Online Appendix for “Centralized versus Decentralized Delegated Portfolio Management under Moral Hazard” (Leung 2015)

This Online Appendix continues its section numbering from the Appendix of the main text.

C Principle of Dynamic Programming

As it is well known in the literature of extending mean-variance analysis to multiple periods, the mean-variance utility over terminal wealth is not directly amenable to a recursive form for dynamic programming. Here, however, we will take effectively the discrete-time approach as motivated by Basak and Chabakauri (2010).

Consider time periods $t = 0, 1, \ldots, T$. Consider the $t = 0$ mean-variance utility over a time $t = T$ random variable $W_T$, 

$$U_0 := \mathbb{E}_0[W_T] - \frac{\eta}{2} \text{Var}_0(W_T).$$

Also let us define the time $t$ continuation utility,

$$U_t := \mathbb{E}_t[W_T] - \frac{\eta}{2} \text{Var}_t(W_T), \quad \text{for } t = 0, 1, \ldots, T,$$

and note that $U_T = W_T$. The law of total variance states that,

$$\text{Var}_t(W_T) = \mathbb{E}_t[\text{Var}_{t+1}(W_T)] + \text{Var}_t(\mathbb{E}_{t+1}[W_T]), \quad \text{for } t = 0, 1, \ldots, T - 1. \quad (C.2)$$

Substituting (C.2) into (C.1), and using the law of iterated expectations, we observe that,

$$U_t = \mathbb{E}_t[W_T] - \frac{\eta}{2} \left( \mathbb{E}_t[\text{Var}_{t+1}(W_T)] + \text{Var}_t(\mathbb{E}_{t+1}[W_T]) \right)$$

$$= \mathbb{E}_t[W_T] - \frac{\eta}{2} \mathbb{E}_t[\text{Var}_{t+1}(W_T)] - \frac{\eta}{2} \text{Var}_t(\mathbb{E}_{t+1}[W_T])$$

$$= \mathbb{E}_t \left[ \mathbb{E}_{t+1}[W_T] - \frac{\eta}{2} \text{Var}_{t+1}(W_T) \right] - \frac{\eta}{2} \text{Var}_t(\mathbb{E}_{t+1}[W_T])$$

$$= \mathbb{E}_t[U_{t+1}] - \frac{\eta}{2} \text{Var}_t(\mathbb{E}_{t+1}[W_T]).$$

Thus, the equations that gives the backward recursive relationship are,

$$U_T = W_T,$$

$$U_t = \mathbb{E}_t[U_{t+1}] - \frac{\eta}{2} \text{Var}_t(\mathbb{E}_{t+1}[W_T]), \quad t = 0, 1, \ldots, T - 1. \quad (C.3)$$

In the actual application considered in this paper, we will use $T = 2$ and so we will consider time periods $t = 0, 1, 2$.

D Dynamic Centralized Delegation

In the dynamic centralized delegation model, the Principal will offer to Manager $C$ a long term contract, consisting of a fixed fee $x_C \in \mathbb{R}$ paid at $t = 2$, and multi-period performance fees $(y_C, y_{C,0}, y_{C,1}) \in \times [0, 1]$.
over, respectively, the ending period wealth $W_{C;1}$ and $W_{C;2}$. Manager $C$ will choose to either accept or reject the contract. If Manager $C$ accepts and commits to the contract, then at $t = 0$ the Principal gives his initial wealth of $1$ to the single centralized Manager $C$ who has $0$ initial wealth. Then Manager $C$ will commit to a long term strategy $(\theta, \tau) \in S$. Subsequently, Manager $C$ will choose portfolio weights $(1 - \psi_0, \psi_0)$ into the strategies with a period return $(R_{\theta;1}, R_{\tau;1})$. At the end of period $t = 1$, the Principal will pay $y_{C;0}R_{(\theta,\tau);1}$ to Manager $C$, where $R_{(\theta,\tau)}$ is the managed portfolio of Manager $C$. Thus, after fees, the $t = 1$ wealth of the Principal is $W_{C;1}$, and the $t = 1$ wealth of Manager $C$ is $W_{C;1}$. For simplicity, with the available wealth of $W_{C;1}$, Manager $C$ will make no further investments. At $t = 1$, the Principal will reinvest with Manager $C$ again, and so Manager $C$ will choose portfolio weights $(1 - \psi_1, \psi_1)$ into the strategies with returns $(R_{\theta;2}, R_{\tau;2})$. At the end of the period, Manager $C$ will be paid $W_{C;1}y_{C;1}R_{(\theta,\tau);2}$.

In all, the optimization problem for centralized delegation is as follows. Please see Figure 14 for a timeline.

$$
\begin{align*}
  \text{Principal offers a linear contract } & x_C, (y_{C;0}, y_{C;1}) \in \mathbb{R} \times [0,1]^2 \\
  \text{to the Manager } & t = 0 \\
  \text{Manager C makes investment strategy choices } & (\theta, \tau) \in S \\
  \text{Principal receives managed portfolio returns } & R_{(\theta,\tau);1}, \\
  & \text{pays Manager } C, \\
  & \text{and has wealth } W_{C;1} \\
  \text{Manager C receives payoff } & y_{C;0}R_{(\theta,\tau);1} \\
  & t = 1 \\
  \text{Manager C chooses portfolio weights } & 1 - \psi_1 \in \mathbb{R} \text{ into } R_{\theta;2}, \\
  & \text{and } \psi_1 \text{ into } R_{\tau;2} \\
  \text{Principal receives managed portfolio returns } & R_{(\theta,\tau);2}, \\
  & \text{pays Manager } C, \\
  & \text{and has wealth } W_{C;2} \\
  \text{Manager C receives payoff } & -(c(\theta) + c(\tau)) + x_C + W_{C;1}y_{C;1}R_{(\theta,\tau);1} \\
  & t = 2 \\
\end{align*}
$$

**Figure 14:** Dynamic centralized delegation time line.

$C$ at $t = 1$. This is clearly very possible, but this is not the raison d’être of extending our discussion from a simple one period model to a dynamic two period model.
The objective function of (DynCen), the Principal wants to maximize the t = 2 terminal wealth, by choosing the optimal t = 0 and t = 1 performance fees yC:=0 and yC:=1, respectively, and also the optimal fixed fee xC. Again, recalling from Assumption 6.1, we assume the Principal only wants to implement the strategy pair (θH, τH). Here, (D.1a) and (D.1f) are the Principal’s t = 1 and t = 2 budget constraints; here we assume that there is no intermediate consumption and the Principal will reinvest all the t = 1 payoffs back into Manager C. Given any contract, Manager C will choose the optimal t = 0 and t = 1 portfolios, in accordance to the budget constraints (D.1h) and (D.1j), and the portfolio choice problem (D.1g), which results in the portfolio return (D.1i). The budget constraints for Manager C, after substituting in the optimal portfolio choices, are (D.1i) and (D.1j); we assume that Manager C will not reinvest his t = 1 wealth. Since Manager C has zero initial wealth, (D.1i) is his individual rationality constraint. And due to moral hazard, (D.1j) is the incentive compatibility constraint for inducing Manager C to choose the Principal’s strict preference for the strategy pairs (θH, τH).

E Dynamic Centralized Delegation in First Best

For the first best centralized delegation case, consider problem (DynCen) without the incentive compatibility constraint (D.1i).

Proposition E.1. Consider the first best centralized delegation problem (DynCen) but without the incentive compatibility constraints (D.1i). Fix any investment strategy pair (θ, τ) ∈ S. Assume Assumption 6.2.

(a) Fix any arbitrary contract (xC, {yC:=0, yC:=1}) and let’s consider the optimal portfolio policy as chosen by Manager C.

(i) Suppose the realized wealth of the Principal at t = 1 is WcP:=1(θ, τ) = WcP:=1(θ, τ) . Then the t = 1 optimal portfolio is,

\[ \hat{\psi}_{1,(\theta, \tau)} := \frac{E_1 R_{\theta,2} - E_1 R_{\theta,2} + \eta_M Var_1 (R_{\theta,2}) - Cov_1 (R_{\theta,2}, R_{\tau,2}) yC:=1 w^0_{cP:=1} (\theta, \tau)}{\eta_M Var_1 (R_{\theta,2} - R_{\tau,2}) yC:=1 w^0_{cP:=1}} \]  

(E.1)
(ii) The \( t = 0 \) optimal portfolio policy chosen by Manager \( C \) is the solution to the following optimization problem. Suppose the realized \( t = 1 \) wealth for the Principal and Manager \( C \) are, respectively, \( W_{cP,1}^{(\theta,\tau)} = w_{cP,1}^{(\theta,\tau)} \), \( W_{C,1}^{(\theta,\tau)} = u_{C,1}^{(\theta,\tau)} \). Define,

\[
\hat{U}_{C,1}^{(\theta,\tau)} := \hat{U}_{C,1}^{(\theta,\tau)}(y_{C,0}, y_{C,1}) = \sup_{\psi \in \mathbb{R}} \mathbb{E}[\hat{W}_{C,2}^{(\theta,\tau)}] - \frac{\eta_M}{2} \text{Var}(\hat{W}_{C,2}^{(\theta,\tau)})
\]

\[
= \hat{u}_{C,1}^{(\theta,\tau)} + \frac{(\mathbb{E}[R_{\theta,2} - \mathbb{E}[R_{\tau,2}])^2}{2\eta_M \text{Var}(R_{\theta,2} - R_{\tau,2})}
\]

\[
+ \frac{(\mathbb{E}[R_{\theta,2} - \mathbb{E}[R_{\tau,2}])^2(\text{Var}(R_{\theta,2}) - \text{Cov}(R_{\theta,2}, R_{\tau,2}))}{\text{Var}(R_{\theta,2} - R_{\tau,2})}
\]

\[
\text{Var}_0(\mathbb{E}[\hat{W}_{C,2}^{(\theta,\tau)}]) = \frac{1}{(\eta_P + \eta_M)^2 \text{Var}(R_{\theta,2} - R_{\tau,2})}
\]

\[
\times (\text{Var}(R_{\theta,2})(1 - \psi_0)^2 + \text{Var}(R_{\tau,2})\psi_0^2 + 2\psi_0(1 - \psi_0)\text{Cov}(R_{\theta,2}, R_{\tau,2}))
\]

(E.2)

And also consider, for any arbitrary \( t = 0 \) portfolio \( \psi_0 \in \mathbb{R} \), the \( t = 0 \) variance of \( \mathbb{E}[\hat{W}_{C,2}^{(\theta,\tau)}] \) is,

\[
\text{Var}_0(\mathbb{E}[\hat{W}_{C,2}^{(\theta,\tau)}]) = \frac{1}{(\eta_P + \eta_M)^2 \text{Var}(R_{\theta,2} - R_{\tau,2})}
\]

\[
\times (\text{Var}(R_{\theta,2})(1 - \psi_0)^2 + \text{Var}(R_{\tau,2})\psi_0^2 + 2\psi_0(1 - \psi_0)\text{Cov}(R_{\theta,2}, R_{\tau,2}))
\]

(E.3)

The \( t = 0 \) optimal portfolio policy \( \hat{\psi}_0 = \hat{\psi}_0(y_{C,0}) \), as a function of an arbitrary \( t = 0 \) performance fee \( y_{C,0} \), chosen by Manager \( C \) is solves the following quadratic concave objective in \( \psi_0 \in \mathbb{R} \),

\[
\hat{U}_{C,1} = \sup_{\psi_0 \in \mathbb{R}} \mathbb{E}[\hat{U}_{C,1}] - \frac{\eta_M}{2} \text{Var}_0(\mathbb{E}[\hat{W}_{C,2}]).
\]

(E.4)

(b) Let’s consider the optimal contract offered by the Principal.

(i) For any performance fees \( y_{C,0}, y_{C,1} \in [0, 1] \), the optimal fixed fee is,

\[
\hat{\psi}_{C,0} = \hat{\psi}_{C,0}(y_{C,0}) = (c(\theta) + c(\tau)) - \mathbb{E}[W_{C,2}^{(\theta,\tau)}] + \frac{\eta_M}{2} \text{Var}(W_{C,2}^{(\theta,\tau)}).
\]

(E.5)

(ii) The \( t = 1 \) optimal performance fee is,

\[
\hat{y}_{C,1}^{FB} = \frac{\eta_P}{\eta_P + \eta_M}.
\]

(E.6)

(iii) The \( t = 0 \) optimal performance fee is given by the following. Suppose also that the \( t = 1 \) realized value of the Principal’s wealth is \( u_{C,1}^{(\theta,\tau)} \). Firstly, define the portfolio return variances \( \hat{\Sigma}_{1, (\theta, \tau)} \), and recall \( \hat{\psi}_1 = \hat{\psi}_1(y_{C,1}^{FB}) \),

\[
\hat{\Sigma}_{1,(\theta,\tau)} = (1 - \hat{\psi}_1)^2 \text{Var}(R_{\theta,2}) + \hat{\psi}_1^2 \text{Var}(R_{\tau,2}) + 2\hat{\psi}_1(1 - \hat{\psi}_1)\text{Cov}(R_{\theta,2}, R_{\tau,2}).
\]

(E.7)

Then define also the \( t = 1 \) continuation utility value for the Principal,

\[
\hat{U}_{C,1}^{(\theta,\tau)} = (1 + (1 - \hat{y}_{C,1})(1 - \hat{\psi}_1)\mathbb{E}[E_{\theta,2} + \hat{\psi}_1 E_{\tau,2}]) - \frac{\eta_P}{2} (1 - \hat{y}_{C,1}^{FB})^2 \hat{\Sigma}_{1,(\theta,\tau)}
\]

(E.8)
The $t=0$ optimal fee $\hat{y}_{C;0}$ is the solution to the problem,

$$
\hat{U}_{C;0} = \sup_{y_{C;0} \in [0,1]} \mathbb{E}_0[\hat{U}_{C;1}(t) - \frac{\eta_C}{2} \text{Var}_0(\mathbb{E}_1 W_{C;1}^{(t,\tau)}) - \frac{\eta_M}{2} \text{Var}_0(\mathbb{E}_1 W_{C;2}^{(t,\tau)})],
$$

(E.9)

where $W_{C;1}^{(t,\tau)}$ is Manager $C$’s time $t$ wealth as per (E.6), (E.10) after choosing the optimal portfolios, and $\hat{U}_{C;1}(t)$ is Manager $C$’s $t=1$ continuation utility value (E.3) after substituting in the optimal $t=1$ performance fee $\hat{y}_{C;1}^{FB}$.

Let’s begin by discussing the optimal portfolio policies (E.11) at $t=1$ of the Manager $C$ under first best dynamic centralized delegation as per Proposition E.1. At $t=1$, the optimal portfolio choice form by Manager $C$ is nearly identical to that of a static mean-variance optimizer, and indeed, the solution form is fairly similar to the static model in Proposition E.1. The beginning wealth at $t=1$ for Manager $C$ is precisely $y_{C;1} w_{C;1}^{(t,\tau)}$. Based on this level of wealth at $t=1$ and risk aversion $\eta_M$, Manager $C$ constructs the optimal portfolio to maximize the one period ahead mean-variance of terminal wealth. Anticipating this, the Principal simply offers a performance contract $\hat{y}_{C;1}^{FB}$ at $t=1$ of (E.0) to optimally risk share with Manager $C$ in accordance to their respective risk aversions $\eta_C$ and $\eta_M$. In all, the $t=1$ policies are essentially akin to that of the static centralized delegation model of Proposition E.1.

The optimal portfolio policy (E.12) at $t=0$ of Manager $C$ is slightly more nuanced. At $t=0$, for any given $t=0$ performance fee $y_{C;0}$, Manager $C$’s hedging motive of future income will matter. Namely, Manager $C$ does not choose portfolios $\psi_0 = \psi_0(y_{C;0})$ at $t=0$ such that:

- Maximizes the $t=0$ expectation of the $t=1$ continuation value $\hat{U}_{C;1}$. In particular, the continuation value $\hat{U}_{C;1}$ is equal to:
  - Manager $C$’s next period (final period) wealth $W_{C;1}^{(t,\tau)}$;
  - Constant term relating to the benefits of executing a long-short strategy;
  - The $t=0$ portfolio choice effects on the $t=2$ performance fees for Manager $C$. Such $t=2$ performance fees depend on the level $t=1$ wealth $W_{C;1}^{(t,\tau)}$ of the Principal, which of course, depends on Manager $C$’s $t=0$ portfolio choice; and
  - Given that Manager $C$ is risk averse, there is also a contract volatility term $(W_{C;1}^{(t,\tau)})^2$ in desiring lower volatility in the $t=2$ performance fees.

- Intertemporal hedging motive $\text{Var}_0(\mathbb{E}_1 W_{C;1}^{(t,\tau)})$ relating to the $t=0$ variance of the next period $t=1$ expectation on the terminal period $t=2$ wealth.

Once the optimal portfolio policies $\hat{\psi}_0$ of Manager $C$ has been determined, the Principal’s optimal performance fee choice procedure is as follows. At $t=1$, the Principal simply wants to optimally risk share based on the Principal and Manager $C$’s risk aversion parameter and offers the performance fee $\hat{y}_{C;1}^{FB}$ of (E.0). However, at $t=0$, the optimal performance fee is chosen, again, to optimally risk share with Manager $C$ but now taking into account the intertemporal hedging motive of both the Principal himself and also that of Manager $C$:

- The Principal wants to choose small performance fees to maximize his next period $t=1$ expected wealth and minimize his $t=1$ wealth volatility: $\mathbb{E}_0[\hat{U}_{C;1}(t)] = \mathbb{E}_0[W_{C;1}^{(t,\tau)}] - \frac{\eta_C}{2} \text{Var}_0(W_{C;1}^{(t,\tau)})$.
- The Principal’s $t=0$ fees affects his $t=1$ wealth, and in turn affects both the amount of resulting wealth in $t=2$, depending on the portfolio policy of Manager $C$ and the then realized returns. Thus, the Principal’s intertemporal hedging motive is to choose performance fees to minimize the terminal wealth volatility $-\frac{\eta_M}{2} \text{Var}_0(\mathbb{E}_1 W_{C;2}^{(t,\tau)})$.
- The Principal has a strong risk sharing motive with Manager $C$ to minimize Manager $C$’s continuation utility $\mathbb{E}_0[U_{C;1}^{(t,\tau)}]$ and an intertemporal incentive motive to minimize Manager $C$’s terminal date $t=2$ wealth volatility $-\frac{\eta_M}{2} \text{Var}_0(\mathbb{E}_1 W_{C;2}^{(t,\tau)})$; note that by the form of the performance fees, lower wealth
volatility \( \text{Var}_0(\mathbb{E}_1 W_{C,2}^{(\theta, \tau)}) \) for Manager \( C \) also implies a lower wealth volatility \( \text{Var}_0(\mathbb{E}_1 W_{c,F,2}^{(\theta, \tau)}) \) for the Principal, and again since all individuals are risk averse, this is beneficial for the Principal.

By now explicitly imposing Assumption \( 6.3 \), we get greater transparency of the solution form.

**Corollary E.2.** Consider again the first best dynamic centralized delegation problem \([\text{DynCen}]\) and Proposition \([\text{Dyn}1]\). In addition to Assumption \( 6.2 \), assume also Assumption \( 6.3 \). Recall that the Principal wants to implement investment strategy pairs \((\theta_H, \tau_H)\).

(a) The optimal portfolio policies chosen by Manager \( C \) are given as follows.

(i) The \( t = 1 \) optimal portfolio chosen by Manager \( C \) is independent of the performance fees \( y_1 \), and is,

\[
\dot{\psi}_{FB}^C = \frac{1}{2}. 
\]

(ii) The \( t = 0 \) optimal portfolio chosen by Manager \( C \) is independent of the performance fees \( y_0 \), and is,

\[
\dot{\psi}_{FB}^C = \frac{1}{2}. 
\]

(b) The optimal contract chosen by the Principal is given as follows.

(i) The optimal fixed fee form is given in \((E.5)\).

(ii) The \( t = 1 \) optimal performance fee chosen by the Principal is,

\[
\dot{y}^{FB}_{C,1} = \frac{\eta_p}{\eta_p + \eta_M}. 
\]

(iii) The \( t = 0 \) optimal (interior solution) performance fee chosen by the Principal is, \( ^{\text{II}} \)

\[
\dot{y}^{FB}_{C,0} = \frac{\dot{y}^N_{C,0}}{\dot{y}^D_{C,0}}, \tag{E.10} 
\]

provided that \( \dot{y}^{FB}_{C,0} \in (0, 1) \), and where,

\[
\dot{y}^N_{C,0} = 2\eta_p(-2\mu^2 + \eta_p\sigma^2) + (-4\mu^2 + 2\mu[1 + 2\mu(1 + \mu)]\sigma^2 + \eta_p\sigma^4)\eta_M \\
+ 2\eta_p\sigma^2(\eta_p + [1 + 2\mu(1 + \mu) + \sigma^2]\eta_M)\rho_{\theta_H,\tau_H} + \eta_p\eta_M\sigma^4\rho_{\theta_H,\tau_H}, \\
\dot{y}^D_{C,0} = \sigma^2(1 + \rho_{\theta_H,\tau_H})[2\eta_p^2 + 2\eta_M^2 + \eta_p\eta_M(4(1 + \mu^2) + \sigma^2(1 + \rho_{\theta_H,\tau_H}))]. 
\]

See Section \([\text{E.4}]\) for numerical illustrations of Corollary \([\text{E.2}]\).

### E.1 Numerical Illustrations of Dynamic Delegation in First Best

Using the analytical solutions for first best dynamic delegation from Corollary \([\text{E.2}]\) and Proposition \([\text{Dyn}1]\), we can easily numerically illustrate the Principal’s value functions under centralization and decentralization, and also their associated optimal policies at \( t = 0 \). It should be noted that from Corollary \([\text{E.2}]\) and Proposition \([\text{Dyn}1]\), the \( t = 1 \) optimal policies for both centralization and decentralization take on explicit and simple forms. Hence, we will focus the numerical illustrations on the \( t = 0 \) optimal policies, for which its comparative statics may not be obvious at first glance. The optimal portfolios at \( t = 0 \) for both centralization and decentralization take on a simple explicit form, and hence are not plotted. In decentralization, the \( t = 0 \) optimal performance fee for both Manager \( A \) and Manager \( B \) are identical, and hence only one of them is plotted. The base parameters are all identical to that of Table \([\text{I}]\).

\[^{31} \text{i.e. “N” for numerator, and “D” for denominator.} \]
Figure 15: Plot of the first best dynamic delegation model against the compliant investment strategy pair’s correlations $\rho_{H_t, \tau H}$, with several scenarios on the means $\mu$ of the compliant investment strategy. Similar to the static first best case of Proposition 4.3, due to optimal risk sharing, higher correlations $\rho_{H_t, \tau H}$ favor decentralization, while lower correlations favor centralization. And also, naturally, higher mean returns $\mu$ will increase the Principal’s $t = 0$ value function. While the first best performance fees in the centralized static model of Proposition 4.1 only consist of the Principal’s and Manager $C$’s risk aversions, this is clearly not the case for the $t = 0$ centralization performance fees due to the intertemporal hedging incentive of all individuals involved.
Figure 16: Plot of the first best dynamic delegation model against the Managers' risk aversion $\eta_M$, with several scenarios on the correlations of the compliant investment strategy pair $\rho_{\theta_i, \nu_i}$. As Managers' risk aversion $\eta_M$ increases, it becomes more expensive to compensate the Managers for taking on the volatility of the contract. This is the same effect as per the first best static delegation model of Section 3.
Figure 17: Plot of the first best dynamic delegation model against the compliant strategies’ mean return $\mu$, with several scenarios of the strategies’ volatilities $\sigma$. Given that all individuals have mean-variance preferences, it is no surprise that the Principal’s $t=0$ value function increases with higher mean $\mu$, and is lower with a higher volatility $\sigma$. However, as compared to the first best static delegation model of Section 4, especially that of Proposition 4.1 for centralization and Proposition 4.2 for decentralization, the optimal static performance fees distinctly do not depend on the mean return $\mu$. In this first best dynamic delegation model, as the mean return $\mu$ of the compliant strategies, the $t=0$ performance fees under both first best dynamic centralization and decentralization decrease. Likewise, in the first best static delegation model, largely thanks to Assumption 3.2, the volatility $\sigma$ of returns also does not enter the static first best performance fees. In contrast, here in first best dynamic delegation, as volatility $\sigma$ increases, the $t=0$ performance fees decreases under both centralization and decentralization.
F Dynamic Centralized Delegation in Second Best

Proposition F.1. Consider the second best centralized delegation problem \((\text{DynCen})\) in its entirety. Assume Assumption A.3 and Assumption B.3. Note and recall that the Principal wants to implement the investment strategy pair \((\theta_H, \tau_H)\).

(a) For any given contract \((x_C, \{y_C, 0, y_C, 1\})\) and any investment strategy pair \((\theta, \tau)\), the optimal portfolio policies \(\hat{\psi}_{t,(\theta, \tau)}\) for \(t = 0, 1\) is equivalent to the form as in Proposition E.3 for first best centralized delegation.

(b) The optimal fixed fee \(\hat{x}_{C, (\theta_H, \tau_H)}\) is equivalent to the form \((E.3)\), when evaluated at \((\theta, \tau) = (\theta_H, \tau_H)\), for any performance fees \((y_0, y_1)\).

(c) Suppose the \(t = 1\) realized wealths are \(w_{cP, 1}^{(\theta, \tau)} = w_{cP, 1}^{(\theta_H, \tau_H)}\) and \(w_{C, 1}^{(\theta, \tau)} = w_{C, 1}^{(\theta_H, \tau_H)}\), for the Principal and Manager C, respectively. Then there exists some vector \(\lambda_C = (\lambda_C^{(\theta_H, \tau_H)}, \lambda_C^{(\theta_H, \tau_H)}, \lambda_C^{(\theta_H, \tau_H)}) \in \mathbb{R}^3\) of which only one of the elements is nonzero and the other two will be zero, and let \(\lambda_C^{(\theta_H, \tau_H)}\) be that nonzero element. Then the optimal \(t = 1\) performance fee is,

\[
\hat{y}_{C, 1}^{\lambda_C} = \hat{y}_{C, 1}^{\lambda_C} (w_{C, 1}, w_{P, 1}) = \frac{w_{C, 1}^{(\theta_H, \tau_H)}(2\lambda_C^{(\theta_H, \tau_H)} + w_{C, 1}^{(\theta_H, \tau_H)})(1 + \rho_{\theta_t, \tau_t}) - \lambda_C^{(\theta_H, \tau_H)}(\mu_{\theta_t} + \mu_{\tau_t})w_{P, 1}^{(\theta_H, \tau_H)}}{\sigma^2 (\eta_{\theta_t} + \eta_{\tau_t}(1 + \lambda_C^{(\theta_H, \tau_H)}))(1 + \rho_{\theta_t, \tau_t}w_{P, 1}^{(\theta_H, \tau_H)})^2 - \lambda_C^{(\theta_H, \tau_H)}\eta_{\theta_t}(w_{P, 1}^{(\theta_H, \tau_H)})^2 (1 + \rho_{\theta_t, \tau_t})}.
\]

(F.1)

(d) The \(t = 1\) continuation value of Manager C taking on investment strategy pair \((\theta, \tau)\) and the Principal are,

\[
\hat{U}_{C, 1}^{(\theta, \tau)} = \hat{U}_{C, 1}^{(\theta, \tau)} |_{y_{C, 1} = \hat{y}_{C, 1}^{\lambda_C}} = u_{C, 1}^{(\theta, \tau)} + \hat{y}_{C, 1}^{\lambda_C} [(1 - \hat{\psi}_{1,(\theta, \tau)})\mu_\theta + \hat{\psi}_{1,(\theta, \tau)}\mu_\tau] - \frac{\eta_{\theta_t}}{2} \left(\hat{y}_{C, 1}^{\lambda_C} u_{C, 1}^{(\theta, \tau)}\right)^2 \Sigma_{1,(\theta, \tau)},
\]

(F.2a)

\[
\hat{U}_{cP, 1}^{(\theta, \tau)} = \hat{U}_{cP, 1}^{(\theta, \tau)} |_{y_{C, 1} = \hat{y}_{C, 1}^{\lambda_C}} = u_{cP, 1}^{(\theta, \tau)} (1 + (1 - \hat{y}_{C, 1}^{\lambda_C})\mu_t - \frac{\eta_{\theta_t}}{2} \left(1 - \hat{y}_{C, 1}^{\lambda_C}\right) u_{C, 1}^{(\theta, \tau)}^2 \Sigma_{1,(\theta, \tau)},
\]

(F.2b)

where we recall that the optimal \(t = 1\) performance fee \(\hat{y}_{C, 1}^{\lambda_C} = \hat{y}_{C, 1}^{\lambda_C} (w_{C, 1}^{(\theta_H, \tau_H)}, w_{cP, 1}^{(\theta_H, \tau_H)}).

(e) The optimal \(t = 0\) performance fee \(\hat{y}_{C, 0}^{\lambda_C} \in [0, 1]\) is the solution to the following optimization problem,

\[
\hat{U}_{cP, 0}^{\lambda_C} = \sup_{y_{C, 0}\in[0,1]} \mathbb{E}_0[\hat{U}_{C, 1}^{(\theta_H, \tau_H)}] - \frac{\eta_{\theta_t}}{2} \text{Var}_0(\mathbb{E}_1 W_{C, 2}^{(\theta_H, \tau_H)})
\]

\[
+ \mathbb{E}_0[\hat{U}_{cP, 1}^{(\theta_H, \tau_H)}] - \frac{\eta_{\theta_t}}{2} \text{Var}_0(W_{C, 2}^{(\theta_H, \tau_H)})
\]

\[
- \lambda_C^{(\theta_H, \tau_H)} \left[\mathbb{E}_0[\hat{U}_{C, 1}^{(\theta_H, \tau_H)}] - \frac{\eta_{\theta_t}}{2} \text{Var}_0(\mathbb{E}_1 W_{C, 2}^{(\theta_H, \tau_H)})\right]
\]

\[
- \left(\left(c(\theta^b) + c(\tau^b)\right) - 2c + \mathbb{E}_0[\hat{U}_{C, 1}^{(\theta_H, \tau_H)}] - \frac{\eta_{\theta_t}}{2} \text{Var}_0(\mathbb{E}_1 W_{C, 2}^{(\theta_H, \tau_H)})\right).\]

(F.3)

\footnote{As one can see in the proofs, this is the vector of Lagrange multipliers associated with the three possible incentive compatibility constraints. As argued in the static centralized delegation model, if a binding solution exists, only one of these will bind. And hence only the binding constraint will have a nonzero unsigned Lagrange multiplier, while the slack constraints will have zero valued multipliers. If the optimal solution is non-binding, then we return back to the first best case, and in that case, \(\lambda_C = 0\) will be the zero vector — we refer this case back to the first best setup of Proposition E.3 and is not treated here for succinctness in exposition.}
The optimal constant $\lambda_C = (\lambda_C^{(\theta_H, \tau_1)}, \lambda_C^{(\theta_H, \tau_2)}, \lambda_C^{(\theta_L, \tau_1)})$ is the solution to,
\[
\inf_{\lambda_C^{(\theta_H, \tau_1)}, \lambda_C^{(\theta_H, \tau_2)}, \lambda_C^{(\theta_L, \tau_1)} \in \mathbb{R}^3} \hat{U}_{C_P, 0}^{W_C},
\]
only one of $\lambda_C^t$ is nonzero

If a finite value is not reached for the infimum, no second best contract will exist.

At $t = 1$, for any arbitrary $t = 1$ performance fee $y_1$, Manager $C$ who is taking on the compliant investment strategy pair $(\theta_H, \tau_1)$ will be entitled to the total performance fee amount of $y_{C, 1} w_{C, 1}^{(\theta_H, \tau_1)}$, where $W_{C, 1}^{(\theta_H, \tau_1)} = w_{C, 1}^{(\theta_H, \tau_1)}$ is the Principal's $t = 1$ realized wealth. But if Manager $C$ decides to deviate to the deviant investment strategy pair $(\theta_H, \tau_1)$, then Manager $C$'s total performance fee becomes $y_{C, 1} w_{C, 1}^{(\theta_H, \tau_1)}$. Because of the long term investment strategy $(\theta, \tau)$ commitment by Manager $C$ at $t = 0$, there is a strong path dependence on the Principal's $t = 1$ wealth $w_{C, 1}^{(\theta_h, \tau)}$, which then affects Manager $C$'s $t = 1$ performance fee compensation. Thus, analogous to the idea in the static centralization model of Proposition 44, the Principal needs to balance out the cost of Manager's deviant strategy. In all, the optimal $t = 1$ performance fee is precisely (44), and as usual $\lambda_C$ represents the shadow price on Manager $C$'s incentive compatibility.

Now let's discuss the $t = 0$ optimal performance fee choice in (44). From the Principal's perspective, the optimal performance fee $\tilde{y}_{C, 0}$ serves two objectives: (i) optimal risk sharing; and (ii) incentive compatibility. As it was also true in the first best case of Proposition 44, the Principal wants to choose $t = 0$ performance fee to maximize his continuation utility $E_0[W_{C, 1}^{(\theta_H, \tau_1)}]$ while minimizing Manager $C$'s continuation utility $E_0[\tilde{y}_{C, 1}^{(\theta_H, \tau_1)}]$. Since both the Principal and Manager $C$ are both risk averse, again as per the first best case of Proposition 44, the Principal has an intertemporal hedging motive, in which case the $t = 0$ performance fees should minimize the terminal wealth volatility $-(\frac{2}{\sigma_4} \text{Var}_0(E_1 W_{C, 1}^{(\theta_H, \tau_1)}) + \frac{2}{\sigma_2} \text{Var}_0(E_1 W_{C, 2}^{(\theta_H, \tau_1)}))$. But under second best, the performance fees must also induce Manager $C$ to take on the Principal's strictly preferred strategy pair $(\theta_H, \tau_1)$, which are the terms multiplied by $\lambda_C^{(\theta_H, \tau_1)}$. In particular, in equilibrium, the performance fees are set such that Manager $C$'s private costs, the continuation value and Manager $C$'s intertemporal hedging motive $-2c + E_0[\tilde{y}_{C, 1}^{(\theta_H, \tau_1)} - \frac{\sigma_1}{\sigma_4} \text{Var}_0(E_1 W_{C, 2}^{(\theta_H, \tau_1)})]$ under $(\theta_H, \tau_1)$, would equate to the payoff for Manager $C$ under the most profitable investment deviation pair $(\theta_H, \tau_1)$, which consists of the private costs, continuation value and the intertemporal hedging motive, $-c(\theta_H) + c(\tau_1) + E_0[\hat{U}_{C, 1}^{(\theta_H, \tau_1)} - \frac{\sigma_1}{\sigma_4} \text{Var}_0(E_1 W_{C, 2}^{(\theta_H, \tau_1)})]$. Taking all these effects in account, it implies also that the determination of the optimal $t = 0$ performance fee $\tilde{y}_{C, 0}$ must also depend on the full joint distribution of both the compliant and deviant strategy returns $(R_{H, t}, R_{H, t}, R_{D, t}, R_{D, t})$.

The economic reason for why the full joint distribution is perhaps most interesting. The incentive compatibility constraint of Manager $C$ act as an endogenous value-at-risk (VaR) constraint on the performance fee policies across time. The equilibrium wealth paths if Manager $C$ is compliant is $W_{C, t}^{(\theta_H, \tau_1)}$, and the wealth paths of the most profitable deviation is $W_{C, t}^{(\theta_H, \tau_1)}$. To incentivize Manager $C$, it implies that the performance fees over time must be constructed such that the terminal mean-variance of $W_{C, 2}^{(\theta_H, \tau_1)}$, taking into account Manager $C$'s private costs, must weakly exceed that of $W_{C, 2}^{(\theta_H, \tau_1)}$, while due to the individuals’ risk aversions, intertemporal wealth smoothing is also taken into account. As it is common with VaR type constraints, the tail probabilities of returns are of first order importance. In particular, we are now concerned with the joint tail probabilities wealths under the on-equilibrium compliant strategies and off-equilibrium deviant strategies.

The technical reason for why the full joint distribution is required is that we are now dealing with ratios of random variables. One can observe that $\tilde{y}_{C, 1}^{\lambda_C} W_{C, 1}^{(\theta_H, \tau_1)}$ is a ratio of function of wealths $W_{C, 1}^{(\theta_H, \tau_1)}$ and $W_{C, 2}^{(\theta_H, \tau_1)}$. Thus when the Principal needs to decide on the $t = 0$ optimal performance fee $\tilde{y}_{C, 0}$, the Principal needs to consider the $t = 0$ expectation of his $t = 1$ continuation value $\hat{U}_{C, 1}^{\lambda_C}$. And since the expectation of a ratio is generically not equal to the ratio of expectations, it means that one would indeed need the full multivariate
distribution of \((R_{\theta_1}, R_{\tau_1}, R_{\theta_L}, R_{\tau_L})\). Indeed, only when \(\lambda_C = 0\) (i.e. the incentive compatibility constraint is non-binding, or effectively the first best setup), does \(\tilde{g}_{C,1}^{\lambda_C} |_{\lambda_C=0}\) become linear in \(W_{eP,1}(\theta_1, \tau_1)\) and hence overall quadratic in the wealth for \(\tilde{g}_{C,1}^{\lambda_C} |_{\lambda_C=0}\), and hence only in this case, it suffices to just consider the first and second moments of the returns.

G Distribution restrictions in second best

Throughout this paper, we have been relatively silent on the existence of the moment quantities involved in both the optimization of our static and dynamic models. This is especially since in the static models (Section 3), as we work with mean-variance preferences, it is clear that having well defined and finite first and second moments of the investment strategy returns \((R_{\theta}, R_{\tau})\) will suffice for our optimization problem. And indeed, again in the first best dynamic models for both centralized and decentralized delegation, first and second moments existence will also suffice; it should be noted that even though we had made strong independence and identical distribution assumptions (i.e. Assumption and Assumption \(H.3\)), even if we relax these assumptions, it is clear that as long as certain conditional versions of first and second moments exist, everything will still pass through.

But as we discuss the second best dynamic delegation problem, the ratio of functions involving the investment strategy returns will naturally arise from the incentive compatibility constraints. This immediately places a strong restriction on the forms of multivariate distributions \((R_{\theta_1}, R_{\tau_1}, R_{\theta_L}, R_{\tau_L})\) that are permissible in order to have finite first and second moments in the computation of the agents’ continuation utilities and variances of wealth. As an important special case, this immediately rules out \((R_{\theta_1}, R_{\tau_1}, R_{\theta_L}, R_{\tau_L})\) having a jointly Gaussian distribution. This is worthy of mention since numerous theoretical and empirical papers in the asset pricing literature that implicitly or explicitly invoke Gaussian assumptions in the distribution of returns. The key point here is that the introduction of moral hazard in delegated portfolio management, through the incentive compatibility constraints, should give the researcher further pause on how one should think about the returns distribution of not only the equilibrium investment strategies, but also the off-equilibrium returns distributions. We have further remarks on this issue of modeling joint distribution of returns in Section 11, where we will use copulas to model said joint dependence.

H Dependence Modeling and Copulas

As noted in Section 8, where we’d discussed centralized and decentralized delegation under second best, simply knowing the first and second moments of the investment return strategies \((R_{\theta}, R_{\tau})\) for each of \(\theta, \tau\) is not sufficient — one needs to have the full joint distribution \((R_{\theta_1}, R_{\tau_1}, R_{\theta_L}, R_{\tau_L})\) of the return strategies. From Assumption \(\boxplus.2\) of Section 8 and Assumption \(\boxplus.2\) and Assumption \(\boxplus.3\) of Section 8, we have already in place several restrictions on the moments of the investment strategies, which implies that we already have some a priori restrictions on their respective marginal distributions. To further model the joint distribution of these investment strategies, when we already have some specified restrictions on their marginal distributions, the most direct method is via copulas.

H.1 Why Copulas?

The first order of business is to answer a seemingly obvious question — if one wants a multivariate distribution involving four random variables, isn’t the multivariate Gaussian the most convenient and obvious choice?

H.1.1 Why not multivariate Gaussian? Why discrete distributions and copulas?

The reader might naturally wonder why would one not use a four dimensional joint Gaussian distribution on \((R_{\theta_1}, R_{\tau_1}, R_{\theta_L}, R_{\tau_L})\), where we can conveniently impose our restrictions on the means and the variances, and then subsequently model the correlations. While this is statement is true in principle, but in practice we encounter several issues, both on the mathematical aspect and also on the numerical computation aspect.

33 For instance, it is well known that if \(X \sim N(\mu, \sigma^2)\), then \(\mathbb{E}[1/X^k]\) does not exist for any integer \(k\).
As seen in Proposition 2.1 and Proposition 3.1, we require to take the expectation of a ratio of random variables. And for Gaussian distributions, it is well known that something as simple as $X/Y$, where $X$ and $Y$ are independent standard normals, follows the Cauchy distribution, of which no moments exist. A similar issue arises precisely in this setting, as already mentioned in Section 3. As a result, it is highly inconvenient or even wrong to use the seemingly innocuous multivariate Gaussian distribution in our model.

The second reason is a numerical issue. Even if we brush aside probabilistic and integration issues on the existence of moments, from a numerical computation perspective, computing the moments of a ratio still brings substantial challenges. For one, computing a multivariate integral using a four dimensional joint Gaussian is rather computationally intensive, and when the integrands involve ratios of such random variables, we also have to handle numerical stability issues. Indeed, one could make the same remark for general multivariate continuous distributions (say, elliptical distributions) where the density function is non-trivial to compute and subsequently numerically integrate. And we also need to recall, since we are optimizing over an endogenous choice of optimal portfolio and performance fee policies, that means we need to iteratively compute such moments numerous times. In all, this brings about a highly computationally intensive and delicate task.

For these two reasons above, it is far easier to use discrete distributions so that: (i) We do not need to worry about moment existence and integral convergence issues, since we’ll be working with finite sums; and (ii) Discrete distributions with small number of states are far more computationally easily to execute than that of continuous distributions with infinite number of states. And once we recognize the mathematical and numerical practical need to step away from the multivariate Gaussian distributions, where the correlation matrix being the single quantity that governs all dependence behaviors, one then needs to be more delicate in modeling the joint dependence of random variables. As far as this author is aware, the most direct and well-established method is via copulas.

### H.2 Copulas — bare basics

The study of copulas is well established and extensive; see Embrechts et al. (2003) and Patton (2009b, 2012).

Here, we make no attempt to summarize the theories but rather just extract out the minimalistic bare elements that are necessary to achieve two goals for the purpose of this paper: (i) A way to construct joint distributions from the marginal distributions of random variables; and (ii) Parametric copula choices that can qualitatively inform us on the dependence behavior of the random variables.

For this section only, let’s denote $I := [0, 1]$ to be the unit interval, let $\mathbb{R}$ be the extended real line and denote $\text{Ran } f$ to be the range of a function $f$.

We start with the definition of a copula.

**Definition H.1** (Nelsen (2007), Definition 2.10.6). An $n$-dimensional copula (or $n$-copula) is a function $C : I^n \to I$ such that:

1. For every $u \in I^n$,
   
   \[ C(u) = 0 \text{ if at least one coordinate of } u \text{ is } 0, \]

   and,

   \[ C(u) \text{ if all coordinates of } u \text{ are } 1 \text{ except } u_k, \text{ then } C(u) = u_k; \]

2. For all $u_0 = (u_{0,1}, \ldots, u_{0,n}) \in I^n$ and $u_1 = (u_{1,1}, \ldots, u_{1,n}) \in I^n$ such that $u_0 \leq u_1$ (i.e. $u_{0,j} \leq u_{1,j}$ for all $j = 1, \ldots, n$),

   \[ \sum_{i_1=0}^{1} \cdots \sum_{i_n=0}^{1} (-1)^{i_1 + \cdots + i_n} C(u_{i_1,1}, \ldots, u_{i_n,n}) \geq 0. \]

The next key theorem connects copulas to multivariate distributions.

**Theorem H.1** (Sklar’s theorem in $n$-dimensions; Nelsen (2007), Theorem 2.10.9). Let $H$ be an $n$-dimensional distribution function with margins $F_1, F_2, \ldots, F_n$. Then there exists an $n$-copula $C$ such that for all $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$,

\[ H(x_1, x_2, \ldots, x_n) = C(F_1(x_1), F_2(x_2), \ldots, F_n(x_n)). \]
If $F_1, F_2, \ldots, F_n$ are all continuous, then $C$ is unique; otherwise, $C$ is uniquely determined on $\text{Ran} F_1 \times \text{Ran} F_2 \times \cdots \times \text{Ran} F_n$. Conversely, if $C$ is an $n$-copula and $F_1, F_2, \ldots, F_n$ are distribution functions, then the function $H$ defined by (1.1) is an $n$-dimensional distribution function with margins $F_1, F_2, \ldots, F_n$.

To be concrete, let’s consider the probability mass function when the marginal distributions are discrete. For notational simplicity, we assume, only for the expression below, that the random vector $Y$ is discrete and moreover that $Y \in \mathbb{N}^n$, where $\mathbb{N}$ is the set of natural numbers. For the actual application in mind, the support of our random vector will be discrete and finite, and it should be clear from the below that the notation there is straightforward to modify. The probability mass function (pmf) of $Y = (Y_1, \ldots, Y_n)$, and where $F_j$ is the marginal distribution of $Y_j$, for $j = 1, \ldots, n$, is given by the $2^n$ finite differences,

$$
P(Y = y) = \sum_{i_1=0,1} \cdots \sum_{i_n=0,1} (-1)^{i_1+\cdots+i_n} P(Y_1 \leq y_1, \ldots, Y_n \leq y_n - i_n)
= \sum_{i_1=0,1} \cdots \sum_{i_n=0,1} (-1)^{i_1+\cdots+i_n} C(F_1(y_1 - i_1), \ldots, F_n(y_n - i_n)).
$$

For our model, we have $n = 4$ and that $Y_j$’s will be the investment strategy returns, which will be defined on a finite discrete support (which only has few states). Given that the marginal $F_j$’s will be defined on finite discrete supports, these should be relatively quick to compute numerically. However, even if the $F_j$’s are quick to compute, we must not forget that we will need to evaluate these $F_j(\cdot)$’s on the copula $C$ with $n$ arguments, before then computing the $2^n$ finite differences. And this is only for one possible value of the $y$’s. And since we need to be working with moments of ratios of random variables, we need to consider summing over all possibilities of $y$’s with respect to the pmf $P(Y = y)$. Thus, it is imperative that we pick a copula $C$ that is very quick to compute.

**Remark H.2.** As already motivated above, we do not consider numerically intensive procedures that do not admit closed form solutions of the copula $C$. An example of a copula without closed form solutions are elliptical copulas (in which Gaussian copulas are a special case), of which the pmf of a multivariate discrete random variable $Y$ has the form,

$$
P(Y = y) = \int_{\Phi^{-1}(F_j^+)}^{\Phi^{-1}(F_j^-)} \cdots \int_{\Phi^{-1}(F_n^+)}^{\Phi^{-1}(F_n^-)} \phi_n(x_1, \ldots, x_n; \Gamma) \, dx_1 \cdots dx_n,
$$

where here, $\phi_n(\cdot; \Gamma)$ denotes the probability density function (pdf) of an $n$-dimensional elliptical distribution at location 0 and with scale parameter (correlation matrix) $\Gamma$, and $\Phi^{-1}$ denotes the inverse cumulative distribution function (cdf) of the univariate margins of the said elliptical distribution; and $F_j^+ := P(Y_j \leq y_j)$ and $F_j^- := P(Y_j \leq y_j - 1)$. Clearly, there are no closed form solutions for $\Phi^{-1}$. See [Joe (2014)] for details.

The main point for the purpose of our paper is that these distributions that do not admit closed form solutions for the copula entails highly numerically intensive computations (i.e. for both $\Phi^{-1}$, and the multivariate numerical integration) for even one single computation of $P(Y = y)$. In particular, we need to consider expectations of the form $E[g(Y; v)] = \sum g(y; v)P(Y = y)$, where for our purposes, $g$ itself already is somewhat complicated in $y$ and may not have closed form solutions, and moreover, that we need to further numerically optimize over the endogenous variable(s) $v$. All such numerical computations make this copula family numerically unsuitable for our paper — even though we fully acknowledge that, with sufficient computing resources, it would be interesting to explore this copula family since the correlation scale parameter $\Gamma$ allows for a richer dependence structure than the Archimedean family that we discuss below.

### H.3 Archimedean copulas

For the purpose of this paper, we will only consider the family of Archimedean copulas. Let’s begin with a technical definition.

---

34 In particular, we do not consider the multivariate Gaussian copula (see Remark H.2), which is widely popular in mathematical finance (say [Joe (2014)]), largely only because of computational speed problems for our model at hand. [Nelsen (2006)] and [Joe (1997, 2000)] are standard references that contain excellent overviews of various types of copulas and their properties.
Definition H.2 (Completely monotonic function: Nelsen (2007), Definition 4.6.1.) A function $g(t)$ is completely monotonic on an interval $J$ if it is continuous there and has derivatives of all orders that alternate in sign; i.e., if it satisfies,

$(-1)^k \frac{d^k}{dt^k} g(t) \geq 0,$

for all $t$ in the interior of $J$ and $k = 0, 1, 2, \ldots$

Now, we can give the definition of an $n$-dimensional (exchangeable) Archimede copula.\footnote{The theorem statement of Theorem 4.6.2. in Nelsen (2007) yields the definition of a multivariate Archimedean copula.}

Definition H.3 (Multivariate Archimedian copula: Nelsen (2007), Theorem 4.6.2.). Let $\varphi$ be a continuous strictly decreasing function from $I$ to $[0, \infty]$ such that $\varphi(0) = \infty$ and $\varphi(1) = 0$, and let $\varphi^{-1}$ denote the inverse of $\varphi$. If $C$ is the function from $I^n$ to $I$ given by,

$$C(u) = \varphi^{-1}(\varphi(u_1) + \varphi(u_2) + \ldots + \varphi(u_n)),$$

then $C$ is an $n$-copula for all $n \geq 2$ if and only if $\varphi^{-1}$ is completely monotonic on $[0, \infty)$. The function $\varphi$ is called the generator of the copula.\footnote{For most purposes and also in the literature, this generator is parametrized by a single scalar $\delta$. The theorem statement of Theorem 4.6.2. in Nelsen (2007) yields the definition of a multivariate Archimedean copula.}

There are numerous properties associated with the Archimedean copula (see Nelsen (2007)) but for our purposes, one of the most restrictive implications of this copula is that it implies an exchangeable distribution. That is, if the multivariate distribution of $(Y_1, \ldots, Y_n)$ is constructed from an Archimedean copula, then it is equivalent in distribution to $(Y_{\sigma(1)}, \ldots, Y_{\sigma(n)})$, where $\sigma(1), \ldots, \sigma(n)$ are any permutations of $1, \ldots, n$. This is admittedly restrictive. But in return, we get a multivariate distribution that is based on a single parameter that can then generate various tail dependence behaviors; these tail dependence behaviors are what’s most important for the purpose of our application.

H.3.1 Examples

In what follows, we will list out both the bivariate and multivariate forms of the classical Archimedean copulas. We will use these examples as fundamental building blocks in the subsequent constructions of hierarchical Archimedean copulas. In the examples below, we will also record the inverse of the generator as it will become useful in the subsequent sections.

Example 1 (Clayton). For $0 \leq \delta < \infty$, the bivariate Clayton copula is,

$$C(u, v; \delta) = (u^{-\delta} + v^{-\delta} - 1)^{-1/\delta}, \quad u, v \in I,$$

where its generator is,

$$\varphi(s; \delta) = (1 + s)^{-1/\delta}, \quad \varphi^{-1}(t; \delta) = t^{-\delta} - 1, \quad t \in I.$$\footnote{There are several different measures of tail dependence and details can be found in Joe (2014). For our purposes, a qualitative description suffices. We say that a pair of random variables exhibit upper tail dependence when one realizes high extreme values, the other also realizes high extreme values. Likewise, we say a pair of random variables exhibit lower tail dependence when one realizes low extreme values, the other also realizes low extreme values. It should also be noted that by their constructions, it is fairly difficult for Archimedean copulas to generate distributions whereby one random variable realizes high extreme values, while the other one realizes low extreme values. Thus, when one computes (Pearson’s) correlation on such random variables, they will tend to be nonnegative. But it should be noted that correlation computations obscure a critical qualitative behavior. Namely, it is possible to have high positive correlations with or without tail dependence. Moreover, if two random variables exhibit upper or lower tail dependencies, they’ll both result in positive (Pearson’s) correlation, and yet their qualitative behaviors are completely different and indeed opposing.}

The multivariate Clayton copula is,

$$C(u; \delta) = [u_1^{-\delta} + \ldots + u_n^{-\delta} - (n - 1)]^{-1/\delta}, \quad u \in I^n.$$

See Joe (2014), Section 4.6.1) for details.
Figure 18: A scatter plot of \((X_1, X_2)\), whose joint distribution is generated by the Clayton copula of Example 11 with various parameters. The marginal distributions are \(X_i \sim \mathcal{N}(0.20, 0.35)\) for \(i = 1, 2\).

Example 2 (Frank). For \(-\infty < \delta < \infty\), the \textit{bivariate Frank} copula is,

\[
C(u, v; \delta) = -\delta^{-1} \log \left( \frac{1 - e^{-\delta} - (1 - e^{-\delta u})(1 - e^{-\delta v})}{1 - e^{-\delta}} \right), \quad u, v \in \mathbf{I},
\]

where its generator is,

\[
\phi(s; \delta) = -\delta^{-1} \log \left(1 - (1 - e^{-\delta})e^{-s}\right), \quad s \in \mathbf{I},
\]

\[
\phi^{-1}(t; \delta) = -\log \frac{1 - e^{-\delta t}}{1 - e^{-\delta}}, \quad t \in \mathbf{I}.
\]

The \textit{multivariate Frank} copula is,

\[
C(u; \delta) = -\delta^{-1} \log \left(1 - \prod_{j=1}^{n} (1 - e^{-\delta u_j}) \right), \quad u \in \mathbf{I}^n.
\]

See Joe (2014, Section 4.5.1) for details.

Figure 19: A scatter plot of \((X_1, X_2)\), whose joint distribution is generated by the Frank copula of Example 2 with various parameters. The marginal distributions are \(X_i \sim \mathcal{N}(0.20, 0.35)\) for \(i = 1, 2\).

Example 3 (Gumbel). For \(1 \leq \delta < \infty\), the \textit{bivariate Gumbel} copula is,

\[
C(u, v; \delta) = \exp \left\{ - \left( -\log u \right)^{\delta} + \left( -\log v \right)^{\delta} \right\}^{1/\delta}, \quad u, v \in \mathbf{I},
\]

where its generator is,

\[
\phi(s; \delta) = \exp(-s^{1/\delta}), \quad s \in \mathbf{I},
\]

\[
\phi^{-1}(t; \delta) = (- \log t)^{\delta}, \quad t \in \mathbf{I}.
\]
The **multivariate Gumbel** copula is,

\[
C(u; \delta) = \exp \left\{ - \left( \sum_{j=1}^{n} [- \log u_j]^{\delta} \right)^{1/\delta} \right\}, \quad u \in I^n.
\]  

(H.12)

See [Joe (2014), Section 4.8.1] for details.

![Figure 20](image)

**Figure 20:** A scatter plot of \((X_1, X_2)\), whose joint distribution is generated by the Gumbel copula of Example 3 with various parameters. The marginal distributions are \(X_i \sim N(0.20, 0.35)\) for \(i = 1, 2\).

### H.4 Approximating continuous marginals by discrete marginals

As mentioned at the beginning of Section H, due to both moment existence issues on the theoretical end and also on numerical computation issues in practice, in this paper we will use discrete distributions rather than continuous distributions. In particular, in light of the copula discussion above, we just need to focus on describing a method to construct marginal distributions \(F_i\)'s, and the joint distribution will then be applied via the copula.

So suppose we have finite number of states. Rather than assigning probability weights in a potentially arbitrary fashion, we will use a more systematic method to construct the probabilities. We will construct a discrete distribution that matches the moments of a corresponding parametric continuous distribution. We follow the procedure described in [Luceno (1999)]; although we acknowledge that the core ideas are much older and are already described in [Abramowitz and Stegun (1972)] (see also [Stoer and Bulirsch (2002)]). In particular, we approximate the Gaussian distribution. In the actual applications to this paper, we will model the returns of all strategies \(R_\theta, R_\tau\) to have Gaussian marginals, with moments matching Assumption 3.2, Assumption 6.2 and Assumption 6.3.

I Additional Results and Proofs for Section 7

**Proof of Proposition E.1.** (a) Fix any investment strategy pair \((\theta, \tau)\).

(i) For any given contract \((x, \{y_{C,0}, y_{C,1}\})\), the problem of solving for the portfolio policy of Manager \(C\) is,

\[
\sup_{\psi_0, \psi_1 \in \mathbb{R}} \left\{ -(c(\theta) + c(\tau)) + \mathbb{E}_0 [x + \bar{W}^{(\theta, \tau)}_{C, t}] - \frac{\eta_0}{2} \text{Var}_0 (\bar{W}^{(\theta, \tau)}_{C, t}) \right\} = -(c(\theta) + c(\tau)) + x_C + \bar{U}^{(\theta, \tau)}_{C, 0},
\]

where recall that \(\bar{W}^{(\theta, \tau)}_{C, t} = \bar{W}^{(\theta, \tau)}_{C, t} (\psi_0, \psi_1; x_C, \{y_{C,0}, y_{C,1}\})\), as per (E.1f), (E.1g), is the time \(t\) wealth of Manager \(C\) (except for the fixed fee \(x_C\)) for an arbitrary contract, and also an arbitrary portfolio policy. Using the dynamic programming principle as per Section C, we are motivated to

\[^{38}\text{In the application, we’ll be taking small number of states, say three.}\]
Next, again holding for an arbitrary contract, we consider the

Thus from before, for any given arbitrary contract \( \{ y_{C,0}, y_{C,1} \} \), we have solved for Manager \( C \)'s optimal portfolio policies \( \hat{\psi}_{t,(\theta, \tau)}(x_C, \{ y_{C,0}, y_{C,1} \}) \) for all times \( t = 0, 1 \). We now solve the optimal contact as chosen by the Principal.

(i) From the individual rationality constraint (I.11), it will bind and implying the fixed fee \( x_C \) is (E.2).

(ii) Let us now optimize over the performance fees \( \{ y_{C,0}, y_{C,1} \} \). Recalling the objective function (DynCen), and substituting in the fixed fee \( x_{C,(\theta, \tau)} \), the optimization problem of the Principal at \( t = 0 \) is,

\[
\hat{U}_{c,P,0} = \sup_{x_C \in \mathbb{R}, \ y_{C,0}, y_{C,1} \in [0,1]} \left[ -x_C + \mathbb{E}_0[W_{c,P,2}^{(\theta, \tau)}] - \frac{\eta_p}{2} \text{Var}_0(W_{c,P,2}^{(\theta, \tau)}) \right] \\
= \sup_{y_{C,0}, y_{C,1} \in [0,1]} \left[ -(c(\theta) + c(\tau)) + \mathbb{E}_0[W_{C,2}^{(\theta, \tau)}] - \frac{\eta_p}{2} \text{Var}_0(W_{C,2}^{(\theta, \tau)}) \right] \\
+ \mathbb{E}_0[W_{C,2}^{(\theta, \tau)}] - \frac{\eta_p}{2} \text{Var}_0(W_{C,2}^{(\theta, \tau)}) \\
= \sup_{y_{C,0}, y_{C,1} \in [0,1]} \left[ -(c(\theta) + c(\tau)) + \hat{U}_{t,0} + U_{C,0} \right],
\]

where we have defined, for any \( y_{C,0}, y_{C,1} \in [0,1] \),

\[
\hat{U}_{c,P,0}^{(\theta, \tau)} := \mathbb{E}_0[W_{c,P,2}^{(\theta, \tau)}] - \frac{\eta_p}{2} \text{Var}_0(W_{c,P,2}^{(\theta, \tau)}), \\
U_{C,0}^{(\theta, \tau)} := \mathbb{E}_0[W_{C,2}^{(\theta, \tau)}] - \frac{\eta_p}{2} \text{Var}_0(W_{C,2}^{(\theta, \tau)}).
\]

Now, applying the dynamic programming principle of Section C, this motivates the definition that for \( t = 0, 1 \),

\[
\hat{U}_{c,P,t} = \mathbb{E}_t \hat{U}_{C,P,t+1} - \frac{\eta_p}{2} \text{Var}_t(\mathbb{E}_{t+1} W_{c,P,2}), \\
U_{C,t}^{(\theta, \tau)} = \mathbb{E}_t U_{C,t+1} - \frac{\eta_p}{2} \text{Var}_t(\mathbb{E}_{t+1} W_{C,2}).
\]
This implies we can now consider the problem

$$
\hat{U}_{cP,1}^{(\theta, \tau)} = \sup_{y_{C,1} \in [0,1]} \hat{U}_{cP,1}^{(\theta, \tau)} + \hat{U}_{C,1}^{(\theta, \tau)},
$$

(I.9)

to determine the $t = 1$ fees. Taking first order conditions for optimization, and actually analogous to the static centralized delegation problem in first best of Proposition (I.1), the first order conditions associated with the performance fees $y_{C,1}$ will have four roots, them being,

$$
\left\{ \begin{array}{l}
\eta_P \\
\frac{\eta_F + \eta_M}{\eta_P + \eta_M} - \frac{(E_1 R_{o,2} - E_1 R_{r,2})^{2/3}}{(E_1 R_{o,2} - E_1 R_{r,2})^{2/3}} \\
+ \frac{(-1)^{2/3}(E_1 R_{o,2} - E_1 R_{r,2})^{2/3}}{(E_1 R_{o,2} - E_1 R_{r,2})^{2/3}} \\
\end{array} \right.
$$

but the only real-valued solution in $[0, 1]$ is clearly the first one. Thus, we have determined Principal’s $t = 1$ optimal performance fee policy (E.3). Let $\hat{U}_{cP,1}^{(\theta, \tau)}$ be the value of $\hat{U}_{cP,1}^{(\theta, \tau)}$ evaluated at $\hat{y}_{C,1}$, and also let $\hat{U}_{C,1}^{(\theta, \tau)}$ be the value of $U_{C,1}^{(\theta, \tau)}$ evaluated at $\hat{y}_{C,1}$, so that $\hat{U}_{cP,1}^{(\theta, \tau)} = \hat{U}_{cP,1}^{(\theta, \tau)} + \hat{U}_{C,1}^{(\theta, \tau)}$.

(iii) Now, we continue to optimize for the $t = 0$ performance fees $y_{C,0} \in [0, 1]$. From the dynamic programming principle, we can consider the problem,

$$
\hat{U}_{cP,0}^{(\theta, \tau)} = \sup_{y_{C,0} \in [0,1]} E_0 [\hat{U}_{cP,1}^{(\theta, \tau)} - \frac{\eta_P}{2} \text{Var}_0 (E_1 W_{cP,2}^{(\theta, \tau)}) - \frac{\eta_M}{2} \text{Var}_0 (E_1 W_{C,2}^{(\theta, \tau)})],
$$

(I.10)

which is (E.4).

**Proof of Corollary (I.2).** The $t = 1$ results are a simple application of Assumption (E.2) to the results of Proposition (I.1). However, by using Assumption (E.3), we can get substantial simplifications and clarity in the $t = 0$ results. One can readily verify that with the simplifications provided by Assumption (E.3), the objective function for the $t = 0$ fees optimization becomes a concave quadratic in $y_{0}$, and thus if an interior maximizer exists in $(0, 1)$, then it must be unique and can be characterized by first order conditions.

---

**I.1 Decentralization**

**Proposition I.1.** Consider the first best decentralized delegation problem (DynDec) but without the incentive compatibility constraints (E.11) and (I.13). Fix any strategy pair $(\theta, \tau) \in \mathcal{S}$. Assume Assumption (E.3).

(a) For any portfolio policy and performance fees $(\pi_0, \pi_1, y_{A,0}, y_{A,1}, y_{A,0}, y_{A,1})$, the optimal fixed fees for Manager A and Manager B are, respectively,

$$
\hat{x}_A(\theta, \tau) = c(\theta) - E_0 [W_{A,2}^{(\theta, \tau)}] + \frac{\eta_M}{2} \text{Var}_0 (W_{A,2}^{(\theta, \tau)}),
$$

(I.11a)

$$
\hat{x}_B(\theta, \tau) = c(\tau) - E_0 [W_{B,2}^{(\theta, \tau)}] + \frac{\eta_M}{2} \text{Var}_0 (W_{B,2}^{(\theta, \tau)}).
$$

(I.11b)

(b) Suppose the $t = 1$ realized value of the Principal’s wealth is $W_{P,1}^{(\theta, \tau)} = u_{P,1}^{(\theta, \tau)}$. Then the $t = 1$ optimal policies are given as follows.

(i) The $t = 1$ optimal portfolio chosen by the Principal is $\pi_1^{N}$,

$$
\hat{\pi}_1(\theta, \tau) = \frac{\hat{\pi}_1^{1N}}{\pi_1^{1D}} + \frac{\hat{\pi}_1^{2N}}{\pi_1^{1D}},
$$

(I.12)

\[\text{i.e. } "N" \text{ for numerator, and } "D" \text{ for denominator.}\]
provided that $\hat{\pi}_{1, (\theta, \tau)} \in (0, 1)$ and where,

$$\hat{\pi}_{1}^N := \text{Var}_1(R_{\theta, 2}) \left[ (\eta_P + \eta_M)\text{Var}_1(R_{\theta, 2})\text{Var}_1(R_{\tau, 2}) - \text{Cov}_1(R_{\theta, 2}, R_{\tau, 2})(\eta_M\text{Var}_1(R_{\tau, 2}) + \eta_P\text{Cov}_1(R_{\theta, 2}, R_{\tau, 2})) \right],$$

$$\hat{\pi}_{1}^D := (\eta_P + \eta_M)\text{Var}_1(R_{\theta, 2})\text{Var}_1(R_{\tau, 2})[\text{Var}_1(R_{\theta, 2}) + \text{Var}_1(R_{\tau, 2})] - 2\eta_P\text{Var}_1(R_{\theta, 2})\text{Var}_1(R_{\tau, 2})\text{Cov}_1(R_{\theta, 2}, R_{\tau, 2}) - \eta_P(\text{Var}_1(R_{\theta, 2}) + \text{Var}_1(R_{\tau, 2}))\text{Cov}_1(R_{\theta, 2}, R_{\tau, 2})^2,$$

and,

$$\hat{\pi}_{1}^{2N} := (E_1R_{\tau, 2} - E_1R_{\theta, 2}) \left[ (\eta_P^2\text{Cov}_1(R_{\theta, 2}, R_{\tau, 2})^2 - (\eta_P + \eta_M)^2\text{Var}_1(R_{\theta, 2})\text{Var}_1(R_{\tau, 2})) \right],$$

$$\hat{\pi}_{1}^{2D} := \eta_P\eta_Mw_{P, 1}^{(\theta, \tau)} \hat{\pi}_{1}^D.$$

(ii) The optimal $t = 1$ performance fee chosen by the Principal to compensate Manager $A$ is,

$$\hat{y}_{A, 1, (\theta, \tau)} = \frac{\hat{y}_{A, 1}^{N}}{\hat{y}_{A, 1}^{D}}, \quad (I.13)$$

provided that $\hat{y}_{A, 1, (\theta, \tau)} \in (0, 1)$, and where,

$$\hat{y}_{A, 1}^{N} := \eta_P^2\eta_M\text{Var}_1(R_{\theta, 2})[\text{Var}_1(R_{\theta, 2})\text{Var}_1(R_{\tau, 2}) - \text{Cov}_1(R_{\theta, 2}, R_{\tau, 2})^2]w_{P, 1}$$

$$+ \eta_P(E_1R_{\theta, 2} - E_1R_{\tau, 2}) \left[ (\eta_P + \eta_M)\text{Var}_1(R_{\theta, 2})\text{Var}_1(R_{\tau, 2}) - \text{Cov}_1(R_{\theta, 2}, R_{\tau, 2})(\eta_M\text{Var}_1(R_{\tau, 2}) + \eta_P\text{Cov}_1(R_{\theta, 2}, R_{\tau, 2})) \right],$$

$$\hat{y}_{A, 1}^{D} := (E_1R_{\theta, 2} - E_1R_{\tau, 2})[\eta_P + \eta_M]^2\text{Var}_1(R_{\theta, 2})\text{Var}_1(R_{\tau, 2}) - \eta_P^2\text{Cov}_1(R_{\theta, 2}, R_{\tau, 2})^2]$$

$$+ \eta_P\eta_M\text{Var}_1(R_{\tau, 2}) \left[ - \eta_P\text{Cov}_1(R_{\theta, 2}, R_{\tau, 2})^2 + \text{Var}_1(R_{\theta, 2})[(\eta_P + \eta_M)\text{Var}_1(R_{\tau, 2}) - \eta_M\text{Cov}_1(R_{\theta, 2}, R_{\tau, 2})] \right]w_{P, 1}^{(\theta, \tau)}.$$

(iii) The optimal $t = 1$ performance fee chosen by the Principal to compensate Manager $B$ is,

$$\hat{y}_{B, 1, (\theta, \tau)} = \frac{\hat{y}_{B, 1}^{N}}{\hat{y}_{B, 1}^{D}}, \quad (I.14)$$

provided that $\hat{y}_{B, 1, (\theta, \tau)} \in (0, 1)$, and where,

$$\hat{y}_{B, 1}^{N} := \eta_P^2\eta_M\text{Var}_1(R_{\theta, 2})[-\text{Var}_1(R_{\theta, 2})\text{Var}_1(R_{\tau, 2}) + \text{Cov}_1(R_{\theta, 2}, R_{\tau, 2})^2]w_{P, 1}^{(\theta, \tau)}$$

$$+ \eta_P(E_1R_{\theta, 2} - E_1R_{\tau, 2}) \left[ - \eta_P\text{Cov}_1(R_{\theta, 2}, R_{\tau, 2})^2 + \text{Var}_1(R_{\theta, 2})[(\eta_P + \eta_M)\text{Var}_1(R_{\tau, 2}) - \eta_M\text{Cov}_1(R_{\theta, 2}, R_{\tau, 2})] \right],$$

$$\hat{y}_{B, 1}^{D} := (E_1R_{\theta, 2} - E_1R_{\tau, 2})[\eta_P + \eta_M]^2\text{Var}_1(R_{\theta, 2})\text{Var}_1(R_{\tau, 2}) - \eta_P^2\text{Cov}_1(R_{\theta, 2}, R_{\tau, 2})^2]$$

$$+ \eta_P\eta_M\text{Var}_1(R_{\theta, 2}) \left[ - (\eta_P + \eta_M)\text{Var}_1(R_{\theta, 2})\text{Var}_1(R_{\tau, 2})$$

$$+ \text{Cov}_1(R_{\theta, 2}, R_{\tau, 2})(\eta_M\text{Var}_1(R_{\tau, 2}) + \eta_P\text{Cov}_1(R_{\theta, 2}, R_{\tau, 2})) \right]w_{P, 1}^{(\theta, \tau)}.$$

(c) The $t = 0$ optimal portfolio and fee policies $(\hat{\pi}_{0, (\theta, \tau)}, \hat{y}_{A, 0, (\theta, \tau)}, \hat{y}_{B, 0, (\theta, \tau)})$ are obtained by solving the following.
(i) Define the $t = 1$ continuation utilities for the Principal, Manager $A$ and Manager $B$, respectively:

\[ \tilde{U}^{(\theta, \tau)}_{P,1} := E_1 W^{(\theta, \tau)}_{P,2} - \frac{\eta_p}{2} \text{Var}_1( W^{(\theta, \tau)}_{P,2}), \quad \text{I.15a} \]

\[ U^{(\theta, \tau)}_{A,1} := E_1 W^{(\theta, \tau)}_{A,2} - \frac{\eta_M}{2} \text{Var}_1( W^{(\theta, \tau)}_{A,2}), \quad \text{I.15b} \]

\[ U^{(\theta, \tau)}_{B,1} := E_1 W^{(\theta, \tau)}_{B,2} - \frac{\eta_M}{2} \text{Var}_1( W^{(\theta, \tau)}_{B,2}), \quad \text{I.15c} \]

where the $t = 2$ wealth expressions, $W^{(\theta, \tau)}_{P,2}, W^{(\theta, \tau)}_{A,2}, W^{(\theta, \tau)}_{B,2}$, have substituted in the $t = 1$ optimal policies $(\tilde{\pi}_1, (\theta, \tau), \tilde{y}_{A,1}, (\theta, \tau), \tilde{y}_{B,1}, (\theta, \tau))$ found above. Then one can write,

\[ \tilde{U}^{(\theta, \tau)}_{P,1} = -\frac{u^N_P}{u^P_P} (W^{(\theta, \tau)}_{P,1})^2 + \frac{u^2N_P}{u^2P_P} W^{(\theta, \tau)}_{P,1} + \frac{u^3N_P}{u^3P_P}, \quad \text{I.16a} \]

\[ U^{(\theta, \tau)}_{A,1} = W^{(\theta, \tau)}_{A,1} - \frac{u^N_A}{u^A_A} (W^{(\theta, \tau)}_{P,1})^2 - \frac{u^2N_A}{u^2A_A} W^{(\theta, \tau)}_{P,1} + \frac{u^3N_A}{u^3A_A}, \quad \text{I.16b} \]

\[ U^{(\theta, \tau)}_{B,1} = W^{(\theta, \tau)}_{B,1} - \frac{u^N_B}{u^B_B} (W^{(\theta, \tau)}_{P,1})^2 - \frac{u^2N_B}{u^2B_B} W^{(\theta, \tau)}_{P,1} + \frac{u^3N_B}{u^3B_B}, \quad \text{I.16c} \]

where $u^{ij}_k$, for $k = P, A, B$, $i = 1, 2, 3$ and $j = N, D$, are constants that relate to the risk aversion parameters $(\eta_p, \eta_M)$ of the Principal and Managers $A$ and $B$, and to the time $t = 1$ expectations, variances and covariances of the investment strategy returns pair $(R_{0,2}, R_{t,2})$ at $t = 2$. In particular, $u^{1N}_k, u^{1D}_k$, for $k = P, A, B$, are strictly positive terms. (The precise analytical forms of these $u^{ij}_k$’s are in the proof).

(ii) After substituting in the optimal $t = 1$ policies, the $t = 1$ expectation of the $t = 2$ terminal wealths are,

\[ E_1 W^{(\theta, \tau)}_{P,2} = \frac{w^{1N}_P}{w^P_P} W^{(\theta, \tau)}_{P,1} + \frac{w^{2N}_P}{w^2P_P}, \quad \text{I.17a} \]

\[ E_1 W^{(\theta, \tau)}_{A,2} = W^{(\theta, \tau)}_{A,1} - \frac{w^{1N}_A}{w^A_A} W^{(\theta, \tau)}_{P,1} + \frac{w^{2N}_A}{w^2A_A}, \quad \text{I.17b} \]

\[ E_1 W^{(\theta, \tau)}_{B,2} = W^{(\theta, \tau)}_{B,1} - \frac{w^{1N}_B}{w^B_B} W^{(\theta, \tau)}_{P,1} + \frac{w^{2N}_B}{w^2B_B}, \quad \text{I.17c} \]

where $w^{ij}_k$, for $k = P, A, B$, $i = 1, 2$ and $j = N, D$, are constants that relate to the risk aversion parameters $(\eta_p, \eta_M)$ of the Principal and Managers $A$ and $B$, and to the time $t = 1$ expectations, variances and covariances of the investment strategy returns pair $(R_{0,2}, R_{t,2})$ at $t = 2$. (The precise analytical forms of these $w^{ij}_k$’s are in the proof).

(iii) The optimal $t = 0$ portfolio and performance fees policies $(\tilde{\pi}_0, (\theta, \tau), \tilde{y}_{A,0}, (\theta, \tau), \tilde{y}_{B,0}, (\theta, \tau))$ are obtained by maximizing $(\pi_0, y_{A,0}, y_{B,0})$,

\[ \tilde{U}^{(\theta, \tau)}_{P,0} = \sup_{y_{A,0}, y_{B,0} \in [0,1]} \sup_{\pi_0 \in \mathbb{R}} E_0 \left[ \tilde{U}^{(\theta, \tau)}_{P,1} + t^{(\theta, \tau)}_{A,1} + U^{(\theta, \tau)}_{B,1} \right] \]

\[ - \frac{\eta_p}{2} \text{Var}_0( E_1 W^{(\theta, \tau)}_{P,2} ) - \frac{\eta_M}{2} \text{Var}_0( E_1 W^{(\theta, \tau)}_{A,2} ) - \frac{\eta_M}{2} \text{Var}_0( E_1 W^{(\theta, \tau)}_{B,2} ). \quad \text{I.18} \]

Let’s begin by discussing the $t = 1$ optimal portfolio and fees policies, $(\tilde{\pi}_1, (\theta, \tau), \tilde{y}_{A,1}, (\theta, \tau), \tilde{y}_{B,1}, (\theta, \tau))$, of the Principal in first best dynamic decentralized delegation as per Proposition 14. At $t = 1$, the realized wealths of the Principal, Manager $A$ and Manager $B$ become, respectively, $W^{(\theta, \tau)}_{P,1} = w^{(\theta, \tau)}_{P,1}, W^{(\theta, \tau)}_{A,1} = w^{(\theta, \tau)}_{A,1}$ and $W^{(\theta, \tau)}_{B,1} = w^{(\theta, \tau)}_{B,1}$. By the linearity of the contracts offered and since $t = 2$ is the terminal contracting date, it implies that from the Principal’s perspective, the $t = 1$ wealths $w^{(\theta, \tau)}_{A,1}, w^{(\theta, \tau)}_{B,1}$ of Managers $A$ and $B$ do not come into his decision making. Thus at $t = 1$, the Principal simply needs to choose portfolios $\pi_1$ to
maximize his $t = 2$ returns, while simultaneously using the portfolio and fee policies $(\pi_1, y_{A,1}, y_{B,1})$ to risk share with Managers $A$ and $B$. We should note, however, for a generic pair of investment strategies $(\theta, \tau)$, there is a distinct Principal $t = 1$ wealth effect $u_{P,1}^{(\theta, \tau)}$ that enters into the optimal portfolio $\hat{x}_{1, (\theta, \tau)}$ and the optimal performance fees $(\hat{\gamma}_{A,1,(\theta, \tau)}, \hat{\gamma}_{B,1,(\theta, \tau)})$.

Next, let’s consider the $t = 0$ optimal policies (I.13) for the Principal. The Principal needs to take into account the motives of himself and the other two agents. In particular, by defining the continuation utilities (I.18) into account the motives of himself and the other two agents. In particular, by defining the continuation utilities (I.18) of the Principal, Manager $A$ and Manager $B$, the $t = 0$ optimal portfolio and performance fee policies effectively maximize the Principal’s continuation utility $E_0[T_{P,1}^{(\theta, \tau)}]$, while minimizing Manager $A$’s and Manager $B$’s continuation utilities $E_0[T_{A,1}^{(\theta, \tau)} + U_{B,1}^{(\theta, \tau)}]$. Simultaneously, given that all individuals have mean-variance preferences over terminal wealth, and hence an intertemporal hedging motive is in effect, the Principal’s optimal $t = 0$ policies must minimize the volatility of all individuals’ terminal wealths $E_1 W_{k,2}^{(\theta, \tau)}$, for $k = P, A, B$.

**Proof of Proposition I.1.** (a) As it is usual, the individual rationality constraints (I.19), (I.11) for both Manager $A$ and $B$ will bind. This pins down the optimal fixed fees for Manager $A$ and $B$ as given in (I.11).

(b) From the objective function (DynDec) and using the optimal form of the fixed fees (I.11), we consider the optimization problem,

\[
\sup_{x_A, x_B \in \mathbb{R}, \ y_{A,0}, y_{A,1} \in [0,1], \ y_{B,0}, y_{B,1} \in [0,1]} \sup_{\pi_0, \pi_1} -x_A - x_B + E_0[W_{P,2}^{(\theta, \tau)}] - \frac{\eta_P}{2} \text{Var}_0(W_{P,2}^{(\theta, \tau)})
\]

\[
= \sup_{y_{A,0}, y_{A,1} \in [0,1], \ \pi_0, \pi_1 \in \mathbb{R}} \sup_{y_{B,0}, y_{B,1} \in [0,1]} -(c(\theta) + c(\tau)) + E_0[W_{A,2}^{(\theta, \tau)}] - \frac{\eta_M}{2} \text{Var}_0(W_{A,2}^{(\theta, \tau)})
\]

\[
+ E_0[W_{B,2}^{(\theta, \tau)}] - \frac{\eta_M}{2} \text{Var}_0(W_{B,2}^{(\theta, \tau)})
\]

\[
+ E_0[W_{P,2}^{(\theta, \tau)}] - \frac{\eta_P}{2} \text{Var}_0(W_{P,2}^{(\theta, \tau)}).
\]

At this point, motivated by the dynamic programming principle of Section I.1, let us further define, recursively for $t = 0, 1$,

\[
\hat{U}_{P,t}^{(\theta, \tau)} = E_t \hat{U}_{P,t+1}^{(\theta, \tau)} - \frac{\eta_P}{2} \text{Var}_t(E_{t+1} W_{P,T}^{(\theta, \tau)}),
\]

\[
U_{A,t}^{(\theta, \tau)} = E_t U_{A,t+1}^{(\theta, \tau)} - \frac{\eta_M}{2} \text{Var}_t(E_{t+1} W_{A,T}^{(\theta, \tau)}),
\]

\[
U_{B,t}^{(\theta, \tau)} = E_t U_{B,t+1}^{(\theta, \tau)} - \frac{\eta_M}{2} \text{Var}_t(E_{t+1} W_{B,T}^{(\theta, \tau)}).
\]

Define,

\[
\hat{U}_{P,0}^{(\theta, \tau)} = \sup_{y_{A,0}, y_{A,1} \in [0,1], \ \pi_0, \pi_1 \in \mathbb{R}} \sup_{y_{B,0}, y_{B,1} \in [0,1]} \hat{U}_{P,0}^{(\theta, \tau)} + U_{A,0}^{(\theta, \tau)} + U_{B,0}^{(\theta, \tau)},
\]

so that our optimization problem (I.11) can be rewritten as,

\[
-(c(\theta) + c(\tau)) + \hat{U}_{P,0}^{(\theta, \tau)}.
\]

Now we can consider the Principal’s problem at $t = 1$. Using the budget constraints, the Principal’s
Next, we consider the objective is to maximize over portfolio policies $\pi_1$ and the fees $(y_{A,1}, y_{B,1}) \in [0, 1]^2$,

$$
\tilde{U}^{(\theta, \tau)}_{P,1} + U^{(\theta, \tau)}_{A,1} + U^{(\theta, \tau)}_{B,1} = E_1 [ W^{(\theta, \tau)}_{P,2} + W^{(\theta, \tau)}_{A,2} + W^{(\theta, \tau)}_{B,2} ] - \frac{\eta_{M}}{2} \text{Var}_{1} (W^{(\theta, \tau)}_{A,2}) - \frac{\eta_{M}}{2} \text{Var}_{1} (W^{(\theta, \tau)}_{B,2}) - \frac{\eta_{M}}{2} \text{Var}_{1} (W^{(\theta, \tau)}_{P,2})
$$

$$
= E_1 \left[ w^{(\theta, \tau)}_{P,1} (1 + \pi_1 (1 - y_{B,1}) R_{\tau,2} + (1 - \pi_1) (1 - y_{A,1}) R_{\theta,2}) \right]
+ E_1 \left[ w^{(\theta, \tau)}_{A,1} + \frac{\eta_{M}}{2} (w^{(\theta, \tau)}_{P,1})^2 (1 - \pi_1)^2 y_{A,1}^2 \text{Var}_{1} (R_{\theta,2}) \right]
+ E_1 \left[ w^{(\theta, \tau)}_{B,1} + \frac{\eta_{M}}{2} (w^{(\theta, \tau)}_{P,1})^2 \pi_1^2 y_{B,1}^2 \text{Var}_{1} (R_{\tau,2}) \right]
- \frac{\eta_{P}}{2} \left[ (1 - \pi_1)^2 \text{Var}_{1} (R_{\tau,2}) + (1 - \pi_1)^2 (1 - y_{A,1})^2 \text{Var}_{1} (R_{\theta,2}) \right]
+ 2\pi_1 (1 - \pi_1) (1 - y_{B,1}) (1 - y_{A,1}) \text{Cov}_{1} (R_{\theta,2}, R_{\tau,2})
$$

Optimizing for an interior solution over $(\pi_1, y_{A,1}, y_{B,1})$ we get the described solution. Substitute the optimal $t = 1$ policies back into $\tilde{U}^{(\theta, \tau)}_{P,1}, U^{(\theta, \tau)}_{A,1}, U^{(\theta, \tau)}_{B,1}$ and denote them, respectively as $\tilde{U}^{(\theta, \tau)}_{P,1}, U^{(\theta, \tau)}_{A,1}, U^{(\theta, \tau)}_{B,1}$. After substituting and simplifying, this results in the expressions (1.21).

(c) Next, we consider the $t = 0$ optimal portfolio and performance fee policies. At $t = 0$, the Principal considers the problem,

$$
\hat{U}_{P,0} = \sup_{y_{A,0}, y_{B,0} \in [0, 1]} \sup_{\pi_0 \in \mathbb{R}} \mathbb{E}_0 [ \tilde{U}^{(\theta, \tau)}_{P,1} + \tilde{U}^{(\theta, \tau)}_{A,1} + \tilde{U}^{(\theta, \tau)}_{B,1} ] - \frac{\eta_{P}}{2} \text{Var}_{0} (\mathbb{E}_1 W^{(\theta, \tau)}_{P,2}) - \frac{\eta_{M}}{2} \text{Var}_{0} (\mathbb{E}_1 W^{(\theta, \tau)}_{A,2}) - \frac{\eta_{M}}{2} \text{Var}_{0} (\mathbb{E}_1 W^{(\theta, \tau)}_{B,2}).
$$

Firstly, we should note that after substituting in the optimal portfolio and fee policies $(\hat{\pi}_1, \hat{y}_{A,1}, \hat{y}_{B,1})$ and substituting them back and simplifying, we get (1.21). Where for the terms involved in $U_{A,1}$,

$$
u_A^{1N} := \eta_{P} \eta_{M} \text{Var}_{1} (R_{\theta,2}) \text{Var}_{1} (R_{\tau,2})^2 (\text{Var}_{1} (R_{\theta,2}) \text{Var}_{1} (R_{\tau,2}) - \text{Cov}_{1} (R_{\theta,2}, R_{\tau,2})^2)^2,$$

$$
u_A^{1D} := 2 \left[ - (\eta_{P} + \eta_{M}) \text{Var}_{1} (R_{\theta,2}) \text{Var}_{1} (R_{\tau,2}) (\text{Var}_{1} (R_{\theta,2}) + \text{Var}_{1} (R_{\tau,2}))
+ 2\eta_{M} \text{Var}_{1} (R_{\theta,2}) \text{Var}_{1} (R_{\tau,2}) \text{Cov}_{1} (R_{\theta,2}, R_{\tau,2})
+ \eta_{P} (\text{Var}_{1} (R_{\theta,2}) + \text{Var}_{1} (R_{\tau,2})) \text{Cov}_{1} (R_{\theta,2}, R_{\tau,2}) \right]^2
$$

86
and,

\[ u_{3N}^2 := \eta_p \text{Var}(R_{\theta,2})[\text{Var}(R_{\theta,2})\text{Var}(R_{\tau,2}) - \text{Cov}(R_{\theta,2}, R_{\tau,2})]^2 \]

\times \left[ \eta_p \text{Cov}(R_{\theta,2}, R_{\tau,2})^2[\text{Var}(R_{\theta,2})\mathbb{E}_1 R_{\tau,2} + \text{Var}(R_{\tau,2})\mathbb{E}_1 R_{\theta,2}] \\
+ \eta_m \text{Var}(R_{\theta,2})\text{Var}(R_{\tau,2})\text{Cov}(R_{\theta,2}, R_{\tau,2})\mathbb{E}_1 R_{\theta,2} + \mathbb{E}_1 R_{\tau,2}) \\
- (\eta_p + \eta_m)\text{Var}(R_{\theta,2})\text{Var}(R_{\tau,2})[\text{Var}(R_{\theta,2})\mathbb{E}_1 R_{\tau,2} + \text{Var}(R_{\tau,2})\mathbb{E}_1 R_{\theta,2}] \right],

\[ u_{3D}^2 := \left[ \eta_p \text{Cov}(R_{\theta,2}, R_{\tau,2})^2(\text{Var}(R_{\theta,2}) + \text{Var}(R_{\tau,2})) \\
+ 2\eta_m \text{Var}(R_{\theta,2})\text{Var}(R_{\tau,2})\text{Cov}(R_{\theta,2}, R_{\tau,2}) \\
- (\eta_p + \eta_m)\text{Var}(R_{\theta,2})\text{Var}(R_{\tau,2})(\text{Var}(R_{\theta,2}) + \text{Var}(R_{\tau,2})) \right]^2, \]

and,

\[ u_{4N}^3 := (\mathbb{E}_1 R_{\theta,2} - \mathbb{E}_1 R_{\tau,2}) \times \left[ \eta_p \text{Cov}(R_{\theta,2}, R_{\tau,2})^2 + \eta_m \text{Var}(R_{\theta,2})\text{Cov}(R_{\theta,2}, R_{\tau,2}) \\
- (\eta_p + \eta_m)\text{Var}(R_{\theta,2})\text{Var}(R_{\tau,2}) \right] \]

\times \left[ \eta_p \text{Cov}(R_{\theta,2}, R_{\tau,2})^2[2\text{Var}(R_{\tau,2})\mathbb{E}_1 R_{\theta,2} + \text{Var}(R_{\theta,2})(\mathbb{E}_1 R_{\theta,2} + \mathbb{E}_1 R_{\tau,2})] \\
+ \eta_m \text{Var}(R_{\theta,2})\text{Var}(R_{\tau,2})\text{Cov}(R_{\theta,2}, R_{\tau,2})(3\mathbb{E}_1 R_{\theta,2} + \mathbb{E}_1 R_{\tau,2}) \\
- (\eta_p + \eta_m)\text{Var}(R_{\theta,2})\text{Var}(R_{\tau,2})[2\text{Var}(R_{\theta,2})\mathbb{E}_1 R_{\theta,2} + (\mathbb{E}_1 R_{\theta,2} + \mathbb{E}_1 R_{\tau,2})\text{Var}(R_{\theta,2})] \right],

\[ u_{4D}^3 := 2\eta_m \left[ \eta_p(\text{Var}(R_{\theta,2}) + \text{Var}(R_{\tau,2}))[\text{Var}(R_{\theta,2})\text{Var}(R_{\tau,2}) - \text{Cov}(R_{\theta,2}, R_{\tau,2})^2] \\
+ \eta_m \text{Var}(R_{\theta,2})\text{Var}(R_{\tau,2})(\text{Var}(R_{\theta,2}) + \text{Var}(R_{\tau,2})) \right]^2. \]

For the terms involved in \( U_{B,1} \),

\[ u_{1B}^N := \frac{2}{\eta_p} \eta_m \text{Var}(R_{\theta,2})^2\text{Var}(R_{\tau,2})[\text{Var}(R_{\theta,2})\text{Var}(R_{\tau,2}) - \text{Cov}(R_{\theta,2}, R_{\tau,2})^2], \]

\[ u_{1D}^B := u_{A}^D, \]

and,

\[ u_{2B}^N := u_{A}^N, \]

\[ u_{2B}^D := u_{B}^D, \]

87
and,
\[
 u_{B}^{3N} := (E_1 R_{t,2} - E_1 R_{t,2}) \times \left[ \eta \text{Cov}_1(R_{\theta,2}, R_{t,2})^2 + \eta_M \text{Var}_1(R_{\theta,2}) \text{Cov}_1(R_{\theta,2}, R_{t,2}) \
- (\eta_p + \eta_M) \text{Var}_1(R_{\theta,2}) \text{Var}_1(R_{t,2}) \right] \\
\times \left[ \eta_p \text{Cov}_1(R_{\theta,2}, R_{t,2})^2[2 \text{Var}_1(R_{\theta,2})E_1 R_{t,2} + \text{Var}_1(R_{t,2})(E_1 R_{\theta,2} + E_1 R_{t,2})] \\
+ \eta_M \text{Var}_1(R_{\theta,2}) \text{Var}_1(R_{t,2}) \text{Cov}_1(R_{\theta,2}, R_{t,2})(E_1 R_{\theta,2} + 3E_1 R_{t,2}) \\
- (\eta_p + \eta_M) \text{Var}_1(R_{\theta,2}) \text{Var}_1(R_{t,2})[2 \text{Var}_1(R_{t,2})E_1 R_{\theta,2} + (E_1 R_{\theta,2} + E_1 R_{t,2}) \text{Var}_1(R_{t,2})] \right],
\]
\[
u_B := u_{A}^{3D}.
\]

For the terms involving \( \tilde{U}_{P,1} \),
\[
u_{P}^{1N} := \eta \eta_M^2 \text{Var}_1(R_{\theta,2})^2 \text{Var}_1(R_{t,2})[\text{Var}_1(R_{\theta,2}) \text{Var}_1(R_{t,2}) - \text{Cov}_1(R_{\theta,2}, R_{t,2})^2],
\]
\[
u_{P}^{1D} := 2 \left[ \eta_p \text{Var}_1(R_{\theta,2}) + \text{Var}_1(R_{t,2}) \right] \left[ \text{Var}_1(R_{\theta,2}) \text{Var}_1(R_{t,2}) - \text{Cov}_1(R_{\theta,2}, R_{t,2})^2 \right]
\]
\[
+ \eta_M \text{Var}_1(R_{\theta,2}) \text{Var}_1(R_{t,2}) \left[ \text{Var}_1(R_{\theta,2}) + (2 + E_1 R_{\theta,2}) \text{Var}_1(R_{t,2}) \right] \\
\times \left[ (2 + E_1 R_{t,2}) \text{Var}_1(R_{\theta,2}) + (2 + E_1 R_{\theta,2}) \text{Var}_1(R_{t,2}) \right]
\]
\[
+ \eta_M^2 \text{Var}_1(R_{\theta,2}) \text{Var}_1(R_{t,2}) \text{Var}_1(R_{\theta,2} - R_{t,2}) \times \\
[(1 + E_1 R_{t,2}) \text{Var}_1(R_{\theta,2}) + (1 + E_1 R_{\theta,2}) \text{Var}_1(R_{t,2}) - (2 + E_1 R_{\theta,2} + E_1 R_{t,2}) \text{Cov}_1(R_{\theta,2}, R_{t,2})],
\]
\[
u_{P}^{2D} := u_{A}^{2D}.
\]

Furthermore, the \( t = 1 \) expectation of the \( t = 2 \) wealth, after substituting in the \( t = 1 \) optimal policies are given in (27). The terms involved in \( W_{A,2} \) are,
\[
w_{A}^{1N} := \eta \text{E}_1[R_{\theta,2}] \text{Var}_1(R_{t,2})[\text{Var}_1(R_{\theta,2}) \text{Var}_1(R_{t,2}) - \text{Cov}_1(R_{\theta,2}, R_{t,2})^2],
\]
\[
w_{A}^{1D} := \eta_p \text{Cov}_1(R_{\theta,2}, R_{t,2})^2[\text{Var}_1(R_{\theta,2}) + \text{Var}_1(R_{t,2})] \\
+ 2\eta_M \text{Var}_1(R_{\theta,2}) \text{Var}_1(R_{t,2}) \text{Cov}_1(R_{\theta,2}, R_{t,2}) \\
- (\eta_p + \eta_M) \text{Var}_1(R_{\theta,2}) \text{Var}_1(R_{t,2})[\text{Var}_1(R_{\theta,2}) + \text{Var}_1(R_{t,2})],
\]
\[
w_{A}^{2N} := \text{E}_1[R_{\theta,2}](\text{E}_1 R_{\theta,2} - \text{E}_1 R_{t,2}) \left[ \eta_p \text{Cov}_1(R_{\theta,2}, R_{t,2})^2 \\
+ \eta_M \text{Var}_1(R_{t,2}) \text{Cov}_1(R_{\theta,2}, R_{t,2}) - (\eta_p + \eta_M) \text{Var}_1(R_{\theta,2}) \text{Var}_1(R_{t,2}) \right],
\]
\[
w_{A}^{2D} := \eta_M w_A^{1D}.
\]
Proof of Proposition 7.1. This is clearly a special case of Proposition 11.1.

J Proofs for Section 11

Proof of Proposition 11.1. (a) For any given contract \((x_A, \{y_{A,0}, y_{A,1}\})\) and any investment strategy pairs \((\theta, \tau) \in \mathcal{S}\), it is clear that Manager C’s optimal portfolio choice would be equivalent to the form as given in Proposition 11.1.

(b) Recall that the Principal wants to induce the strategy pair \((\theta_H, \tau_H)\). Thus, by binding the individual rationality constraint \((14.14)\), we obtain the optimal fixed fee form as per Proposition 11.1.

(c) Now we consider the Principal’s second best optimization problem. Recall that the Principal wants to implement the strategy pair \((\theta_H, \tau_H)\), but Manager C could deviate to \((\theta, \tau) \in \mathcal{S} \setminus \{\theta_H, \tau_H\}\). As it was in the case in static centralized delegation, although we have three incentive compatibility constraints, clearly only one of them will bind, while the others will be slack. Following Bellman (1950), let us introduce the Lagrange multipliers \(\lambda^{(\theta, \tau)}_C, \lambda^{(\theta_H, \tau_H)}_C, \lambda^{(\theta, \tau)}_C\) into our dynamic optimization problem. Recall also the notations from the proof of Proposition 11.1 where we had denoted \(U_C^{(\theta, \tau)}\) as Manager C’s time \(t\)
continuation value for implementing strategy \((\theta, \tau)\) along the optimal portfolio strategies \(\hat{\psi}_{t, (\theta, \tau)}\); note and recall also that \(\hat{U}_{C, t}^{(\theta, \tau)}\) is still a function of the performance fees \((y_0, y_1)\). Also note that using our notations, we can rewrite the incentive compatibility constraints \((\text{J.1})\) as,

\[
-2c + U_{C, 0}^{(\theta_H, \tau_H)} \geq -c + U_{C, 0}^{(\theta_H, \tau_L)},
\]

\[
-2c + U_{C, 0}^{(\theta_L, \tau_H)} \geq -c + U_{C, 0}^{(\theta_L, \tau_L)},
\]

\[
-2c + U_{C, 0}^{(\theta_H, \tau_L)} \geq U_{C, 0}^{(\theta_L, \tau_L)}.
\]

Thus at \(t = 0\), recalling the objective function form \((\text{DynCon})\), substituting in the optimal fixed fee form, and incorporating the rewritten form of the incentive compatibility constraint, the Principal considers the sequence of problems indexed by \(\lambda_C := (\lambda_C^{(\theta_H, \tau_H)}, \lambda_C^{(\theta_L, \tau_H)}, \lambda_C^{(\theta_L, \tau_L)}) \in \mathbb{R}^3\),

\[
\sup_{y_{A, 0}, y_{A, 1} \in [0, 1]} -x_A + E_0[W_{cP, 2}^{(\theta_H, \tau_H)}] - \eta^M \frac{2}{2} \text{Var}_0(W_{cP, 2}^{(\theta_H, \tau_H)}) - \lambda_C^{(\theta_H, \tau_H)} [\hat{U}^{(\theta_H, \tau_H)}_{C, 0} - (-c + \hat{U}^{(\theta_H, \tau_H)}_{C, 0})]
\]

\[
- \lambda_C^{(\theta_L, \tau_H)} [\hat{U}^{(\theta_L, \tau_H)}_{C, 0} - (-c + \hat{U}^{(\theta_L, \tau_H)}_{C, 0})] - \lambda_C^{(\theta_L, \tau_L)} [\hat{U}^{(\theta_L, \tau_L)}_{C, 0} - (-c + \hat{U}^{(\theta_L, \tau_L)}_{C, 0})]
\]

\[
\sup_{y_{A, 0}, y_{A, 1} \in [0, 1]} -2c + E_0[W_{cP, 2}^{(\theta_H, \tau_H)}] - \eta^M \frac{2}{2} \text{Var}_0(W_{cP, 2}^{(\theta_H, \tau_H)})
\]

\[
+ E_0[W_{cP, 2}^{(\theta_L, \tau_H)}] - \eta^M \frac{2}{2} \text{Var}_0(W_{cP, 2}^{(\theta_L, \tau_H)}) - \lambda_C^{(\theta_L, \tau_H)} [U_{C, 0}^{(\theta_L, \tau_H)} - (-2c + U_{C, 0}^{(\theta_H, \tau_H)})]
\]

\[
- \lambda_C^{(\theta_L, \tau_L)} [U_{C, 0}^{(\theta_L, \tau_L)} - (-2c + U_{C, 0}^{(\theta_L, \tau_L)})]
\]

\[
- \lambda_C^{(\theta_L, \tau_L)} [U_{C, 0}^{(\theta_L, \tau_L)} - (-c + U_{C, 0}^{(\theta_L, \tau_L)})]
\]

\[
= -2c + \hat{U}_{cP, 0}^{\lambda_C}.
\]

In particular, again motivated by the dynamic programming principle of Section \(\text{Q}\), let’s recall the notations \((\text{LS})\). In particular, this implies that if we focus on the value function term \(U_{C, 0}^{\lambda_C}\), and denoting \(U_{cP, 0}^{\lambda_C}\) as the associated objective function, we can consider the recursive relationship,

\[
\hat{U}_{cP, 1}^{\lambda_C} = \sup_{y_{A, 1} \in [0, 1]} U_{C, 1}^{(\theta_H, \tau_H)} + \hat{U}_{cP, 1}^{(\theta_H, \tau_H)} - \lambda_C^{(\theta_L, \tau_L)} [U_{C, 1}^{(\theta_L, \tau_L)} - (-2c + U_{C, 1}^{(\theta_H, \tau_H)})]
\]

\[
- \lambda_C^{(\theta_L, \tau_H)} [U_{C, 0}^{(\theta_L, \tau_H)} - (-c + U_{C, 0}^{(\theta_L, \tau_H)})] - \lambda_C^{(\theta_L, \tau_L)} [\hat{U}_{C, 0}^{(\theta_L, \tau_L)} - (-c + \hat{U}_{C, 0}^{(\theta_L, \tau_L)})]
\]

\[
\hat{U}_{cP, 0}^{\lambda_C} = \sup_{y_{A, 0} \in [0, 1]} E_0[\hat{U}_{cP, 1}^{\lambda_C}] - \eta^M \frac{2}{2} \text{Var}_0(E_1 W_{cP, 2}^{(\theta_H, \tau_H)}) - \eta^M \frac{2}{2} \text{Var}_0(W_{cP, 2}^{(\theta_H, \tau_H)})
\]

\[
- \lambda_C^{(\theta_H, \tau_L)} [\eta^M \frac{2}{2} \text{Var}_0(E_1 W_{cP, 2}^{(\theta_H, \tau_H)}) - (-\eta^M \frac{2}{2} \text{Var}_0(E_1 W_{cP, 2}^{(\theta_H, \tau_H)}))]
\]

\[
- \lambda_C^{(\theta_H, \tau_L)} [\eta^M \frac{2}{2} \text{Var}_0(E_1 W_{cP, 2}^{(\theta_H, \tau_H)}) - (-\eta^M \frac{2}{2} \text{Var}_0(E_1 W_{cP, 2}^{(\theta_H, \tau_H)}))]
\]

\[
- \lambda_C^{(\theta_H, \tau_L)} [\eta^M \frac{2}{2} \text{Var}_0(E_1 W_{cP, 2}^{(\theta_H, \tau_H)}) - (-\eta^M \frac{2}{2} \text{Var}_0(E_1 W_{cP, 2}^{(\theta_H, \tau_H)}))]
\]

\[
= \sup_{y_{A, 0} \in [0, 1]} U_{C, 0}^{(\theta_H, \tau_H)} + \hat{U}_{cP, 0}^{(\theta_H, \tau_H)} - \lambda_C^{(\theta_L, \tau_H)} [U_{C, 0}^{(\theta_L, \tau_H)} - (-2c + U_{C, 0}^{(\theta_H, \tau_H)})]
\]

\[
- \lambda_C^{(\theta_L, \tau_H)} [U_{C, 0}^{(\theta_L, \tau_H)} - (-c + U_{C, 0}^{(\theta_L, \tau_H)})]
\]

\[
- \lambda_C^{(\theta_L, \tau_L)} [U_{C, 0}^{(\theta_L, \tau_L)} - (-c + U_{C, 0}^{(\theta_L, \tau_L)})]
\]

\[
= -2c + \hat{U}_{cP, 0}^{\lambda_C}.
\]

where we have denoted \(\hat{U}_{C, 1}^{(\theta, \tau)}\) as the value of Manager \(C\)’s continuation value after substituting in the optimal \(t = 1\) performance fee \(\hat{y}_1\) into \(U_{C, 1}^{(\theta, \tau)}\), and likewise for \(\hat{U}_{C, 0}^{(\theta, \tau)}\). But as it was argued in the static centralized delegation case, we know that at equilibrium, if a binding solution exists, only one of the three incentive compatibility constraints will bind. And thus, to further save on notations, if \((\theta^b, x^b)\) is the pair of deviant strategies associated with the binding incentive compatibility constraints, we allow
that Lagrange multiplier $\lambda_C^{(b, \tau)}$ to be nonzero, and set the remaining other two to zero. Then after this simplification, we get the displayed equations we see in the proposition.

And also note that, with some abuse of notations, the wealth values $W_C^{(b, \tau)}$ and $W_{C,2}^{(b, \tau)}$ in (J.3) and (J.4) are different; in (J.3), after conditioning on the $t = 1$ realized wealths of both the Principal and Manager $C$, those $t = 2$ wealth terms are only a function of $y_{A,0}$; in (J.4), the resulting optimal fee $\hat{y}_{A,1}$ has been substituted in, and since at $t = 0$, the $t = 2$ wealth is a function of $t = 1$ wealth, those $t = 2$ wealth terms are only a function of $y_{A,0}$. Finally in (J.5), we simply reuse the notations in the proof of Proposition E.1, and in particular noting that in $U_C^{(b, \tau)}$ and $\tilde{U}_{C,0}^{(b, \tau)}$ is the $t = 0$ wealth value for Manager $C$ and the Principal, after substituting for Manager $C$’s optimal portfolio policy for any arbitrary contract; hence in the dynamic programming formulation, $U_C^{(b, \tau)}$ and $\tilde{U}_{C,0}^{(b, \tau)}$ are only a function of $y_{A,0}$.

Indeed, taking first order conditions, which is both sufficient and necessary for optimization here, the optimal $t = 1$ performance fees are (F.1).

(d) With the $t = 1$ optimal performance fees (F.1), we substitute it back to the $t = 1$ continuation value of the Principal and Manager $C$; that is, and recalling the optimal portfolio policy form $\hat{\psi}_{1, (b, \tau)} := \hat{\psi}_{1, (b, \tau)}(y_{A,1})$ of (E.1), we get (F.2).

(e) Substituting those expressions back into (F.3), and further substituting in the budget constraints (D.1a) and (D.1c), we can find an optimizer $\hat{y}_{A,0}^{\lambda_C} \in [0, 1]$.

(f) Now, once $\hat{y}_{A,0}^{\lambda_C}$ has been found, then we obtain the $t = 0$ value function $\hat{U}_{C,0}^{\lambda_C}$ for the Principal. At this point, we still need to choose the optimal $\lambda_C$. Assuming that the conditions for Strong Duality Theorem holds, we then the optimal $\lambda_C$ is precisely the solution to (F.4).

\[ \Box \]

Proof of Proposition 8.1. The proof is quite analogous to that of Proposition E.1 but repeated here for completeness.

(a) By the same arguments as in Proposition E.1, the individual rationality constraints (6.1g) and (6.1h) will bind. This pins down the form of the fixed fees, and indeed it has the same form as the first best form of Proposition E.1.

(b) We now consider the Principal’s second best optimization problem. Again, following Bellman (1957), we introduce the Lagrange multipliers $\lambda_A, \lambda_B$ associated with the incentive compatibility constraints (6.1i), (6.1j), respectively. We will also recycle the notations from the proof of Proposition E.1 of the first best case. Note that using those notations, we can write the incentive compatibility constraints (6.1i) and (6.1j), respectively, as,

\[ -c + U_{A,0}^{(b, \tau)} \geq U_{A,0}^{(b, \tau)} , \]  

\[ -c + U_{B,0}^{(b, \tau)} \geq U_{B,0}^{(b, \tau)} . \]  

(J.6a)  

(J.6b)

Thus, the Principal’s optimization problem from (DynDec), binding the individual rationality con-
By the dynamic programming principle of Section I.1, and also recall the analogous argument in the proof of Proposition I.1, we are lead to consider the following sequence of problems:

\[
\hat{U}_{P,0}^{\lambda_A,\lambda_B} = \sup_{\pi_0 \in \mathbb{R}, y_{A,0}, y_{B,1} \in [0,1]} \hat{U}_{P,0}^{(\theta_{\pi, \gamma})} + \hat{U}_{A,1}^{(\theta_{\pi, \gamma})} + \hat{U}_{B,1}^{(\theta_{\pi, \gamma})}
\]

\[
\begin{align*}
- \lambda_A \left( \hat{U}_{A,0}^{(\theta_{\pi, \gamma})} - (-c + U_{A,0}^{(\theta_{\pi, \gamma})}) \right) &- \lambda_B \left( \hat{U}_{B,0}^{(\theta_{\pi, \gamma})} - (-c + U_{B,0}^{(\theta_{\pi, \gamma})}) \right), \\
= &\sup_{y_{A,0}, y_{A,1} \in [0,1], \pi_0} \sup_{y_{B,0}, y_{B,1} \in [0,1]} -2c + \hat{U}_{A,0}^{(\theta_{\pi, \gamma})} - (-c + U_{A,0}^{(\theta_{\pi, \gamma})})
\end{align*}
\]

\[
\hat{U}_{P,1}^{\lambda_A,\lambda_B} = \sup_{\pi_1 \in \mathbb{R}, y_{A,1} \in [0,1]} \hat{U}_{P,1}^{(\theta_{\pi, \gamma})} + \hat{U}_{A,1}^{(\theta_{\pi, \gamma})} + \hat{U}_{B,1}^{(\theta_{\pi, \gamma})}
\]

\[
\begin{align*}
- \lambda_A \left( \hat{U}_{A,1}^{(\theta_{\pi, \gamma})} - (-c + U_{A,1}^{(\theta_{\pi, \gamma})}) \right) &- \lambda_B \left( \hat{U}_{B,1}^{(\theta_{\pi, \gamma})} - (-c + U_{B,1}^{(\theta_{\pi, \gamma})}) \right), \\
= &\sup_{y_{A,1}, y_{A,2} \in [0,1], \pi_1} \sup_{y_{B,0}, y_{B,1} \in [0,1]} \mathbb{E}_0[\hat{U}_{P,1}^{\lambda_A,\lambda_B}] - \frac{\eta_M}{2} \text{Var}_0(\mathbb{E}_1 W_{A,2}^{(\theta_{\pi, \gamma})}) + \frac{\eta_M}{2} \text{Var}_0(\mathbb{E}_1 W_{B,2}^{(\theta_{\pi, \gamma})})
\end{align*}
\]

\[
\hat{U}_{P,0}^{\lambda_A,\lambda_B} = \sup_{\pi_0 \in \mathbb{R}, y_{A,0}, y_{B,0} \in [0,1]} \hat{U}_{P,0}^{(\theta_{\pi, \gamma})} + \hat{U}_{A,1}^{(\theta_{\pi, \gamma})} + \hat{U}_{B,1}^{(\theta_{\pi, \gamma})}
\]

\[
\begin{align*}
- \lambda_A \left( \hat{U}_{A,0}^{(\theta_{\pi, \gamma})} - (-c + U_{A,0}^{(\theta_{\pi, \gamma})}) \right) &- \lambda_B \left( \hat{U}_{B,0}^{(\theta_{\pi, \gamma})} - (-c + U_{B,0}^{(\theta_{\pi, \gamma})}) \right), \\
= &\sup_{y_{A,0}, y_{A,1} \in [0,1], \pi_0} \sup_{y_{B,0}, y_{B,1} \in [0,1]} \mathbb{E}_0[\hat{U}_{P,0}^{\lambda_A,\lambda_B}] - \frac{\eta_M}{2} \text{Var}_0(\mathbb{E}_1 W_{A,2}^{(\theta_{\pi, \gamma})}) + \frac{\eta_M}{2} \text{Var}_0(\mathbb{E}_1 W_{B,2}^{(\theta_{\pi, \gamma})})
\end{align*}
\]

Considering the optimizing problem (I.7), and optimizing for the 1 policies \((\pi_1, y_{A,1}, y_{B,1})\), we
arrive at (8.1), (8.2) and (8.3). In particular,

\[\hat{\theta}_{1,1}^{(J.9a)} = w_{P_1}^{(0_{\theta_1,\tau_1})} \left[ \left(-w_{P_1}^{(0_{\theta_1,\tau_1})} \right)^2 \eta_p + \eta_M(\lambda_A - \lambda_B) + \left(w_{P_1}^{(0_{\theta_1,\tau_1})} - w_{P_1}^{(0_{\theta_1,\tau_1})} \right) \left(w_{P_1}^{(0_{\theta_1,\tau_1})} + w_{P_1}^{(0_{\theta_1,\tau_1})} \right) \eta_M \lambda_A \lambda_B \right] \mu \\
+ \eta_M \left(-w_{P_1}^{(0_{\theta_1,\tau_1})} \right)^2 \lambda_A + \left(w_{P_1}^{(0_{\theta_1,\tau_1})} \right)^2 (1 + \lambda_A) \left(-w_{P_1}^{(0_{\theta_1,\tau_1})} \right)^2 \eta_M \lambda_B + \left(w_{P_1}^{(0_{\theta_1,\tau_1})} \right)^2 (\eta_p + \eta_M(1 + \lambda_B)) \sigma^2 \]

\[+ w_{P_1}^{(0_{\theta_1,\tau_1})} \left(-w_{P_1}^{(0_{\theta_1,\tau_1})} \right)^2 \eta_M \lambda_A + \left(w_{P_1}^{(0_{\theta_1,\tau_1})} \right)^2 (\eta_p + \eta_M(1 + \lambda_A)) \right) \lambda_B \mu_{\tau_1} \]

\[+ \left(\eta_M \right)^2 \left[ \left(-w_{P_1}^{(0_{\theta_1,\tau_1})} \right)^2 \lambda_A - \left(w_{P_1}^{(0_{\theta_1,\tau_1})} \right)^2 \lambda_B + \left(w_{P_1}^{(0_{\theta_1,\tau_1})} \right)^2 (2 + \lambda_A + \lambda_B) \right) \rho_{\theta_1,\tau_1} \right]. \] (J.8a)

And,

\[\hat{\theta}_{1,1}^{(J.9b)} = -w_{P_1}^{(0_{\theta_1,\tau_1})} \lambda_A \mu_{\tau_1} \left( -2\left(\left(-w_{P_1}^{(0_{\theta_1,\tau_1})} \right)^2 \eta_M \lambda_B + \left(w_{P_1}^{(0_{\theta_1,\tau_1})} \right)^2 (\eta_p + 2\eta_M(1 + \lambda_B)) + \left(w_{P_1}^{(0_{\theta_1,\tau_1})} \right)^2 \eta_M \rho_{\theta_1,\tau_1} \right) \right) \]

\[+ w_{P_1}^{(0_{\theta_1,\tau_1})} \left[ \left(-w_{P_1}^{(0_{\theta_1,\tau_1})} \right)^2 \eta_M \lambda_A + \left(w_{P_1}^{(0_{\theta_1,\tau_1})} \right)^2 (\eta_p + \eta_M(1 + \lambda_A)) \right] \left(-w_{P_1}^{(0_{\theta_1,\tau_1})} \right)^2 \lambda_B + \left(w_{P_1}^{(0_{\theta_1,\tau_1})} \right)^2 (1 + \lambda_B) \sigma^2 \]

\[+ \eta_M \left(-w_{P_1}^{(0_{\theta_1,\tau_1})} \right)^2 \lambda_B + \left(w_{P_1}^{(0_{\theta_1,\tau_1})} \right)^2 (1 + \lambda_B) \right) \rho_{\theta_1,\tau_1} \right]. \] (J.9a)

And,

\[\hat{\theta}_{1,1}^{(J.9b)} = w_{P_1}^{(0_{\theta_1,\tau_1})} \left[ \left(\left(-w_{P_1}^{(0_{\theta_1,\tau_1})} \right)^2 \eta_M \lambda_A + \left(w_{P_1}^{(0_{\theta_1,\tau_1})} \right)^2 (\eta_p + \eta_M(1 + \lambda_A)) \right] \left(-w_{P_1}^{(0_{\theta_1,\tau_1})} \right)^2 \lambda_B + \left(w_{P_1}^{(0_{\theta_1,\tau_1})} \right)^2 (1 + \lambda_B) \sigma^2 \]

\[+ \eta_M \left(-w_{P_1}^{(0_{\theta_1,\tau_1})} \right)^2 \lambda_B + \left(w_{P_1}^{(0_{\theta_1,\tau_1})} \right)^2 (1 + \lambda_B) \right) \rho_{\theta_1,\tau_1} \right]. \] (J.9b)
And,

\[ y_{B,1}^{\theta_{h,\tau_h}, \lambda_A, \lambda_B} = w_{P,1}^{(\theta_{h,\tau_h})} \left[ -2(w_{P,1}^{(\theta_{h,\tau_h})})^2 \eta_M \lambda_A \lambda_B + (w_{P,1}^{(\theta_{h,\tau_h})})^2 (-\eta_P \lambda_A + (\eta_P + 2\eta_M(1 + \lambda_A)) \lambda_B) \right] \mu \\
+ w_{P,1}^{(\theta_{h,\tau_h})} \eta_P \eta_M \left( -(w_{P,1}^{(\theta_{h,\tau_h})})^2 \lambda_A + (w_{P,1}^{(\theta_{h,\tau_h})})^2 (1 + \lambda_A) \sigma^2 \right) \\
- w_{P,1}^{(\theta_{h,\tau_h})} \left[ -2(w_{P,1}^{(\theta_{h,\tau_h})})^2 \eta_M \lambda_A + (w_{P,1}^{(\theta_{h,\tau_h})})^2 (\eta_P + 2\eta_M(1 + \lambda_A)) \right] \lambda_B \mu_{\tau_h} \\
+ (w_{P,1}^{(\theta_{h,\tau_h})})^2 \eta_P \left[ w_{P,1}^{(\theta_{h,\tau_h})} (-\lambda_A + \lambda_B) \mu + \eta_M \left( -(w_{P,1}^{(\theta_{h,\tau_h})})^2 \lambda_A + (w_{P,1}^{(\theta_{h,\tau_h})})^2 (1 + \lambda_A) \right) \sigma^2 - w_{P,1}^{(\theta_{h,\tau_h})} \lambda_B \mu_{\tau_h} \right] \rho_{\theta_{h,\tau_h}} \\
+ (w_{P,1}^{(\theta_{h,\tau_h})})^2 \eta_P \lambda_A \mu_{\theta_{h,\tau_h}} (1 + \rho_{\theta_{h,\tau_h}}) \right], \quad (J.10a) \]

\[ y_{B,1}^{D,\lambda_A, \lambda_B} = w_{P,1}^{(\theta_{h,\tau_h})} \left[ -\left( w_{P,1}^{(\theta_{h,\tau_h})} \right)^2 (\eta_P + \eta_M)(\lambda_A - \lambda_B) + (w_{P,1}^{(\theta_{h,\tau_h})} - w_{P,1}^{(\theta_{h,\tau_h})})(w_{P,1}^{(\theta_{h,\tau_h})} + w_{P,1}^{(\theta_{h,\tau_h})}) \eta_M \lambda_A \lambda_B \right] \mu \\
+ \eta_M \left[ -(w_{P,1}^{(\theta_{h,\tau_h})})^2 \lambda_A + (w_{P,1}^{(\theta_{h,\tau_h})})^2 (1 + \lambda_A) \right] \left[ -(w_{P,1}^{(\theta_{h,\tau_h})})^2 \eta_M \lambda_B + (w_{P,1}^{(\theta_{h,\tau_h})})^2 (\eta_P + \eta_M(1 + \lambda_B)) \right] \sigma^2 \\
- w_{P,1}^{(\theta_{h,\tau_h})} \left[ -(w_{P,1}^{(\theta_{h,\tau_h})})^2 \eta_M \lambda_A + (w_{P,1}^{(\theta_{h,\tau_h})})^2 (\eta_P + \eta_M(1 + \lambda_A)) \right] \lambda_B \mu_{\theta_{h,\tau_h}} \\
+ (w_{P,1}^{(\theta_{h,\tau_h})})^2 \eta_P \left[ w_{P,1}^{(\theta_{h,\tau_h})} (-\lambda_A + \lambda_B) \mu + \eta_M \left( -(w_{P,1}^{(\theta_{h,\tau_h})})^2 \lambda_A + (w_{P,1}^{(\theta_{h,\tau_h})})^2 (1 + \lambda_A) \right) \sigma^2 - w_{P,1}^{(\theta_{h,\tau_h})} \lambda_B \mu_{\tau_h} \right] \rho_{\theta_{h,\tau_h}} \\
+ w_{P,1}^{(\theta_{h,\tau_h})} \lambda_A \mu_{\theta_{h,\tau_h}} \left[ -(w_{P,1}^{(\theta_{h,\tau_h})})^2 \eta_M \lambda_B + (w_{P,1}^{(\theta_{h,\tau_h})})^2 (\eta_P + \eta_M(1 + \lambda_B)) + (w_{P,1}^{(\theta_{h,\tau_h})})^2 \eta_P \rho_{\theta_{h,\tau_h}} \right]. \quad (J.10b) \]

(c) This is just to gather some results in preparation for the next subpart by substituting in the aforementioned \( t = 1 \) optimal policies.

(d) This is equivalent to (L.74).

(e) This is immediate by the Strong Duality theorem.

\[ \blacksquare \]