lemma 4.3, we find $\hat{z} = z(\bar{\pi})$, which implies (v). Constraint (vii) and $z(\pi)$ non-decreasing imply that non-negativity reduces to (vi). The (IC) constraint (4.3) follows from (v) and the (IC) constraint (4.2). The (IC) constraint (4.4) follows from Lemma 4.3 and (v).

Using $(\Delta(\pi), \Gamma(\pi), \hat{\Delta}, \hat{\Gamma})$ and (v), we can write (A.41) as

$$\begin{aligned} U = &- 2(1-\lambda) \int_{\frac{1}{2}}^{\bar{\pi}} \left[\pi e^{-\rho\left[z(\pi)d - \frac{\Delta(\pi)}{2}\right]} + (1-\pi) e^{\rho\left[z(\pi)d - \frac{\Delta(\pi)}{2}\right]} \right] e^{\rho\Gamma(\pi)} h(\pi) d\pi \\ &- \frac{\lambda}{2} \left[e^{-\rho\left[z(\bar{\pi})d - \frac{\Delta(\bar{\pi})}{2}\right]} + e^{\rho\left[z(\bar{\pi})d - \frac{\Delta(\bar{\pi})}{2}\right]} \right] e^{\rho\Gamma(\bar{\pi})}. \end{aligned} \tag{A.42}$$

The problem $(\mathcal{P})$ reduces to maximizing (A.42) over $(z(\pi), \Delta(\pi), \Gamma(\pi))$ for $\pi \in [\frac{1}{2}, \bar{\pi}]$, subject to (i), (ii), (iii), (iv), (vi) and (vii). Since $\Delta(\pi) \equiv f(z(\pi)d) - f(-z(\pi)d)$, we must impose the additional constraint (viii) $\Delta(\pi)$ is non-negative, $\Delta(\pi) = 0$ when $z(\pi) = 0$, and $\Delta(\pi)$ is constant in any interval where $z(\pi)$ is constant.

The limit when ϵ goes to zero of the solution to $(\mathcal{P})$ is the solution to maximizing (A.42) over $(z(\pi), \Delta(\pi), \Gamma(\pi))$ for $\pi \in [\frac{1}{2}, \bar{\pi}]$, subject to (i), (ii), (iii), (iv), (vi), (viii) and $z(\pi)$ non-decreasing. The reason why we can replace (vii) by $z(\pi)$ non-decreasing when deriving the limit is that for all $\epsilon > 0$ (ii) and (vii) imply $z(\pi)$ non-decreasing, but for $\epsilon = 0$ (vii) is implied by $z(\pi)$ non-decreasing and (4.2) (or equivalently (ii) and (iii)) and is hence redundant.

Using (A.21) and $\Gamma(\bar{\pi}) = \frac{\Delta(\bar{\pi})}{2}$, which follow from (iii) and (vi), respectively, we can write (A.42) as

$$\begin{aligned} U = &- 2(1-\lambda) \int_{\frac{1}{2}}^{\bar{\pi}} \left[\pi e^{-\rho\left[z(\pi)d - \frac{\Delta(\pi)}{2}\right]} + (1-\pi) e^{\rho\left[z(\pi)d - \frac{\Delta(\pi)}{2}\right]} \right] \\ &\times \exp\left[-\frac{\rho}{\bar{\rho}} \int_{\pi}^{\bar{\pi}} H(\Delta(\pi'), \pi') d\pi' \right] \left[\frac{\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} + (1-\pi) e^{\frac{\bar{\rho}}{2}\Delta(\pi)}}{\bar{\pi} e^{-\bar{\rho}\Delta(\bar{\pi})} + 1 - \bar{\pi}} \right]^{\frac{\rho}{\bar{\rho}}} h(\pi) d\pi \\ &- \frac{\lambda}{2} \left[e^{-\rho[z(\bar{\pi})d - \Delta(\bar{\pi})]} + e^{\rho z(\bar{\pi})d} \right]. \end{aligned} \tag{A.43}$$

The problem $(\mathcal{P})$ reduces to maximizing (A.43) over $(z(\pi), \Delta(\pi))$ for $\pi \in [\frac{1}{2}, \bar{\pi}]$, subject to (i), (ii), (iv), (viii) and $z(\pi)$ non-decreasing. In other words, (A.43) must be maximized over non-negative and non-decreasing $(z(\pi), \Delta(\pi))$, subject to the (IC) constraint (4.10) and the constraint that $\Delta(\pi) = 0$ when $z(\pi) = 0$ and that $\Delta(\pi)$ is constant in any interval where $z(\pi)$ is constant.

Without loss of generality, we can assume $z(\frac{1}{2}) = 0$ and hence $\Delta(\frac{1}{2}) = 0$. Indeed, if $z(\frac{1}{2}) > 0$, then we can set $z(\frac{1}{2})$ and $\Delta(\frac{1}{2})$ to zero. Since the density $h(\pi)$ is continuous, this change does not affect the (IC) constraint (4.10) and the investor's utility (A.43). Moreover, the functions $(z(\pi), \Delta(\pi))$ remain non-negative and non-decreasing, and the constraint that $\Delta(\pi) = 0$ when $z(\pi) = 0$ and that $\Delta(\pi)$ is constant in any interval where $z(\pi)$ is constant, remains satisfied.

We next consider

$$\hat{\pi} \equiv \inf\{\pi : z(\pi) = z(\bar{\pi}) \, \forall \, \pi \in [\hat{\pi}, \bar{\pi}]\},$$

and show that $\hat{\pi}$ is equal to π^* defined in (4.15) and that $z(\bar{\pi})$ is given by (4.16). Since $z(\pi)$ is constant in $[\hat{\pi}, \bar{\pi}]$, $\Delta(\pi)$ is also constant in that interval. Equation (4.15) defines $\pi^* \in (\frac{1}{2}, \bar{\pi})$ uniquely because the left-hand side is decreasing in π^*, is positive for $\pi^* = \frac{1}{2}$, and is negative for $\pi^* = \bar{\pi}$. Equation (4.15) implies

$$\begin{aligned} 2(1-\lambda)\int_{\pi^*}^{\bar{\pi}} \pi h(\pi) d\pi + \frac{\lambda}{2} &= \frac{2(1-\lambda)}{\pi^* - \frac{1}{2}} \left[\left(\pi^* - \frac{1}{2}\right) \int_{\pi^*}^{\bar{\pi}} \pi h(\pi) d\pi + \frac{1}{2} \int_{\pi^*}^{\bar{\pi}} (\pi - \pi^*) h(\pi) d\pi \right] \\ &= \frac{2(1-\lambda)\pi^*}{\pi^* - \frac{1}{2}} \int_{\pi^*}^{\bar{\pi}} \left(\pi - \frac{1}{2}\right) h(\pi) d\pi, \end{aligned} \tag{A.44}$$

and

$$\begin{aligned} 2(1-\lambda)\int_{\pi^*}^{\bar{\pi}} (1-\pi) h(\pi) d\pi + \frac{\lambda}{2} &= \frac{2(1-\lambda)}{\pi^* - \frac{1}{2}} \left[\left(\pi^* - \frac{1}{2}\right) \int_{\pi^*}^{\bar{\pi}} (1-\pi) h(\pi) d\pi + \frac{1}{2} \int_{\pi^*}^{\bar{\pi}} (\pi - \pi^*) h(\pi) d\pi \right] \\ &= \frac{2(1-\lambda)(1-\pi^*)}{\pi^* - \frac{1}{2}} \int_{\pi^*}^{\bar{\pi}} \left(\pi - \frac{1}{2}\right) h(\pi) d\pi. \end{aligned} \tag{A.45}$$

To show that $\hat{\pi}$ is equal to π^* defined in (4.15) and that $z(\bar{\pi})$ is given by (4.16), we proceed by contradiction. We assume that these properties are not true, and show that the investor can raise his utility by changing $z(\pi)$ while leaving $\Delta(\pi)$ the same. Since $\Delta(\pi)$ does not change, it remains non-negative and non-deceasing, and the (IC) constraint (4.10) remains satisfied. Moreover, under all the changes that we consider, $z(\pi)$ remains non-negative and non-decreasing, and $\Delta(\pi)$ remains equal to zero when $z(\pi) = 0$ and remains constant in any interval where $z(\pi)$ is constant.

Suppose that $\hat{\pi} < \pi^*$, and consider first the case where

$$z(\bar{\pi}) \leq \frac{1}{2\rho d} \log \left(\frac{2(1-\lambda)\int_{\hat{\pi}}^{\bar{\pi}} \pi h(\pi) d\pi + \frac{\lambda}{2}}{2(1-\lambda)\int_{\hat{\pi}}^{\bar{\pi}} (1-\pi) h(\pi) d\pi + \frac{\lambda}{2}} \right) + \frac{\Delta(\bar{\pi})}{2d}. \tag{A.46}$$

If the investor replaces $z(\pi)$ for all $\pi \in [\pi^*, \bar{\pi}]$ by $z(\bar{\pi}) + \phi$, for small $\phi > 0$, then his utility (A.42) becomes

$$\begin{aligned} U =& -2(1-\lambda)\left\{\int_{\frac{1}{2}}^{\pi^*}\left[\pi e^{-\rho\left[z(\pi)d-\frac{\Delta(\pi)}{2}\right]}+(1-\pi)e^{\rho\left[z(\pi)d-\frac{\Delta(\pi)}{2}\right]}\right]e^{\rho\Gamma(\pi)}h(\pi)d\pi \right.\\ &\left. +\left[e^{-\rho\left[(z(\bar\pi)+\phi)d-\frac{\Delta(\bar\pi)}{2}\right]}\int_{\pi^*}^{\bar\pi}\pi h(\pi)d\pi+e^{\rho\left[(z(\bar\pi)+\phi)d-\frac{\Delta(\bar\pi)}{2}\right]}\int_{\pi^*}^{\bar\pi}(1-\pi)h(\pi)d\pi\right]e^{\rho\Gamma(\bar\pi)}\right\}\\ &-\frac{\lambda}{2}\left[e^{-\rho\left[(z(\bar\pi)+\phi)d-\frac{\Delta(\bar\pi)}{2}\right]}+e^{\rho\left[(z(\bar\pi)+\phi)d-\frac{\Delta(\bar\pi)}{2}\right]}\right]e^{\rho\Gamma(\bar\pi)}. \end{aligned} \tag{A.47}$$

Equation (A.47) implies

$$\begin{aligned} \left.\frac{\partial U}{\partial\phi}\right|_{\phi=0} =&\rho d\left[e^{-\rho\left[z(\bar\pi)d-\frac{\Delta(\bar\pi)}{2}\right]}\left(2(1-\lambda)\int_{\pi^*}^{\bar\pi}\pi h(\pi)d\pi+\frac{\lambda}{2}\right)\right.\\ &\left.-e^{\rho\left[z(\bar\pi)d-\frac{\Delta(\bar\pi)}{2}\right]}\left(2(1-\lambda)\int_{\pi^*}^{\bar\pi}(1-\pi)h(\pi)d\pi+\frac{\lambda}{2}\right)\right]e^{\rho\Gamma(\bar\pi)}. \end{aligned}$$

Hence, utility increases if

$$\left.\frac{\partial U}{\partial\phi}\right|_{\phi=0}>0\Leftrightarrow z(\bar\pi)<\frac{1}{2\rho d}\log\left(\frac{2(1-\lambda)\int_{\pi^*}^{\bar\pi}\pi h(\pi)d\pi+\frac{\lambda}{2}}{2(1-\lambda)\int_{\pi^*}^{\bar\pi}(1-\pi)h(\pi)d\pi+\frac{\lambda}{2}}\right)+\frac{\Delta(\bar\pi)}{2d}. \tag{A.48}$$

Equation (A.46) implies (A.48) if

$$\begin{aligned} &\frac{2(1-\lambda)\int_{\hat\pi}^{\bar\pi}\pi h(\pi)d\pi+\frac{\lambda}{2}}{2(1-\lambda)\int_{\hat\pi}^{\bar\pi}(1-\pi)h(\pi)d\pi+\frac{\lambda}{2}}<\frac{2(1-\lambda)\int_{\pi^*}^{\bar\pi}\pi h(\pi)d\pi+\frac{\lambda}{2}}{2(1-\lambda)\int_{\pi^*}^{\bar\pi}(1-\pi)h(\pi)d\pi+\frac{\lambda}{2}}\\ \Leftrightarrow&\frac{\int_{\hat\pi}^{\pi^*}\pi h(\pi)d\pi+\frac{\pi^*}{\pi^*-\frac{1}{2}}\int_{\pi^*}^{\bar\pi}\left(\pi-\frac{1}{2}\right)h(\pi)d\pi}{\int_{\hat\pi}^{\pi^*}(1-\pi)h(\pi)d\pi+\frac{1-\pi^*}{\pi^*-\frac{1}{2}}\int_{\pi^*}^{\bar\pi}\left(\pi-\frac{1}{2}\right)h(\pi)d\pi}<\frac{\pi^*}{1-\pi^*}, \end{aligned} \tag{A.49}$$

where the equivalence follows from (A.44) and (A.45). Equation (A.49) is equivalent to

$$\frac{\int_{\hat\pi}^{\pi^*}\pi h(\pi)d\pi}{\int_{\hat\pi}^{\pi^*}(1-\pi)h(\pi)d\pi}<\frac{\pi^*}{1-\pi^*},$$

which holds because $\hat\pi<\pi^*$.

Consider next the case where

$$z(\bar\pi)>\frac{1}{2\rho d}\log\left(\frac{2(1-\lambda)\int_{\hat\pi}^{\bar\pi}\pi h(\pi)d\pi+\frac{\lambda}{2}}{2(1-\lambda)\int_{\hat\pi}^{\bar\pi}(1-\pi)h(\pi)d\pi+\frac{\lambda}{2}}\right)+\frac{\Delta(\bar\pi)}{2d}. \tag{A.50}$$

In the sub-case $\hat\pi>\frac{1}{2}$, consider a small $\eta>0$ and suppose that the investor replaces $z(\pi)$ for all $\pi\in[\hat\pi-\eta,\bar\pi]$ by $(1-\phi)z(\pi)+\phi z(\hat\pi-\eta)$, for small $\phi>0$. Under this change $z(\pi)$ decreases for all $\pi\in[\hat\pi-\eta,\bar\pi]$ because $z(\pi)\geq z(\hat\pi-\eta)$. Moreover, the decrease is strict for all $\pi\in(\hat\pi,\bar\pi]$ because

$z(\pi) = z(\bar{\pi}) > z(\hat{\pi} - \eta)$. The investor's utility (A.42) becomes

$$\begin{aligned} U =& -2(1-\lambda)\left\{\int_{\frac{1}{2}}^{\hat{\pi}-\eta}\left[\pi e^{-\rho\left[z(\pi)d-\frac{\Delta(\pi)}{2}\right]}+(1-\pi)e^{\rho\left[z(\pi)d-\frac{\Delta(\pi)}{2}\right]}\right]e^{\rho\Gamma(\pi)}h(\pi)d\pi\right. \\ &+\int_{\hat{\pi}-\eta}^{\hat{\pi}}\left[\pi e^{-\rho\left[[(1-\phi)z(\pi)+\phi z(\hat{\pi}-\eta)]d-\frac{\Delta(\pi)}{2}\right]}+(1-\pi)e^{\rho\left[[(1-\phi)z(\pi)+\phi z(\hat{\pi}-\eta)]d-\frac{\Delta(\pi)}{2}\right]}\right]e^{\rho\Gamma(\pi)}h(\pi)d\pi \\ &+\left[e^{-\rho\left[[(1-\phi)z(\bar{\pi})+\phi z(\hat{\pi}-\eta)]d-\frac{\Delta(\bar{\pi})}{2}\right]}\int_{\hat{\pi}}^{\bar{\pi}}\pi h(\pi)d\pi\right. \\ &\left.\left.+e^{\rho\left[[(1-\phi)z(\bar{\pi})+\phi z(\hat{\pi}-\eta)]d-\frac{\Delta(\bar{\pi})}{2}\right]}\int_{\hat{\pi}}^{\bar{\pi}}(1-\pi)h(\pi)d\pi\right]e^{\rho\Gamma(\bar{\pi})}\right\} \\ &-\frac{\lambda}{2}\left[e^{-\rho\left[[(1-\phi)z(\bar{\pi})+\phi z(\hat{\pi}-\eta)]d-\frac{\Delta(\bar{\pi})}{2}\right]}+e^{\rho\left[[(1-\phi)z(\bar{\pi})+\phi z(\hat{\pi}-\eta)]d-\frac{\Delta(\bar{\pi})}{2}\right]}\right]e^{\rho\Gamma(\bar{\pi})}. \end{aligned} \tag{A.51}$$

Equation (A.51) implies

$$\begin{aligned} \left.\frac{\partial U}{\partial\phi}\right|_{\phi=0} =& 2(1-\lambda)\rho d\int_{\hat{\pi}-\eta}^{\hat{\pi}}[z(\hat{\pi}-\eta)-z(\pi)]\left[\pi e^{-\rho\left[z(\pi)d-\frac{\Delta(\pi)}{2}\right]}+(1-\pi)e^{\rho\left[z(\pi)d-\frac{\Delta(\pi)}{2}\right]}\right]e^{\rho\Gamma(\pi)}h(\pi)d\pi \\ &+\rho d[z(\hat{\pi}-\eta)-z(\bar{\pi})]\left[e^{-\rho\left[z(\bar{\pi})d-\frac{\Delta(\bar{\pi})}{2}\right]}\left(2(1-\lambda)\int_{\hat{\pi}}^{\bar{\pi}}\pi h(\pi)d\pi+\frac{\lambda}{2}\right)\right. \\ &\left.-e^{\rho\left[z(\bar{\pi})d-\frac{\Delta(\bar{\pi})}{2}\right]}\left(2(1-\lambda)\int_{\hat{\pi}}^{\bar{\pi}}(1-\pi)h(\pi)d\pi+\frac{\lambda}{2}\right)\right]e^{\rho\Gamma(\bar{\pi})}. \end{aligned} \tag{A.52}$$

Since (A.50) implies

$$e^{-\rho\left[z(\bar{\pi})d-\frac{\Delta(\bar{\pi})}{2}\right]}\left(2(1-\lambda)\int_{\hat{\pi}}^{\bar{\pi}}\pi h(\pi)d\pi+\frac{\lambda}{2}\right)-e^{\rho\left[z(\bar{\pi})d-\frac{\Delta(\bar{\pi})}{2}\right]}\left(2(1-\lambda)\int_{\hat{\pi}}^{\bar{\pi}}(1-\pi)h(\pi)d\pi+\frac{\lambda}{2}\right)<0,$$

the second term in (A.52) is non-zero and dominates the first term for small η. Since, in addition, $z(\bar{\pi}) > z(\hat{\pi} - \eta)$, the second term in (A.52) is positive. Hence, $\left.\frac{\partial U}{\partial\phi}\right|_{\phi=0} > 0$.

In the sub-case $\hat{\pi} = \frac{1}{2}$, suppose that the investor replaces $z(\pi)$ for all $\pi \in (\hat{\pi}, \bar{\pi}]$ by $z(\bar{\pi}) - \phi$, for small $\phi > 0$. The investor's utility (A.42) becomes

$$\begin{aligned} U =& -2(1-\lambda)\left[e^{-\rho\left[(z(\bar{\pi})-\phi)d-\frac{\Delta(\bar{\pi})}{2}\right]}\int_{\frac{1}{2}}^{\bar{\pi}}\pi h(\pi)d\pi+e^{\rho\left[(z(\bar{\pi})-\phi)d-\frac{\Delta(\bar{\pi})}{2}\right]}\int_{\frac{1}{2}}^{\bar{\pi}}(1-\pi)h(\pi)d\pi\right]e^{\rho\Gamma(\bar{\pi})} \\ &-\frac{\lambda}{2}\left[e^{-\rho\left[(z(\bar{\pi})-\phi)d-\frac{\Delta(\bar{\pi})}{2}\right]}+e^{\rho\left[(z(\bar{\pi})-\phi)d-\frac{\Delta(\bar{\pi})}{2}\right]}\right]e^{\rho\Gamma(\bar{\pi})}. \end{aligned} \tag{A.53}$$

Equation (A.53) implies

$$\left.\frac{\partial U}{\partial \phi}\right|_{\phi=0} = -\rho d \left[e^{-\rho\left[z(\bar{\pi})d-\frac{\Delta(\bar{\pi})}{2}\right]}\left(2(1-\lambda)\int_{\frac{1}{2}}^{\bar{\pi}}\pi h(\pi)d\pi+\frac{\lambda}{2}\right) \right.$$
$$\left. -e^{\rho\left[z(\bar{\pi})d-\frac{\Delta(\bar{\pi})}{2}\right]}\left(2(1-\lambda)\int_{\frac{1}{2}}^{\bar{\pi}}(1-\pi)h(\pi)d\pi+\frac{\lambda}{2}\right)\right] e^{\rho\Gamma(\bar{\pi})}.$$

Since (A.50) implies that the term in square brackets is negative, $\left.\frac{\partial U}{\partial \phi}\right|_{\phi=0} > 0$.

Suppose next that $\hat{\pi} > \pi^*$, and consider first the case where (A.50) holds. The investor can raise his utility through the same change in $z(\pi)$ as in the case where $\hat{\pi} < \pi^*$, (A.50) holds, and $\hat{\pi} > \frac{1}{2}$.

Consider next the case where (A.46) holds. Consider a small $\eta > 0$, and suppose that the investor replaces $z(\pi)$ for all $\pi \in [\hat{\pi}-\eta, \bar{\pi}]$ by $(1-\phi)z(\pi) + \phi z(\hat{\pi}-\eta)$, for small $\phi < 0$. Under this change $z(\pi)$ increases for all $\pi \in [\hat{\pi}-\eta, \bar{\pi}]$ because $z(\pi) \geq z(\hat{\pi}-\eta)$. Moreover, the increase is strict for all $\pi \in (\hat{\pi}, \bar{\pi}]$ because $z(\pi) = z(\bar{\pi}) > z(\hat{\pi}-\eta)$. The investor's utility (A.42) becomes (A.51), and its partial derivative with respect to ϕ is (A.52). When (A.46) holds as a strict inequality, it implies together with $z(\bar{\pi}) > z(\hat{\pi}-\eta)$ that the second term in (A.52) is negative. Hence, $\left.\frac{\partial U}{\partial \phi}\right|_{\phi=0} < 0$, and utility increases. When (A.46) holds as an equality, the second term in (A.52) is zero, and utility increases if

$$\left.\frac{\partial U}{\partial \phi}\right|_{\phi=0} < 0 \Leftrightarrow 2(1-\lambda)\rho d\int_{\hat{\pi}-\eta}^{\hat{\pi}}[z(\hat{\pi}-\eta)-z(\pi)]\left[\pi e^{-\rho\left[z(\pi)d-\frac{\Delta(\pi)}{2}\right]}+(1-\pi)e^{\rho\left[z(\pi)d-\frac{\Delta(\pi)}{2}\right]}\right]e^{\rho\Gamma(\pi)}h(\pi)d\pi < 0. \tag{A.54}$$

For small η, (A.54) is equivalent to

$$\hat{\pi}e^{-\rho\left[z(\hat{\pi}^-)d-\frac{\Delta(\hat{\pi}^-)}{2}\right]}+(1-\hat{\pi})e^{\rho\left[z(\hat{\pi}^-)d-\frac{\Delta(\hat{\pi}^-)}{2}\right]} < 0 \Leftrightarrow z(\hat{\pi}^-) < \frac{1}{2\rho d}\log\left(\frac{\hat{\pi}}{1-\hat{\pi}}\right)+\frac{\Delta(\hat{\pi}^-)}{2d}.$$

Hence, utility may not increase only if (A.46) holds as an equality and

$$z(\hat{\pi}^-) \geq \frac{1}{2\rho d}\log\left(\frac{\hat{\pi}}{1-\hat{\pi}}\right)+\frac{\Delta(\hat{\pi}^-)}{2d}. \tag{A.55}$$

When (A.46) holds as an equality

$$\begin{aligned}
z(\bar{\pi}) &= \frac{1}{2\rho d}\log\left(\frac{2(1-\lambda)\int_{\hat{\pi}}^{\bar{\pi}}\pi h(\pi)d\pi+\frac{\lambda}{2}}{2(1-\lambda)\int_{\hat{\pi}}^{\bar{\pi}}(1-\pi)h(\pi)d\pi+\frac{\lambda}{2}}\right)+\frac{\Delta(\bar{\pi})}{2d}\\
&= \frac{1}{2\rho d}\log\left(\frac{\int_{\hat{\pi}}^{\pi^*}\pi h(\pi)d\pi+\frac{\pi^*}{\pi^*-\frac{1}{2}}\int_{\pi^*}^{\bar{\pi}}\left(\pi-\frac{1}{2}\right)h(\pi)d\pi}{\int_{\hat{\pi}}^{\pi^*}(1-\pi)h(\pi)d\pi+\frac{1-\pi^*}{\pi^*-\frac{1}{2}}\int_{\pi^*}^{\bar{\pi}}\left(\pi-\frac{1}{2}\right)h(\pi)d\pi}\right)+\frac{\Delta(\bar{\pi})}{2d}\\
&< \frac{1}{2\rho d}\log\left(\frac{\pi^*}{1-\pi^*}\right)+\frac{\Delta(\bar{\pi})}{2d},
\end{aligned} \tag{A.56}$$

where the second step follows from (A.44) and (A.45), and the third step follows from $\pi^* < \hat{\pi}$. Equations (A.55), (A.56), $z(\bar{\pi}) = z(\hat{\pi}^+) \geq z(\hat{\pi}^-)$, and $\pi^* < \hat{\pi}$ imply $\Delta(\bar{\pi}) = \Delta(\hat{\pi}^+) > \Delta(\hat{\pi}^-)$, i.e., $\Delta(\pi)$ is discontinuous at $\hat{\pi}$. We next show that this discontinuity contradicts the optimality of $\Delta(\pi)$. This implies that (A.46) cannot hold as an equality together with (A.55), and completes our proof that $\hat{\pi}$ cannot exceed π^*.

To show that $\Delta(\pi)$ discontinuous at $\hat{\pi}$ contradicts the optimality of $\Delta(\pi)$, we distinguish cases. Consider first the case $\hat{\pi} < \bar{\pi}$. Consider a small $\eta > 0$, and suppose that the investor replaces $\Delta(\pi)$ for all $\pi \in [\hat{\pi}-\eta, \hat{\pi})$ by $\Delta(\hat{\pi}^-)+\phi^-$, and replaces $\Delta(\pi)$ for all $\pi \in (\hat{\pi}, \hat{\pi}+\eta]$ by $\Delta(\bar{\pi})-\phi^+$, where

$$\begin{aligned}
&\left.\frac{\partial H(\Delta,\hat{\pi})}{\partial\Delta}\right|_{\Delta=\Delta_{\hat{\pi}^-}}\phi^- = \left.\frac{\partial H(\Delta,\hat{\pi})}{\partial\Delta}\right|_{\Delta=\Delta_{\bar{\pi}}}\phi^+\\
\Leftrightarrow\ &\frac{1}{\left[\hat{\pi}e^{-\bar{\rho}\frac{\Delta(\pi^-)}{2}}+(1-\hat{\pi})e^{\bar{\rho}\frac{\Delta(\pi^-)}{2}}\right]^2}\phi^- = \frac{1}{\left[\hat{\pi}e^{-\bar{\rho}\frac{\Delta(\bar{\pi})}{2}}+(1-\hat{\pi})e^{\bar{\rho}\frac{\Delta(\bar{\pi})}{2}}\right]^2}\phi^+ \equiv \phi,
\end{aligned} \tag{A.57}$$

for small $\phi > 0$. Since the (IC) constraint (4.10) can be written as

$$2e^{\bar{\rho}K}\int_{\frac{1}{2}}^{\bar{\pi}}\exp\left[-\int_{\frac{1}{2}}^{\pi}H(\Delta(\pi'),\pi')d\pi'\right]h(\pi)d\pi \leq 1, \tag{A.58}$$

its left-hand side remains the same except possibly for terms of order smaller than $\phi\eta$. Equation

(A.43) implies that the change in utility is of order $\phi\eta$, and is positive if

$$\left[\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}+(1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\right]^{\frac{\rho}{\bar{\rho}}}\left\{\hat{\pi}e^{-\rho\left[z(\bar{\pi})d-\frac{\Delta(\bar{\pi})}{2}\right]}-(1-\hat{\pi})e^{\rho\left[z(\bar{\pi})d-\frac{\Delta(\bar{\pi})}{2}\right]}\right.$$
$$\left.+\left[\hat{\pi}e^{-\rho\left[z(\bar{\pi})d-\frac{\Delta(\bar{\pi})}{2}\right]}+(1-\hat{\pi})e^{\rho\left[z(\bar{\pi})d-\frac{\Delta(\bar{\pi})}{2}\right]}\right]\frac{(1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}-\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}}{\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}+(1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}}\right\}\phi^{+}$$
$$-\left[\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^{-})}+(1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^{-})}\right]^{\frac{\rho}{\bar{\rho}}}\left\{\hat{\pi}e^{-\rho\left[z(\hat{\pi}^{-})d-\frac{\Delta(\hat{\pi}^{-})}{2}\right]}-(1-\hat{\pi})e^{\rho\left[z(\hat{\pi}^{-})d-\frac{\Delta(\hat{\pi}^{-})}{2}\right]}\right.$$
$$\left.+\left[\hat{\pi}e^{-\rho\left[z(\hat{\pi}^{-})d-\frac{\Delta(\hat{\pi}^{-})}{2}\right]}+(1-\hat{\pi})e^{\rho\left[z(\hat{\pi}^{-})d-\frac{\Delta(\hat{\pi}^{-})}{2}\right]}\right]\frac{(1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^{-})}-\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^{-})}}{\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^{-})}+(1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^{-})}}\right\}\phi^{-}>0. \tag{A.59}$$

Rearranging terms in (A.59), we can write it as

$$\left[\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}+(1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\right]^{\frac{\rho}{\bar{\rho}}}\frac{e^{-\rho\left[z(\bar{\pi})d-\frac{\Delta(\bar{\pi})}{2}\right]}e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}-e^{\rho\left[z(\bar{\pi})d-\frac{\Delta(\bar{\pi})}{2}\right]}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}}{\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}+(1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}}\phi^{+}$$
$$>\left[\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^{-})}+(1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^{-})}\right]^{\frac{\rho}{\bar{\rho}}}\frac{e^{-\rho\left[z(\hat{\pi}^{-})d-\frac{\Delta(\hat{\pi}^{-})}{2}\right]}e^{\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^{-})}-e^{\rho\left[z(\hat{\pi}^{-})d-\frac{\Delta(\hat{\pi}^{-})}{2}\right]}e^{-\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^{-})}}{\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^{-})}+(1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^{-})}}\phi^{-}$$
$$\Leftrightarrow\left[\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}+(1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\right]^{\frac{\rho}{\bar{\rho}}+1}\left[e^{-\rho\left[z(\bar{\pi})d-\frac{\Delta(\bar{\pi})}{2}\right]}e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}-e^{\rho\left[z(\bar{\pi})d-\frac{\Delta(\bar{\pi})}{2}\right]}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\right]$$
$$>\left[\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^{-})}+(1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^{-})}\right]^{\frac{\rho}{\bar{\rho}}+1}\left[e^{-\rho\left[z(\hat{\pi}^{-})d-\frac{\Delta(\hat{\pi}^{-})}{2}\right]}e^{\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^{-})}-e^{\rho\left[z(\hat{\pi}^{-})d-\frac{\Delta(\hat{\pi}^{-})}{2}\right]}e^{-\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^{-})}\right], \tag{A.60}$$

where the second step follows from (A.57). Equation (A.56) implies

$$e^{-\rho\left[z(\bar{\pi})d-\frac{\Delta(\bar{\pi})}{2}\right]}e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}-e^{\rho\left[z(\bar{\pi})d-\frac{\Delta(\bar{\pi})}{2}\right]}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}>\sqrt{\frac{1-\pi^*}{\pi^*}}e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}-\sqrt{\frac{\pi^*}{1-\pi^*}}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}$$
$$=\frac{(1-\pi^*)e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}-\pi^*e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}}{\sqrt{\pi^*(1-\pi^*)}}>0, \tag{A.61}$$

where the last inequality follows from $\Delta(\bar{\pi})>\Delta^*(\bar{\pi})>\Delta^*(\pi^*)$. Equation (A.55) likewise implies

$$e^{-\rho\left[z(\hat{\pi}^{-})d-\frac{\Delta(\hat{\pi}^{-})}{2}\right]}e^{\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^{-})}-e^{\rho\left[z(\hat{\pi}^{-})d-\frac{\Delta(\hat{\pi}^{-})}{2}\right]}e^{-\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^{-})}<\sqrt{\frac{1-\hat{\pi}}{\hat{\pi}}}e^{\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^{-})}-\sqrt{\frac{\hat{\pi}}{1-\hat{\pi}}}e^{-\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^{-})}$$
$$=\frac{(1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^{-})}-\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^{-})}}{\sqrt{\hat{\pi}(1-\hat{\pi})}} \tag{A.62}$$

Equation (A.61) implies that the left-hand side of (A.60) is positive. If $\Delta(\hat{\pi}^{-})\leq\Delta^*(\hat{\pi})$, then (A.62) implies that the right-hand side of (A.60) is negative, and hence (A.60) holds. If instead

$\Delta(\hat{\pi}^-) > \Delta^*(\hat{\pi})$, then $\Delta(\bar{\pi}) > \Delta(\hat{\pi}^-) > \Delta^*(\hat{\pi})$ imply

$$\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})} + (1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})} > \hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^-)} + (1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^-)},$$

and (A.60) holds under the sufficient condition

$$e^{-\rho\left[z(\bar{\pi})d-\frac{\Delta(\bar{\pi})}{2}\right]}e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})} - e^{\rho\left[z(\bar{\pi})d-\frac{\Delta(\bar{\pi})}{2}\right]}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})} > e^{-\rho\left[z(\hat{\pi}^-)d-\frac{\Delta(\hat{\pi}^-)}{2}\right]}e^{\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^-)} - e^{\rho\left[z(\hat{\pi}^-)d-\frac{\Delta(\hat{\pi}^-)}{2}\right]}e^{-\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^-)}. \tag{A.63}$$

Equations (A.61) and (A.62) imply that (A.63) holds if

$$\sqrt{\frac{1-\pi^*}{\pi^*}}e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})} - \sqrt{\frac{\pi^*}{1-\pi^*}}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})} > \sqrt{\frac{1-\hat{\pi}}{\hat{\pi}}}e^{\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^-)} - \sqrt{\frac{\hat{\pi}}{1-\hat{\pi}}}e^{-\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^-)}. \tag{A.64}$$

Equation (A.64) holds because $\hat{\pi} > \pi^*$ implies

$$\sqrt{\frac{1-\pi^*}{\pi^*}}e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})} - \sqrt{\frac{\pi^*}{1-\pi^*}}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})} > \sqrt{\frac{1-\hat{\pi}}{\hat{\pi}}}e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})} - \sqrt{\frac{\hat{\pi}}{1-\hat{\pi}}}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}$$

and because $\Delta(\bar{\pi}) > \Delta(\hat{\pi}^-)$.

Consider next the case $\hat{\pi} = \bar{\pi}$. Suppose that the investor lowers $\Delta(\bar{\pi})$ by a small $\phi > 0$. Since $\Delta(\pi)$ for $\pi < \bar{\pi}$ remains the same, so does the left-hand side of the (IC) constraint (4.10). Moreover, the only effect on the integral in (A.43) is through the term $\bar{\pi}e^{-\bar{\rho}\Delta(\bar{\pi})} + 1 - \bar{\pi}$, which increases. Hence, the term in (A.43) that corresponds to the risk-averse types increases. Since the term that corresponds to the risk-neutral type also increases, utility increases.

Since $\hat{\pi}$ cannot exceed π^* and cannot be smaller than π^*, it is equal to π^*. If

$$z(\bar{\pi}) > \frac{1}{2\rho d}\log\left(\frac{2(1-\lambda)\int_{\pi^*}^{\bar{\pi}}\pi h(\pi)d\pi + \frac{\lambda}{2}}{2(1-\lambda)\int_{\pi^*}^{\bar{\pi}}(1-\pi)h(\pi)d\pi + \frac{\lambda}{2}}\right) + \frac{\Delta(\bar{\pi})}{2d}, \tag{A.65}$$

then the investor can raise his utility through the same change in $z(\pi)$ as in the case where $\hat{\pi} < \pi^*$, (A.50) holds, and $\hat{\pi} > \frac{1}{2}$. If instead (A.65) holds as a strict inequality in the other direction, then the investor can raise his utility through the same change in $z(\pi)$ as in the case where $\hat{\pi} > \pi^*$ and (A.46) holds. Hence, (A.65) holds as an equality, which means from (A.44) and (A.45) that $z(\bar{\pi})$ is given by (4.16).

To show that $z(\pi)$ is given by (4.17) for $\pi \in (\frac{1}{2}, \pi^*)$, we maximize (A.42) point-wise over $z(\pi)$, without requiring that $z(\pi)$ is non-negative and non-decreasing, and that $\Delta(\pi)$ is equal to zero when $z(\pi) = 0$ and is constant in any interval where $z(\pi)$ is constant. This point-wise maximization yields (4.17). Since $\Delta(\pi)$ is non-negative and non-decreasing, (4.17) implies that $z(\pi)$ is positive and increasing for $\pi \in (\frac{1}{2}, \pi^*)$. The properties that $z(\pi)$ is non-negative and non-decreasing extend to the larger interval $[\frac{1}{2}, \pi]$: in the case of $\frac{1}{2}$ because $z(\frac{1}{2}) = 0$, and in the case of $[\pi^*, \bar{\pi}]$ because (4.16), (4.17) and $\Delta(\pi^{*-}) \leq \Delta(\bar{\pi})$ imply $z(\pi^{*-}) \leq z(\bar{\pi})$. Since $z(\pi)$ is positive and increasing for $\pi \in (\frac{1}{2}, \pi^*)$, the constraint that $\Delta(\pi)$ is equal to zero when $z(\pi) = 0$ and is constant in any interval where $z(\pi)$ is constant is trivially satisfied.

We next show that $\Delta(\pi) > \Delta^*(\pi)$ for all $\pi \in (\frac{1}{2}, \bar{\pi}]$. This inequality holds for all $\pi \in [\pi^*, \bar{\pi}]$ because $\Delta(\pi) = \Delta(\bar{\pi}) > \Delta^*(\bar{\pi}) > \Delta^*(\pi)$. To show that it also holds for all $\pi \in (\frac{1}{2}, \pi^*)$, we consider

$$\hat{\pi} \equiv \inf\{\pi : \Delta(\pi) > \Delta^*(\pi) \, \forall \, \pi \in [\hat{\pi}, \bar{\pi}]\}$$

and suppose by contradiction that $\hat{\pi} > \frac{1}{2}$. Consider a small $\eta > 0$, and suppose that the investor replaces $\Delta(\pi)$ for all $\pi \in [\hat{\pi} - \eta, \hat{\pi})$ by $(1 - \phi)\Delta(\pi) + \phi\Delta^*(\hat{\pi})$ for small $\phi > 0$, and changes $z(\pi)$ for all $\pi \in [\hat{\pi} - \eta, \hat{\pi})$ so that (4.17) is satisfied for the new value of $\Delta(\pi)$. Under this change $\Delta(\pi)$ increases for all $\pi \in [\hat{\pi} - \eta, \hat{\pi})$ because the definition of $\hat{\pi}$ implies $\Delta(\pi) < \Delta^*(\hat{\pi})$ for $\pi < \hat{\pi}$. Since $H(\Delta, \pi)$ is increasing in Δ, the left-hand side of the (IC) constraint (4.10) decreases and hence that constraint remains satisfied. The derivative of the utility (A.43) with respect to ϕ is

$$\begin{aligned}
\left.\frac{\partial U}{\partial \phi}\right|_{\phi=0} = & -2\rho(1-\lambda)\left\{\int_{\hat{\pi}-\eta}^{\hat{\pi}} \sqrt{\pi(1-\pi)} \exp\left[-\frac{\rho}{\bar{\rho}}\int_{\pi}^{\bar{\pi}} H(\Delta(\pi'), \pi')d\pi'\right]\right. \\
& \times \frac{\left[(1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)} - \pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)}\right]\left[\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}\right]^{\frac{\rho}{\bar{\rho}}-1}}{\left[\bar{\pi}e^{-\bar{\rho}\Delta(\bar{\pi})} + 1 - \bar{\pi}\right]^{\frac{\rho}{\bar{\rho}}}}[\Delta^*(\hat{\pi}) - \Delta(\pi)]h(\pi)d\pi \\
& - \int_{\hat{\pi}-\eta}^{\hat{\pi}} \frac{2[\Delta^*(\hat{\pi}) - \Delta(\pi)]}{\left[\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}\right]^2} d\pi \int_{\frac{1}{2}}^{\hat{\pi}} \sqrt{\pi(1-\pi)} \exp\left[-\frac{\rho}{\bar{\rho}}\int_{\pi}^{\hat{\pi}} H(\Delta(\pi'), \pi')d\pi'\right] \\
& \left. \times \left[\frac{\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}}{\bar{\pi}e^{-\bar{\rho}\Delta(\bar{\pi})} + 1 - \bar{\pi}}\right]^{\frac{\rho}{\bar{\rho}}} h(\pi)d\pi \right\} \qquad \text{(A.66)}
\end{aligned}$$

plus terms of order smaller than $\int_{\hat{\pi}-\eta}^{\hat{\pi}}[\Delta^*(\hat{\pi}) - \Delta(\pi)]d\pi$. Since $\Delta(\pi) < \Delta^*(\hat{\pi})$ implies

$$(1 - \hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\pi)} - \hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} < 0,$$

the term

$$\int_{\hat{\pi}-\eta}^{\hat{\pi}} \sqrt{\pi(1-\pi)} \exp\left[-\frac{\rho}{\bar{\rho}} \int_{\pi}^{\bar{\pi}} H(\Delta(\pi'), \pi') d\pi'\right]$$
$$\times \frac{\left[(1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)} - \pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)}\right] \left[\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}\right]^{\frac{\rho}{\bar{\rho}}-1}}{\left[\bar{\pi}e^{-\bar{\rho}\Delta(\bar{\pi})} + 1 - \bar{\pi}\right]^{\frac{\rho}{\bar{\rho}}}} [\Delta(\hat{\pi}) - \Delta(\pi)] h(\pi) d\pi$$

in the curly bracket is negative or of order smaller than $\int_{\hat{\pi}-\eta}^{\hat{\pi}} [\Delta(\bar{\pi}) - \Delta(\pi)] d\pi$. Because the other term is negative, utility increases.

We next show that $\Delta(\pi)$ is continuous for $\pi \in (\frac{1}{2}, \bar{\pi}]$. Suppose, by contradiction, that there is a discontinuity at $\hat{\pi} \in (\frac{1}{2}, \pi^*]$. Consider a small $\eta > 0$, and suppose that the investor replaces $\Delta(\pi)$ for all $\pi \in [\hat{\pi} - \eta, \hat{\pi})$ by $\Delta(\hat{\pi}^-) + \phi^-$, and replaces $\Delta(\pi)$ for all $\pi \in (\hat{\pi}, \hat{\pi} + \eta]$ by $\Delta(\hat{\pi}^+) - \phi^+$, where (ϕ^-, ϕ^+) are defined as in (A.57) with $\Delta(\hat{\pi}^+)$ instead of $\Delta(\bar{\pi})$. Since (4.16) and (4.17) imply

$$e^{-\rho\left[z(\hat{\pi}^-)d - \frac{\Delta(\hat{\pi}^-)}{2}\right]} = e^{-\rho\left[z(\hat{\pi}^+)d - \frac{\Delta(\hat{\pi}^+)}{2}\right]} = \sqrt{\frac{1-\hat{\pi}}{\hat{\pi}}}$$

(A.60) implies that the change in utility is positive if

$$\left[\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^+)} + (1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^+)}\right]^{\frac{\rho}{\bar{\rho}}+1} \left[(1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^+)} - \hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^+)}\right]$$
$$> \left[\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^-)} + (1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^-)}\right]^{\frac{\rho}{\bar{\rho}}+1} \left[(1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^-)} - \hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^-)}\right]. \tag{A.67}$$

Since $\Delta(\hat{\pi}^+) > \Delta(\hat{\pi}^-) > \Delta^*(\hat{\pi})$,

$$\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^+)} + (1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^+)} > \hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^-)} + (1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^-)}.$$

Hence, (A.67) holds under the sufficient condition

$$(1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^+)} - \hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^+)} > (1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^-)} - \hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\hat{\pi}^-)} > 0,$$

which is satisfied because $\Delta(\hat{\pi}^+) > \Delta(\hat{\pi}^-) > \Delta^*(\hat{\pi})$.

We finally derive the employment condition (4.14). Using (4.16), (4.17), and $(\Delta(\pi), \Gamma(\pi)) =$

$(\Delta(\bar{\pi}), \frac{\Delta(\bar{\pi})}{2})$ for $\pi \in [\pi^*, \bar{\pi}]$, we can write the investor's utility (A.42) as

$$\begin{aligned} U = & -4(1-\lambda)\int_{\frac{1}{2}}^{\pi^*} e^{\rho\Gamma(\pi)}\sqrt{\pi(1-\pi)}h(\pi)d\pi \\ & -2e^{\frac{\rho}{2}\Delta(\bar{\pi})}\sqrt{\left[2(1-\lambda)\int_{\pi^*}^{\bar{\pi}} \pi h(\pi)d\pi + \frac{\lambda}{2}\right]\left[2(1-\lambda)\int_{\pi^*}^{\bar{\pi}}(1-\pi)h(\pi)d\pi + \frac{\lambda}{2}\right]} \\ = & -4(1-\lambda)\left[\int_{\frac{1}{2}}^{\pi^*} e^{\rho\Gamma(\pi)}\sqrt{\pi(1-\pi)}h(\pi)d\pi + e^{\frac{\rho\Delta(\bar{\pi})}{2}}\sqrt{\pi^*(1-\pi^*)}\frac{\int_{\pi^*}^{\bar{\pi}}\left(\pi-\frac{1}{2}\right)h(\pi)d\pi}{\pi^*-\frac{1}{2}}\right], \end{aligned} \quad \text{(A.68)}$$

where the second step follows from (A.44) and (A.45). The utility (A.68) exceeds the utility $U = -1$ that the investor receives by investing directly if (4.14) holds. □

To prove Proposition **??**, we first show the following Lemma.

Lemma A.1. *Suppose that the pooling condition (4.13) holds. If*

$$\frac{\int_{\hat{\pi}}^{\bar{\pi}}\sqrt{\pi(1-\pi)}h(\pi)d\pi}{\int_{\hat{\pi}}^{\bar{\pi}}(\pi-\hat{\pi})h(\pi)d\pi} > \frac{\int_{\frac{1}{2}}^{\hat{\pi}}\sqrt{\pi(1-\pi)}h(\pi)d\pi}{(2\hat{\pi}-1)\int_{\hat{\pi}}^{\bar{\pi}}(1-\pi)h(\pi)d\pi} \quad \text{(A.69)}$$

for all $\hat{\pi} \in (\frac{1}{2}, \pi^*]$*, then* $\Delta(\pi)$ *is constant in* $(\frac{1}{2}, \bar{\pi}]$*.*

Proof of Lemma A.1. Showing that $\Delta(\pi)$ is constant in $(\frac{1}{2}, \bar{\pi}]$ amounts to showing that $\hat{\pi}$ defined by

$$\hat{\pi} \equiv \inf\{\pi : \Delta(\pi) = \Delta(\bar{\pi})\, \forall\, \pi \in (\hat{\pi}, \bar{\pi}]\}$$

is equal to $\frac{1}{2}$. Since $\Delta(\pi)$ is constant in $[\pi^*, \bar{\pi}]$, $\hat{\pi}$ does not exceed π^*. Suppose, by contradiction, that $\hat{\pi} \in (\frac{1}{2}, \pi^*]$. We will construct a change in $\Delta(\pi)$ that raises the investor's utility while keeping the left-hand side of the (IC) constraint (4.10) unchanged. Before constructing that change, we rewrite the (IC) constraint (4.10) and the investor's utility using that $\Delta(\pi)$ is constant in $(\hat{\pi}, \bar{\pi}]$ and that $z(\pi)$ is given by (4.16) and (4.17).

Since $\Delta(\pi)$ is constant in $(\hat{\pi},\bar{\pi}]$, we can write the (IC) constraint (4.10) as

$$\begin{aligned}
&2e^{\bar{\rho}K}\left[\int_{\frac{1}{2}}^{\hat{\pi}}\exp\left[-\int_{\frac{1}{2}}^{\pi}H(\Delta(\pi'),\pi')d\pi'\right]h(\pi)d\pi\right.\\
&\left.+\exp\left[-\int_{\frac{1}{2}}^{\hat{\pi}}H(\Delta(\pi'),\pi')d\pi'\right]\int_{\hat{\pi}}^{\bar{\pi}}\exp\left[-\int_{\hat{\pi}}^{\pi}H(\Delta(\bar{\pi}),\pi')d\pi'\right]h(\pi)d\pi\right]\leq 1\\
&\Leftrightarrow 2e^{\bar{\rho}K}\left[\int_{\frac{1}{2}}^{\hat{\pi}}\exp\left[-\int_{\frac{1}{2}}^{\pi}H(\Delta(\pi'),\pi')d\pi'\right]h(\pi)d\pi\right.\\
&\left.+\exp\left[-\int_{\frac{1}{2}}^{\hat{\pi}}H(\Delta(\pi'),\pi')d\pi'\right]\frac{e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\int_{\hat{\pi}}^{\bar{\pi}}\pi h(\pi)d\pi+e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\int_{\hat{\pi}}^{\bar{\pi}}(1-\pi)h(\pi)d\pi}{\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}+(1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}}\right]\leq 1,
\end{aligned}\tag{A.70}$$

where the second step follows because the same calculation as in (A.25) implies

$$\exp\left[-\int_{\hat{\pi}}^{\pi}H(\Delta(\bar{\pi}),\pi')d\pi'\right]=\frac{\pi e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}+(1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}}{\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}+(1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}}$$

for $\pi\in[\hat{\pi},\bar{\pi}]$.

Since $\Delta(\pi)$ is constant in $(\hat{\pi},\bar{\pi}]$, we can write the investor's utility (A.43) as

$$\begin{aligned}
U=&-2(1-\lambda)\exp\left[-\frac{\rho}{\bar{\rho}}\int_{\hat{\pi}}^{\bar{\pi}}H(\Delta(\bar{\pi}),\pi')d\pi'\right]\int_{\frac{1}{2}}^{\hat{\pi}}\left[\pi e^{-\rho\left[z(\pi)d-\frac{\Delta(\pi)}{2}\right]}+(1-\pi)e^{\rho\left[z(\pi)d-\frac{\Delta(\pi)}{2}\right]}\right]\\
&\times\exp\left[-\frac{\rho}{\bar{\rho}}\int_{\pi}^{\hat{\pi}}H(\Delta(\pi'),\pi')d\pi'\right]\left[\frac{\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)}+(1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}}{\bar{\pi}e^{-\bar{\rho}\Delta(\bar{\pi})}+1-\bar{\pi}}\right]^{\frac{\rho}{\bar{\rho}}}h(\pi)d\pi\\
&-2(1-\lambda)\int_{\hat{\pi}}^{\bar{\pi}}\left[\pi e^{-\rho\left[z(\pi)d-\frac{\Delta(\bar{\pi})}{2}\right]}+(1-\pi)e^{\rho\left[z(\pi)d-\frac{\Delta(\bar{\pi})}{2}\right]}\right]\\
&\times\exp\left[-\frac{\rho}{\bar{\rho}}\int_{\pi}^{\bar{\pi}}H(\Delta(\bar{\pi}),\pi')d\pi'\right]\left[\frac{\pi e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}+(1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}}{\bar{\pi}e^{-\bar{\rho}\Delta(\bar{\pi})}+1-\bar{\pi}}\right]^{\frac{\rho}{\bar{\rho}}}h(\pi)d\pi\\
&-\frac{\lambda}{2}\left[e^{-\rho[z(\bar{\pi})d-\Delta(\bar{\pi})]}+e^{\rho z(\bar{\pi})d}\right],
\end{aligned}\tag{A.71}$$

Since the same calculation as in (A.25) implies

$$\exp\left[-\int_{\pi}^{\bar{\pi}}H(\Delta(\bar{\pi}),\pi')d\pi'\right]=\frac{\bar{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}+(1-\bar{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}}{\pi e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}+(1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}}$$

for $\pi \in [\hat{\pi}, \bar{\pi}]$, we can write (A.71) as

$$U = -2(1-\lambda)e^{\frac{\rho}{2}\Delta(\bar{\pi})}\int_{\frac{1}{2}}^{\hat{\pi}}\left[\pi e^{-\rho\left[z(\pi)d - \frac{\Delta(\pi)}{2}\right]} + (1-\pi)e^{\rho\left[z(\pi)d - \frac{\Delta(\pi)}{2}\right]}\right]$$
$$\times \exp\left[-\frac{\rho}{\bar{\rho}}\int_{\pi}^{\hat{\pi}} H(\Delta(\pi'),\pi')d\pi'\right]\left[\frac{\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}}{\hat{\pi} e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})} + (1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}}\right]^{\frac{\rho}{\bar{\rho}}} h(\pi)d\pi$$
$$-2(1-\lambda)\int_{\hat{\pi}}^{\bar{\pi}}\left[\pi e^{-\rho[z(\pi)d - \Delta(\bar{\pi})]} + (1-\pi)e^{\rho z(\pi)d}\right]h(\pi)d\pi - \frac{\lambda}{2}\left[e^{-\rho[z(\bar{\pi})d - \Delta(\bar{\pi})]} + e^{\rho z(\bar{\pi})d}\right]$$
$$= -4(1-\lambda)e^{\frac{\rho}{2}\Delta(\bar{\pi})}\left\{\int_{\frac{1}{2}}^{\hat{\pi}}\sqrt{\pi(1-\pi)}\exp\left[-\frac{\rho}{\bar{\rho}}\int_{\pi}^{\hat{\pi}} H(\Delta(\pi'),\pi')d\pi'\right]\left[\frac{\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}}{\hat{\pi} e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})} + (1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}}\right]^{\frac{\rho}{\bar{\rho}}} h(\pi)d\pi\right.$$
$$\left. + \int_{\hat{\pi}}^{\pi^*}\sqrt{\pi(1-\pi)}h(\pi)d\pi + \sqrt{\pi^*(1-\pi^*)}\frac{\int_{\pi^*}^{\bar{\pi}}\left(\pi - \frac{1}{2}\right)h(\pi)d\pi}{\pi^* - \frac{1}{2}}\right\}, \tag{A.72}$$

where the second step follows from using (4.16), (4.17), (A.44) and (A.45).

Consider a small $\eta > 0$ and suppose that the investor replaces $\Delta(\pi)$ for $\pi \in [\hat{\pi}, \bar{\pi}]$ by $\Delta(\bar{\pi}) - \phi^+$, replaces $\Delta(\pi)$ for $\pi \in (\frac{1}{2}, \hat{\pi} - \eta]$ by $\Delta(\pi) + \phi^-$, and replaces $\Delta(\pi)$ for $\pi \in (\hat{\pi} - \eta, \hat{\pi})$ by

$$\left(1 - \frac{\hat{\pi} - \pi}{\eta}\right)\left[\Delta(\bar{\pi}) - \phi^+\right] + \frac{\hat{\pi} - \pi}{\eta}\left[\Delta(\pi - \eta) + \phi^-\right],$$

i.e., the point in the line linking $(\hat{\pi} - \eta, \Delta(\hat{\pi} - \eta) + \phi^-)$ and $(\hat{\pi}, \Delta(\bar{\pi}) - \phi^+)$. The constants (ϕ^-, ϕ^+) are positive, and such that $\Delta(\bar{\pi}) - \Delta(\pi - \eta) > \phi^+ + \phi^-$ so that $\Delta(\pi)$ remains non-decreasing. The investor also changes $z(\pi)$ for all $\pi \in (\frac{1}{2}, \bar{\pi}]$ so that (4.16) and (4.17) are satisfied for the new value of $\Delta(\pi)$. Under this change, $\Delta(\pi)$ becomes more uniform in $(\frac{1}{2}, \bar{\pi}]$: larger for $\pi \in (\frac{1}{2}, \hat{\pi} - \eta]$, and smaller for $\pi \in [\hat{\pi}, \bar{\pi}]$. To compute the change in the investor's utility and on the left-hand side $\mathcal{L}$ of the (IC) constraint (A.70), we compute the separate effects of changing ϕ^+ and ϕ^-.

Changing ϕ^+ affects $\Delta(\pi)$ for $\pi \in [\hat{\pi} - \eta, \bar{\pi}]$. For small η, the dominant term comes from considering values of π in $[\hat{\pi}, \bar{\pi}]$ only, and can be computed by taking the opposite of the derivative of (A.72) and $\mathcal{L}$ with respect to $\Delta(\bar{\pi})$. This yields

$$\left.\frac{\partial U}{\partial \phi^+}\right|_{\phi^+=0} = 2\rho(1-\lambda)e^{\frac{\rho}{2}\Delta(\bar{\pi})}\left\{2\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\right.$$
$$\times \int_{\frac{1}{2}}^{\hat{\pi}}\sqrt{\pi(1-\pi)}\exp\left[-\frac{\rho}{\bar{\rho}}\int_{\pi}^{\hat{\pi}} H(\Delta(\pi'),\pi')d\pi'\right]\frac{\left[\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}\right]^{\frac{\rho}{\bar{\rho}}}}{\left[\hat{\pi} e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})} + (1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\right]^{\frac{\rho}{\bar{\rho}}+1}}h(\pi)d\pi$$
$$\left. + \int_{\hat{\pi}}^{\pi^*}\sqrt{\pi(1-\pi)}h(\pi)d\pi + \sqrt{\pi^*(1-\pi^*)}\frac{\int_{\pi^*}^{\bar{\pi}}\left(\pi - \frac{1}{2}\right)h(\pi)d\pi}{\pi^* - \frac{1}{2}}\right\} + o(\eta), \tag{A.73}$$

and

$$\left.\frac{\partial \mathcal{L}}{\partial \phi^+}\right|_{\phi^+=0} = 2\bar{\rho}e^{\bar{\rho}K} \exp\left[-\int_{\frac{1}{2}}^{\hat{\pi}} H(\Delta(\pi'),\pi')d\pi'\right] \frac{\int_{\hat{\pi}}^{\bar{\pi}}(\pi-\hat{\pi})h(\pi)d\pi}{\left[\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}+(1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\right]^2} + o(\eta). \tag{A.74}$$

Changing ϕ^- affects $\Delta(\pi)$ for $\pi \in (\frac{1}{2}, \hat{\pi}]$. For small η, the dominant term comes from considering values of π in $(\frac{1}{2}, \hat{\pi}-\eta]$ only. This yields

$$\begin{aligned}\left.\frac{\partial U}{\partial \phi^-}\right|_{\phi^-=0} =& -2\rho(1-\lambda)e^{\frac{\rho}{2}\Delta(\bar{\pi})}\left\{\int_{\frac{1}{2}}^{\hat{\pi}} \sqrt{\pi(1-\pi)}\exp\left[-\frac{\rho}{\bar{\rho}}\int_{\pi}^{\hat{\pi}} H(\Delta(\pi'),\pi')d\pi'\right]\right.\\ &\times \frac{\left[(1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}-\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)}\right]\left[\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)}+(1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}\right]^{\frac{\rho}{\bar{\rho}}-1}}{\left[\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}+(1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\right]^{\frac{\rho}{\bar{\rho}}}} h(\pi)d\pi\\ &-\int_{\frac{1}{2}}^{\hat{\pi}} \frac{2}{\left[\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)}+(1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}\right]^2}\int_{\frac{1}{2}}^{\pi}\sqrt{\pi'(1-\pi')}\exp\left[-\frac{\rho}{\bar{\rho}}\int_{\pi'}^{\hat{\pi}} H(\Delta(\pi''),\pi'')d\pi''\right]\\ &\left.\times\left[\frac{\pi' e^{-\frac{\bar{\rho}}{2}\Delta(\pi')}+(1-\pi')e^{\frac{\bar{\rho}}{2}\Delta(\pi')}}{\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}+(1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}}\right]^{\frac{\rho}{\bar{\rho}}} h(\pi')d\pi' d\pi\right\} + o(\eta)\end{aligned} \tag{A.75}$$

and

$$\begin{aligned}\left.\frac{\partial \mathcal{L}}{\partial \phi^-}\right|_{\phi^-=0} =& -2\bar{\rho}e^{\bar{\rho}K}\int_{\frac{1}{2}}^{\hat{\pi}} \frac{1}{\left[\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)}+(1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}\right]^2}\left[\int_{\pi}^{\hat{\pi}}\exp\left[-\int_{\frac{1}{2}}^{\pi'} H(\Delta(\pi''),\pi'')d\pi''\right]h(\pi')d\pi'\right.\\ &\left.+\exp\left[-\int_{\frac{1}{2}}^{\hat{\pi}} H(\Delta(\pi'),\pi')d\pi'\right]\frac{e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\int_{\hat{\pi}}^{\bar{\pi}}\pi h(\pi)d\pi + e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\int_{\hat{\pi}}^{\bar{\pi}}(1-\pi)h(\pi)d\pi}{\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}+(1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}}\right] d\pi + o(\eta).\end{aligned} \tag{A.76}$$

The overall change does not affect the largest-order term in $\mathcal{L}$ if

$$\left.\frac{\partial \mathcal{L}}{\partial \phi^+}\right|_{\phi^+=0}\phi^+ + \left.\frac{\partial \mathcal{L}}{\partial \phi^-}\right|_{\phi^-=0}\phi^- = 0, \tag{A.77}$$

and raises U if

$$\left.\frac{\partial U}{\partial \phi^+}\right|_{\phi^+=0}\phi^+ + \left.\frac{\partial U}{\partial \phi^-}\right|_{\phi^-=0}\phi^- > 0. \tag{A.78}$$

Combining (A.77) and (A.78), and noting from (A.74) and (A.76) that $\left.\frac{\partial \mathcal{L}}{\partial \phi^+}\right|_{\phi^+=0} > 0$ and $\left.\frac{\partial \mathcal{L}}{\partial \phi^-}\right|_{\phi^-=0} <$

0, we find that utility increases if

$$\frac{\left.\frac{\partial U}{\partial \phi^+}\right|_{\phi^+=0}}{\left.\frac{\partial \mathcal{L}}{\partial \phi^+}\right|_{\phi^+=0}} > \frac{\left.\frac{\partial U}{\partial \phi^-}\right|_{\phi^-=0}}{\left.\frac{\partial \mathcal{L}}{\partial \phi^-}\right|_{\phi^-=0}}. \tag{A.79}$$

Equations (A.73), (A.74), (A.75) and (A.76) imply that (A.79) holds under the sufficient condition

$$\frac{\int_{\hat{\pi}}^{\pi^*} \sqrt{\pi(1-\pi)} h(\pi) d\pi + \sqrt{\pi^*(1-\pi^*)} \frac{\int_{\pi^*}^{\bar{\pi}} \left(\pi - \frac{1}{2}\right) h(\pi) d\pi}{\pi^* - \frac{1}{2}}}{\int_{\hat{\pi}}^{\bar{\pi}} (\pi - \hat{\pi}) h(\pi) d\pi}$$
$$> \frac{\int_{\frac{1}{2}}^{\hat{\pi}} \sqrt{\pi(1-\pi)} \exp\left[-\frac{\rho}{\bar{\rho}} \int_{\pi}^{\hat{\pi}} H(\Delta(\pi'), \pi') d\pi'\right] \frac{\left[(1-\pi) e^{\frac{\bar{\rho}}{2}\Delta(\pi)} - \pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)}\right] \left[\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} + (1-\pi) e^{\frac{\bar{\rho}}{2}\Delta(\pi)}\right]^{\frac{\rho}{\bar{\rho}}-1}}{\left[\hat{\pi} e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})} + (1-\hat{\pi}) e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\right]^{\frac{\rho}{\bar{\rho}}+1}} h(\pi) d\pi}{\left[e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})} \int_{\hat{\pi}}^{\bar{\pi}} \pi h(\pi) d\pi + e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})} \int_{\hat{\pi}}^{\bar{\pi}} (1-\pi) h(\pi) d\pi\right] \int_{\frac{1}{2}}^{\hat{\pi}} \frac{d\pi}{\left[\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} + (1-\pi) e^{\frac{\bar{\rho}}{2}\Delta(\pi)}\right]^2}}. \tag{A.80}$$

If (A.31) holds as a strict inequality in the other direction for some $\pi \in (\frac{1}{2}, \hat{\pi})$, then (i) (A.21) implies that $\Gamma(\pi) > \Gamma(\hat{\pi})$ and (ii) the same argument as in Lemma 4.3 implies that $\Delta(\pi)$ exceeds the lower bound in (4.12) and hence exceeds $\Delta^*(\bar{\pi})$. Equations $\Delta(\hat{\pi}) > \Delta(\pi) > \Delta^*(\bar{\pi}) > \Delta^*(\hat{\pi})$ and $\Gamma(\pi) > \Gamma(\hat{\pi})$ yield a contradiction since they imply that the risk-averse type $\hat{\pi}$ cannot be induced to accept $(\Delta(\hat{\pi}), \Gamma(\hat{\pi}))$ over $(\Delta(\pi), \Gamma(\pi))$. Hence (A.31) holds, and (A.80) holds under the sufficient condition

$$\frac{\int_{\hat{\pi}}^{\pi^*} \sqrt{\pi(1-\pi)} h(\pi) d\pi + \sqrt{\pi^*(1-\pi^*)} \frac{\int_{\pi^*}^{\bar{\pi}} \left(\pi - \frac{1}{2}\right) h(\pi) d\pi}{\pi^* - \frac{1}{2}}}{\int_{\hat{\pi}}^{\bar{\pi}} (\pi - \hat{\pi}) h(\pi) d\pi}$$
$$> \frac{\int_{\frac{1}{2}}^{\hat{\pi}} \frac{\sqrt{\pi(1-\pi)} \left[(1-\pi) e^{\frac{\bar{\rho}}{2}\Delta(\pi)} - \pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)}\right] h(\pi) d\pi}{\left[\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} + (1-\pi) e^{\frac{\bar{\rho}}{2}\Delta(\pi)}\right] \left[\hat{\pi} e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})} + (1-\hat{\pi}) e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\right]}}{\left[e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})} \int_{\hat{\pi}}^{\bar{\pi}} \pi h(\pi) d\pi + e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})} \int_{\hat{\pi}}^{\bar{\pi}} (1-\pi) h(\pi) d\pi\right] \int_{\frac{1}{2}}^{\hat{\pi}} \frac{d\pi}{\left[\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} + (1-\pi) e^{\frac{\bar{\rho}}{2}\Delta(\pi)}\right]^2}}. \tag{A.81}$$

Since the function

$$M(\pi^*) \equiv \int_{\hat{\pi}}^{\pi^*} \sqrt{\pi(1-\pi)} h(\pi) d\pi + \sqrt{\pi^*(1-\pi^*)} \frac{\int_{\pi^*}^{\bar{\pi}} \left(\pi - \frac{1}{2}\right) h(\pi) d\pi}{\pi^* - \frac{1}{2}}$$

is decreasing, as can be seen from

$$M'(\pi^*) = \frac{d}{d\pi^*} \left(\frac{\sqrt{\pi^*(1-\pi^*)}}{\pi^* - \frac{1}{2}} \right) \int_{\pi^*}^{\bar{\pi}} \left(\pi - \frac{1}{2}\right) h(\pi) d\pi < 0,$$

(A.81) holds under the sufficient condition

$$\frac{\int_{\hat{\pi}}^{\bar{\pi}} \sqrt{\pi(1-\pi)}h(\pi)d\pi}{\int_{\hat{\pi}}^{\bar{\pi}} (\pi-\hat{\pi})h(\pi)d\pi} > \frac{\int_{\frac{1}{2}}^{\hat{\pi}} \frac{\sqrt{\pi(1-\pi)}\left[(1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}-\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)}\right]h(\pi)d\pi}{\left[\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)}+(1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}\right]\left[\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}+(1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\right]}}{\left[e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\int_{\hat{\pi}}^{\bar{\pi}} \pi h(\pi)d\pi + e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\int_{\hat{\pi}}^{\bar{\pi}}(1-\pi)h(\pi)d\pi\right]\int_{\frac{1}{2}}^{\hat{\pi}} \frac{d\pi}{\left[\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)}+(1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}\right]^2}}. \tag{A.82}$$

Since $\Delta^*(\pi) < \Delta(\pi) < \Delta(\bar{\pi})$ for $\pi \in (\frac{1}{2}, \hat{\pi})$,

$$\begin{aligned}
&\frac{\left[\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})} + (1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\right]\left[e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\int_{\hat{\pi}}^{\bar{\pi}} \pi h(\pi)d\pi + e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\int_{\hat{\pi}}^{\bar{\pi}}(1-\pi)h(\pi)d\pi\right]}{\left[\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}\right]^2} \\
&> \frac{\left[\hat{\pi}e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})} + (1-\hat{\pi})e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\right]\left[e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\int_{\hat{\pi}}^{\bar{\pi}} \pi h(\pi)d\pi + e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\int_{\hat{\pi}}^{\bar{\pi}}(1-\pi)h(\pi)d\pi\right]}{\left[\pi e^{-\frac{\bar{\rho}}{2}\Delta(\bar{\pi})} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\bar{\pi})}\right]^2} \\
&> \frac{(1-\hat{\pi})\int_{\hat{\pi}}^{\bar{\pi}}(1-\pi)h(\pi)d\pi}{(1-\pi)^2}.
\end{aligned}$$

Since, in addition,

$$\frac{(1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)} - \pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)}}{\pi e^{-\frac{\bar{\rho}}{2}\Delta(\pi)} + (1-\pi)e^{\frac{\bar{\rho}}{2}\Delta(\pi)}} < 1,$$

(A.82) holds under the sufficient condition

$$\frac{\int_{\hat{\pi}}^{\bar{\pi}} \sqrt{\pi(1-\pi)}h(\pi)d\pi}{\int_{\hat{\pi}}^{\bar{\pi}} (\pi-\hat{\pi})h(\pi)d\pi} > \frac{\int_{\frac{1}{2}}^{\hat{\pi}} \sqrt{\pi(1-\pi)}h(\pi)d\pi}{(1-\hat{\pi})\int_{\hat{\pi}}^{\bar{\pi}}(1-\pi)h(\pi)d\pi \int_{\frac{1}{2}}^{\hat{\pi}} \frac{d\pi}{(1-\pi)^2}}, \tag{A.83}$$

which is equivalent to (A.69). □

Proof of Proposition ??. Lemma A.1 implies that $\Delta(\pi)$ is constant in $(\frac{1}{2}, \bar{\pi}]$ if (A.69) holds for all $\hat{\pi} \in (\frac{1}{2}, \pi^*]$. Making the change of variable $\pi = \frac{1}{2} + \mu d$, we find that when d is small (A.69) holds for all $\hat{\mu} \in (0, \bar{\mu}]$ if

$$\begin{aligned}
&\frac{\int_{\hat{\mu}}^{\bar{\mu}} \frac{1}{2}\hat{h}(\mu)d\mu}{\int_{\hat{\mu}}^{\bar{\mu}}(\mu-\hat{\mu})\hat{h}(\mu)d\mu} > \frac{\int_0^{\hat{\mu}} \frac{1}{2}\hat{h}(\mu)d\mu}{2\hat{\mu}\int_{\hat{\mu}}^{\bar{\mu}} \frac{1}{2}\hat{h}(\mu)d\mu} \\
&\Leftrightarrow \left[\int_{\hat{\mu}}^{\bar{\mu}} \hat{h}(\mu)d\mu\right]^2 > \int_{\hat{\mu}}^{\bar{\mu}}(\mu-\hat{\mu})\hat{h}(\mu)d\mu \frac{\int_0^{\hat{\mu}} \hat{h}(\mu)d\mu}{\hat{\mu}}.
\end{aligned} \tag{A.84}$$

To show that (A.84) holds, we first show that $\hat{h}(\mu)$ increasing implies

$$\frac{\int_{\hat{\mu}}^{\bar{\mu}} \hat{h}(\mu)d\mu}{\bar{\mu}-\hat{\mu}} \geq \frac{\int_0^{\hat{\mu}} \hat{h}(\mu)d\mu}{\hat{\mu}}. \tag{A.85}$$

Equation (A.85) is equivalent to the function

$$N(\hat{\mu}) \equiv \hat{\mu} \int_{\hat{\mu}}^{\bar{\mu}} \hat{h}(\mu)d\mu - (\bar{\mu}-\hat{\mu}) \int_0^{\hat{\mu}} \hat{h}(\mu)d\mu$$

being positive. The derivative of $N(\hat{\mu})$ is

$$N'(\hat{\mu}) = \int_{\hat{\mu}}^{\bar{\mu}} \hat{h}(\mu)d\mu + \int_0^{\hat{\mu}} \hat{h}(\mu)d\mu - \bar{\mu}\hat{h}(\hat{\mu}) = \int_0^{\bar{\mu}} \hat{h}(\mu)d\mu - \bar{\mu}\hat{h}(\hat{\mu}) = \frac{1}{2} - \bar{\mu}\hat{h}(\hat{\mu}).$$

If $\hat{h}(\mu)$ is increasing, then $N'(\hat{\mu})$ is decreasing and hence $N(\hat{\mu})$ is concave. Since $N(\hat{\mu})$ is zero at $\hat{\mu}=0$ and $\hat{\mu}=\bar{\mu}$, it is non-negative.

Equation (A.85) implies that (A.84) holds under the sufficient condition

$$\left[\int_{\hat{\mu}}^{\bar{\mu}} \hat{h}(\mu)d\mu\right]^2 > \int_{\hat{\mu}}^{\bar{\mu}} (\mu-\hat{\mu})\hat{h}(\mu)d\mu \frac{\int_{\hat{\mu}}^{\bar{\mu}} \hat{h}(\mu)d\mu}{\bar{\mu}-\hat{\mu}}$$
$$\Leftrightarrow \int_{\hat{\mu}}^{\bar{\mu}} \hat{h}(\mu)d\mu > \int_{\hat{\mu}}^{\bar{\mu}} \frac{\mu-\hat{\mu}}{\bar{\mu}-\hat{\mu}}\hat{h}(\mu)d\mu,$$

which holds. Hence, $\Delta(\pi)$ is constant in $(\frac{1}{2},\bar{\pi}]$, and so is $\Gamma(\pi)$ because of the (IC) constraint (4.2).

Since $\Delta(\pi)$ is constant in $(\frac{1}{2},\bar{\pi}]$, it is equal to the lower bound in (4.12). Using this property and setting $\pi=\frac{1}{2}+\mu d$, $K=\frac{kd}{\bar{\rho}}$, and $\bar{\rho}=N\rho$ in (4.12), we find (4.20). Equation (A.18) applied to the right-limits of $(\Delta(\pi),\Gamma(\pi))$ at $\frac{1}{2}$, and combined with $\Delta(\frac{1}{2})=0$, $\Delta(\frac{1}{2}^+)=\Delta(\bar{\pi})$ and $\Gamma(\frac{1}{2}^+)=\frac{\Delta(\bar{\pi})}{2}$, yields (4.21).

The optimal fee in the proposition includes the optimal values of $f(z(\pi)d)$ and $f(-z(\pi)d)$ for $\pi\in[\frac{1}{2},\bar{\pi}]$, and of $f(\hat{z}d)$ and $f(-\hat{z}d)$. Since $f(W)$ is equal to zero for $W<-\frac{\Delta(\bar{\pi})}{2}$ and to a positive value for $W>\frac{\Delta(\bar{\pi})}{2}$, the risk-averse types $\pi\in(\frac{1}{2},\bar{\pi}]$ are indifferent between any position $z>\frac{\Delta(\bar{\pi})}{2d}$. Since (4.21) ensures that the risk-averse type $\frac{1}{2}$ is indifferent between the zero position and any position $z>\frac{\Delta(\bar{\pi})}{2d}$, the risk-averse types $\pi\in(\frac{1}{2},\bar{\pi}]$ prefer any position $z>\frac{\Delta(\bar{\pi})}{2d}$ to the zero position, and so does the risk-neutral type. Therefore, the position $z(\pi)$ is (weakly) optimal for the risk-averse type π, and the position $\hat{z}$ is (weakly) optimal for the risk-neutral type.

We finally derive the employment condition. Equation (4.19) defines $\mu^*\in(0,\bar{\mu})$ uniquely because the left-hand side is decreasing in μ^*, is positive for $\mu^*=0$, and is negative for $\mu^*=\bar{\mu}$. Setting $\pi=\frac{1}{2}+\mu d$ in (4.19), we find $\pi^*=\frac{1}{2}+\mu^* d$. Setting $\Gamma(\pi)=\frac{\Delta(\bar{\pi})}{2}$ for all $\pi\in(\frac{1}{2},\bar{\pi}]$, we can write (4.14) as

$$(1-\lambda)e^{\frac{\rho\Delta(\bar{\pi})}{2}}\left[\int_{\frac{1}{2}}^{\pi^*}\sqrt{\pi(1-\pi)}h(\pi)d\pi+\sqrt{\pi^*(1-\pi^*)}\frac{\int_{\pi^*}^{\bar{\pi}}\left(\pi-\frac{1}{2}\right)h(\pi)d\pi}{\pi^*-\frac{1}{2}}\right]<\frac{1}{4}$$

$$\Leftrightarrow (1-\lambda)e^{\frac{\rho\Delta(\bar{\pi})}{2}}\left[\int_0^{\mu^*}\sqrt{\frac{1}{4}-\mu^2d^2}\hat{h}(\mu)d\mu+\sqrt{\frac{1}{4}-(\mu^*)^2d^2}\frac{\int_{\mu^*}^{\bar{\mu}}\mu\hat{h}(\mu)d\mu}{\mu^*}\right]<\frac{1}{4}$$

$$\Leftrightarrow (1-\lambda)e^{\frac{\rho\Delta(\bar{\pi})}{2}}\left[\int_0^{\mu^*}\sqrt{\frac{1}{4}-\mu^2d^2}\hat{h}(\mu)d\mu+\sqrt{\frac{1}{4}-(\mu^*)^2d^2}\left(\int_{\mu^*}^{\bar{\mu}}\hat{h}(\mu)d\mu+\frac{\lambda}{2(1-\lambda)}\right)\right]<\frac{1}{4}$$

$$\Leftrightarrow (1-\lambda)\left[1+\frac{1}{2N}\log\left(\frac{4\int_0^{\bar{\mu}}\hat{h}(\mu)d\mu+k}{4\int_0^{\bar{\mu}}\hat{h}(\mu)d\mu-k}\right)+o\left(\frac{1}{N}\right)\right]$$
$$\times\left[\frac{1}{4(1-\lambda)}-\int_0^{\mu^*}\mu^2d^2\hat{h}(\mu)d\mu-(\mu^*)^2d^2\left(\int_{\mu^*}^{\bar{\mu}}\hat{h}(\mu)d\mu+\frac{\lambda}{2(1-\lambda)}\right)+o(d^4)\right]<\frac{1}{4}, \qquad \text{(A.86)}$$

where the second step follows by making the change of variable $\pi=\frac{1}{2}+\mu d$ and using $\pi^*=\frac{1}{2}+\mu^* d$, the third step follows from (4.19), and the fourth step follows from (4.20). Rearranging (A.86), we can write it as

$$\frac{1}{2N}\log\left(\frac{4\int_0^{\bar{\mu}}\hat{h}(\mu)d\mu+k}{4\int_0^{\bar{\mu}}\hat{h}(\mu)d\mu-k}\right)$$
$$-4(1-\lambda)d^2\left[\int_0^{\mu^*}\mu^2\hat{h}(\mu)d\mu+(\mu^*)^2\left(\int_{\mu^*}^{\bar{\mu}}\hat{h}(\mu)d\mu+\frac{\lambda}{2(1-\lambda)}\right)\right]+o\left(\frac{1}{N}\right)+o(d^2)<0$$
$$\Leftrightarrow \frac{1}{2N}\log\left(\frac{4\int_0^{\bar{\mu}}\hat{h}(\mu)d\mu+k}{4\int_0^{\bar{\mu}}\hat{h}(\mu)d\mu-k}\right)$$
$$-4(1-\lambda)d^2\left[\int_0^{\mu^*}\mu^2\hat{h}(\mu)d\mu+\mu^*\int_{\mu^*}^{\bar{\mu}}\mu\hat{h}(\mu)d\mu\right]+o\left(\frac{1}{N}\right)+o(d^2)<0, \qquad \text{(A.87)}$$

where the second step follows from (4.19). If $\frac{1}{N}$ is of order smaller than d^2, then the dominant term in the left-hand side of (A.87) is negative, and hence that equation is satisfied. If $\frac{1}{N}$ is of order larger than d^2, then the dominant term in the left-hand side of (A.87) is positive, and hence that equation is not satisfied. If $\frac{1}{N}$ and d^2 are of the same order, then requiring that the dominant term in (A.87) is negative if equivalent to (4.18). □

References

Buffa, Andrea, Dimitri Vayanos, and Paul Woolley, 2019, Asset management contracts and equilibrium prices, working paper London School of Economics.