Appendix C to Hellwig and Schmidt: Discrete-Time Approximations of the Holmström-Milgrom Brownian-Motion Model of Intertemporal Incentive Provision (not to be published)

In this appendix, we provide a detailed proof of Proposition 4. We begin with a few remarks on the mathematical structure underlying the analysis. As before, we let C^N be the space of continuous functions from $[0,1]$ into $I\!R^N$, endowed with the topology of uniform convergence, and for a given set $\hat{K} \subset I\!R^N$, we let $C^N_{\hat{K}}$ be the subspace of those functions $F \in C^N$ that have a Radon-Nikodym derivative f_F taking values in $\hat{K}$, so the representation

$$F(t) = \int_0^t f_F(t')dt' \tag{C.1}$$

is valid for some function $f_F \in L_1([0,1],\hat{K})$, where $L_1([0,1],\hat{K})$ is the space of (equivalence classes of almost everywhere equal) measurable functions from the time interval $[0,1]$ into the set $\hat{K}$, endowed with the usual L_1- norm.

Heuristically, the function $f_F(\cdot)$ in (C.1) is the time path of controls $f_F(t) \in \hat{K}$ which generates the *cumulative* control path $F(\cdot)$. Thus, (C.1) establishes a one-to-one relation between control paths and cumulative control paths, between $L_1([0,1],\hat{K})$ and $C^N_{\hat{K}}$. Given this relation, the cumulative total effort cost associated with a cumulative control path $F(\cdot)$ is defined as

$$\Gamma(F) = \int_0^1 \hat{c}(f_F(t))dt. \tag{C.2}$$

In terms of economic substance, it does not make any difference whether we study the agent's problem in terms of control paths or in terms of cumulative control paths. In terms of mathematics, the following lemma shows that if $\hat{K}$ is compact and convex, then $C^N_{\hat{K}}$ is compact and the function $\Gamma : C^N_{\hat{K}} \to I\!R$ is lower semi-continuous. In contrast, $L_1([0,1],\hat{K})$ is not compact, but cumulative total effort cost is continuous on $L_1([0,1],\hat{K})$. The net effect of these two considerations will be that it is easier to work with cumulative control paths rather than control paths, i.e., with the space $C^N_{\hat{K}}$ rather than $L_1([0,1],\hat{K})$.

Lemma C.1 *Let $\hat{K}$ be a compact and convex subset of $I\!R^N$. Then $C^N_{\hat{K}}$ is compact and the function $\Gamma : C^N_{\hat{K}} \rightarrow I\!R$ that is defined by (C.2) is lower semi-continuous.*

Proof: Let $\|.\|$ denote the Euclidean norm on $I\!R^N$, and let $\bar{k} := \max_{\mu \in \hat{K}} \|\mu\|$. For any $f \in C^N_{\hat{K}}$, one obviously has $\|F(t') - F(t'')\| \leq \bar{k} \; |t' - t''|$ for all t' and t''. Since (C.1) implies $F(0) = 0$, one also has $\|F(t)\| \leq \bar{k}$ for all t. By the Arzelà-Ascoli theorem (Billingsley, 1968, p.221), it follows that the closure of $C^N_{\hat{K}}$ is compact. To establish compactness, it is therefore sufficient to show that $C^N_{\hat{K}}$ is closed.

For this purpose, we show that any function $F \in C^N$ that is a limit of a sequence $\{F^n\}$ of elements of $C^N_{\hat{K}}$ must itself be an element of $C^N_{\hat{K}}$. As a limit of Lipschitz functions with the common Lipschitz constant $\bar{k}$, F must itself be Lipschitzian and hence absolutely continuous. By the Radon-Nikodym theorem, it follows that F satisfies (C.1) for some function $f_F : [0, 1] \rightarrow I\!R^N$. To prove that $F \in C^N_{\hat{K}}$ it suffices to show that $f_F(t) \in \hat{K}$ for almost all t.

By Lebesgue's theorem on the differentiation of a function of bounded variation, for almost all $t \in [0, 1]$, we have

$$f_F(t) = \lim_{h \rightarrow 0} \frac{F(t+h) - F(t-h)}{2h}.$$

By the specification of F as a limit of the sequence $\{F^n\}$, it follows that for almost all $t \in [0, 1]$,

$$f_F(t) = \lim_{h \rightarrow 0} \lim_{n \rightarrow \infty} \frac{F^n(t+h) - F^n(t-h)}{2h} = \lim_{h \rightarrow 0} \lim_{n \rightarrow \infty} q^n(t+h, t-h),$$

where

$$q^n(t+h, t-h) := \frac{\int_{t-h}^{t+h} f_{F^n}(t') dt'}{2h}.$$

Given that $\hat{K}$ is convex and $q^n(t+h, t-h)$ is an average of $f_{F^n}(t')$ for $t' \in [t+h, t-h]$, we have $q^n(t+h, t-h) \in \hat{K}$. Given that $\hat{K}$ is also compact, it follows that $f_F(t) \in \hat{K}$ for almost all t, and hence that $F \in C^N_{\hat{K}}$.

To prove that the function $\Gamma(.)$ is lower semi-continuous, let $\{F^n\}$ be any sequence in $C^N_{\hat{K}}$ that converges to a limit F, and consider the associated sequence $\{\Gamma(F^n)\}$. For any n and $h = 1, \frac{1}{2}, \ldots$, (C.2) yields

$$\Gamma(F^n) \;\; = \;\; \sum_{i=1}^{1/h} h \, \frac{\int_{(i-1)h}^{ih} \hat{c}(f_{F^n}(t)) dt}{h} \geq \sum_{i=1}^{1/h} h \; \hat{c} \left(\frac{\int_{(i-1)h}^{ih} f_{F^n}(t) dt}{h} \right)$$

$$= \sum_{i=1}^{1/h} h\, \hat{c}\left(\frac{F^n(ih) - F^n((i-1)h)}{h}\right),$$

where we use the fact that $\hat{c}(.)$ is a convex function. Given the convergence of $F^n(ih)$ and $F^n((i-1)h)$ to $F(ih)$ and $F((i-1)h)$, it follows that

$$\begin{aligned} \lim_{\bar{n}\to\infty} \inf_{n\geq\bar{n}} \Gamma(F^n) &\geq \sup_h \sum_{i=1}^{1/h} h\, \hat{c}\left(\frac{F(ih) - F((i-1)h)}{h}\right) \\ &\geq \lim_{h\to 0} \sum_{i=1}^{1/h} h\, \hat{c}\left(\frac{F(ih) - F((i-1)h)}{h}\right) \\ &= \int_0^1 \hat{c}(f_F(t))dt = \Gamma(F), \end{aligned}$$

where we again use the fact that for almost all t, $f_F(t) = \lim_{h\to 0}[F([t/h]h + h) - F([t/h]h)]/h$. *Q.E.D.*

Any control processes $\mu^\Delta(.)$, $\mu(.)$ in the discrete-time models and the continuous-time model can be thought of as random variables on some underlying probability space $(\Omega, \mathcal{F}, \nu)$ with values in $L_1([0,1], \hat{K})$. The associated *cumulative* control processes $M^\Delta(.)$ and $M(.)$ can be thought of as random variables on $(\Omega, \mathcal{F}, \nu)$ with values in $C^N_{\hat{K}}$. Finally, as discussed in Appendix B, the associated processes $X^\Delta(.)$ and $X(.)$ of cumulative deviations from the means are treated as random variables on $(\Omega, \mathcal{F}, \nu)$ with values in C^N.

The control processes $\mu^\Delta(.)$, $\mu(.)$ in the discrete-time models and the continuous-time model are determined by the agent's *control strategies* and the cumulative-deviations processes in these models. As discussed in Appendix B, a control strategy in the discrete-time model with period length Δ corresponds to a sequence $\{\hat{\mu}^{\Delta,\tau}\}_{\tau=1}^m$ such that for any τ, $\hat{\mu}^{\Delta,\tau}(.)$ is a function that indicates how the control chosen in period τ depends on the profit levels $\tilde{\pi}^{\Delta,1}, ..., \tilde{\pi}^{\Delta,\tau-1}$ in periods τ' prior to τ. Given that there is a one-to-one relation between the profit level $\tilde{\pi}^{\Delta,\tau}$ in any period τ and the vector $\tilde{X}^{\Delta,\tau} = (\tilde{X}_1^{\Delta,\tau}, ..., \tilde{X}_N^{\Delta,\tau})$ of increments of the cumulative-deviations process from period $\tau - 1$ to period τ, we can equivalently think of a control strategy in the discrete-time model with period length Δ as a sequence $\{\breve{\mu}^{\Delta,\tau}\}_{\tau=1}^m$ of functions $\breve{\mu}^{\Delta,\tau}$ of $\tilde{X}^{\Delta,1}, ..., \tilde{X}^{\Delta,\tau-1}$. Given such a sequence, the corresponding *cumulative-control strategy* is a sequence $\{\breve{M}^{\Delta,\tau}\}_{\tau=1}^m$ of functions $\breve{M}^{\Delta,\tau}$ of $\tilde{X}^{\Delta,1}, ..., \tilde{X}^{\Delta,\tau-1}$ such that for any Δ and $\tau = 2, ...m$,

$$\breve{M}^{\Delta,\tau}(\tilde{X}^{\Delta,1}, ..., \tilde{X}^{\Delta,\tau-1}) = \sum_{\tau'=1}^{\tau-1} \breve{\mu}^{\Delta,\tau'}(\tilde{X}^{\Delta,1}, ..., \tilde{X}^{\Delta,\tau'-1}). \tag{C.3}$$

In terms of continuous-time representations of control processes and cumulative-control processes as well as cumulative-deviations processes, control strategies in the discrete-time models then correspond to suitably measurable functions from C^N into $L_1([0,1],\hat{K})$, and cumulative-control strategies correspond to suitably measurable functions from C^N into $C^N_{\hat{K}}$.

The following result, which is fundamental for the entire analysis, exploits the compactness of the space $C^N_{\hat{K}}$.

Proposition C.1 *For $m = 1, 2, \ldots,$ and $\Delta = \frac{1}{m}$, let $\{\breve{M}^{\Delta,\tau}\}_{\tau=1}^{m}$ be a cumulative-control strategy for the discrete-time model with period length Δ, and let $\{(M^{\Delta}(\cdot), X^{\Delta}(\cdot)\}$ be the induced process of cumulative controls and cumulative deviations from the means. Any subsequence of this sequence has a further subsequence which converges in distribution to a process $(M(\cdot), X(\cdot))$ such that $X(\cdot)$ is the Gaussian process with initial value $X(0) = 0$, zero drift, and covariance matrix $\mathbf{\Sigma}$, and, moreover, for any $t \in [0,1)$, the continuation $\{X(t')\}_{t'>t}$ of the process $X(\cdot)$ and the history $\{(M(\tau), X(\tau)\}_{\tau \leq t}$ are conditionally independent given $X(t)$.*

Proof: For any Δ, let Φ^{Δ} be the distribution of the joint process $(M^{\Delta}(\cdot), X^{\Delta}(\cdot))$. As discussed in Appendix A, the cumulative control process $M^{\Delta}(\cdot)$ takes values in the set $C^N_{\hat{K}}$; by Lemma C.1, this is a compact subset of C^N. The process $X^{\Delta}(\cdot)$ of cumulative deviations from the means take values in C^N. Thus Φ^{Δ} is an element of the space $\mathcal{M}(C^N_{\hat{K}} \times C^N)$ of probability measures on $C^N_{\hat{K}} \times C^N$. By Prohorov's Theorem (Billingsley, 1968, p.240), the sequence $\{\Phi^{\Delta}\}$ is sequentially compact if and only if it is tight, i.e., if and only if for any $\eta > 0$, there exists a compact subset $K_{\eta} \subset C^N_{\hat{K}} \times C^N$ such that, for any Δ, $\Phi^{\Delta}(K_{\eta}) \geq 1 - \eta$. Since $C^N_{\hat{K}}$ itself is a compact subset of C^N, it suffices to show that for any $\eta > 0$, there exists a compact subset $\bar{K}_{\eta} \subset C^N$ such that $\Phi^{\Delta}(C^N_{\hat{K}} \times \bar{K}_{\eta}) \geq 1 - \eta$ for all Δ. For this purpose, consider the marginal distributions Ψ^{Δ} of the cumulative-deviations processes $(X^{\Delta}_1(\cdot), ..., X^{\Delta}_N(\cdot))$. By Proposition B.1 in Appendix B, the sequence $\{\Psi^{\Delta}\}$ converges to the distribution of the process $X(\cdot)$. By Prohorov's Theorem, it follows that the sequence $\{\Psi^{\Delta}\}$ is tight, so for any $\eta > 0$, there exists a compact subset $\bar{K}_{\eta} \subset C^N$ such that, for any Δ, $\Psi^{\Delta}(\bar{K}_{\eta}) \geq 1 - \eta$. Since $\Psi^{\Delta}(\bar{K}_{\eta}) = \Phi^{\Delta}(C^N_{\hat{K}} \times \bar{K}_{\eta})$, this implies that the sequence $\{\Phi^{\Delta}\}$ is tight and, hence, sequentially compact.

Let Φ be a limit point of the sequence $\{\Phi^\Delta\}$, and let $(M(\cdot), X(\cdot))$ have the distribution Φ. If Ψ is the marginal distribution on C^N that is induced by Φ, then by standard arguments, Ψ coincides with the limit of the marginal distributions Ψ^Δ of the processes $X^\Delta(\cdot)$. By Proposition B.1 in Appendix B, it follows that $X(\cdot)$ is Gaussian with initial value $X(0) = 0$, zero drift, and covariance matrix $\mathbf{\Sigma}$.

It remains to be shown that for any $t \in [0, 1)$, the continuation $\{X(t')\}_{t'>t}$ of the process $X(\cdot)$ and the history $\{(M(\tau), X(\tau)\}_{\tau \leq t}$ are conditionally independent given $X(t)$. For this purpose we note that, by (C.3), formulae (B.7) - (B.12) in the proof of Proposition B.1 in Appendix B remain valid if the symbol $\mathcal{F}_t^\Delta$ in these formulae refers to the σ-algebra that is generated by the random variables $\tilde{M}^{\Delta,1}, \tilde{\pi}^{\Delta,1}, ..., \tilde{M}^{\Delta,[t/\Delta]}, \tilde{\pi}^{\Delta,[t/\Delta]}$ rather than just the σ-algebra generated by $\tilde{\pi}^{\Delta,1}, ..., \tilde{\pi}^{\Delta,[t/\Delta]}$. Appealing to Theorems 6 and 7 of Gihman and Skorokhod (1979, p. 195) and using the equation $X^\Delta(.) = (Q')^{-1}X_Q^\Delta$, as before, we may therefore conclude that the (regular) conditional distributions of the continuations $\{X^\Delta(t')\}_{t' \in (t,1]}$ given $\tilde{M}^{\Delta,1}, \tilde{\pi}^{\Delta,1}, ..., \tilde{M}^{\Delta,[t/\Delta]}, \tilde{\pi}^{\Delta,[t/\Delta]}$ converge in distribution[C-1] to the conditional distribution of the continuation $\{X(t')\}_{t' \in (t,1]}$ of the process $X(\cdot)$ given the "initial" value $X(t)$. By an argument given in Hellwig (1996, pp. 452 ff.), this convergence property of conditional distributions is sufficient for the desired conditional-independence property in the limit. *Q.E.D.*

Let $\{(M^{\Delta'}(\cdot), X^{\Delta'}(\cdot))\}$ be the subsequence of joint cumulative-control and cumulative-deviations processes which corresponds to the specified subsequence $\{s^{\Delta'}(\cdot)\}$ of incentive schemes. Any subsequence of $\{(M^{\Delta'}(\cdot), X^{\Delta'}(\cdot))\}$ has yet a further subsequence which converges in distribution to a pair $(M(\cdot), X(\cdot))$ such that $X(\cdot)$ is the driftless Brownian motion specified in Theorem 1, and moreover, for any t, the history of the process $(M(\cdot), X(\cdot))$ up to t and the continuation of the process $X(\cdot)$ from t on are conditionally independent given $X(t)$. Convergence of the sequence $\{(M^{\Delta'}(\cdot), X^{\Delta'}(\cdot))\}$ itself is assured if all convergent subsequences have the same limit. For this it suffices to show that the control process $\mu(\cdot)$ which corresponds to the cumulative control process $M(\cdot)$ is actually an optimal control process for the agent in the continuous-time model with incentive scheme $s(\cdot)$. This is so because, with a strictly convex cost function, by the results of

[C-1] Note that the conditional distributions themselves can be regarded as measure-valued random variables, i.e., functions from the appropriate spaces of histories into spaces of measures on continuations.

Schättler and Sung (1993, Theorems 4.1 and 4.2), *regardless of the incentive scheme,* the solution to the agent's problem in the continuous-time model is unique in the sense that any two optimal control processes must have the same values almost surely for almost all $t \in [0,1]$.

In an abuse of notation, we also write $\{(M^{\Delta'}(\cdot), X^{\Delta'}(\cdot))\}$ for any convergent subsequence, and we show that if $\{(M^{\Delta'}(\cdot), X^{\Delta'}(\cdot))\}$ converges in distribution to a pair $(M(\cdot), X(\cdot))$, then the limiting cumulative control process $M(\cdot)$ must almost surely coincide with the process $M^*(\cdot)$ which is optimal for the agent in the continuous-time model with incentive scheme $s(\cdot)$. As discussed in the text, the argument proceeds in three distinct steps. The **first step** involves showing that the agent's expected payoff from the limit pair $(M(.), X(.))$ in the continuous-time model with incentive scheme $s(\cdot)$ is not significantly worse than his expected payoffs from the pairs $\{(M^{\Delta'}(.), X^{\Delta'}(.))\}$ in the discrete-time models with incentive schemes $s^{\Delta'}$ when Δ' is small.

Proposition C.2 *Under the assumptions of Proposition 4, if* $\{(M^{\Delta'}(\cdot), X^{\Delta'}(\cdot))\}$ *converges in distribution to* $(M(\cdot), X(\cdot))$*, then*

$$\begin{aligned}&\lim_{\bar\Delta\to 0}\sup_{\Delta'\leq\bar\Delta}[-E\exp\{-r[s^{\Delta'}(z^{\Delta'}) - \Gamma(M^{\Delta'}(.))]\}] \\ &\leq\ -E\exp\{-r[s\left(\sum_{i=1}^{N}(M_i(1)+X_i(1))\right) - \Gamma(M(.))]\},\end{aligned} \tag{C.4}$$

where $z^{\Delta'} = \sum_{i=1}^{N} Z_i^{\Delta'}(1) = \sum_{i=1}^{N}[M_i^{\Delta'}(1) + X_i^{\Delta'}(1)]$.

Proof: By Skorokhod's theorem (see, e.g., Hildenbrand (1974), pp. 50 f.), convergence in distribution of $\{(M^{\Delta'}(.), X^{\Delta'}(.))\}$ to $(M(.), X(.))$ implies the existence of a measure space $(A, \mathcal{A}, \alpha)$ and measurable functions $(f_M^{\Delta'}, f_X^{\Delta'})$ and (f_M, f_X) of A into $C_{\hat K}^N \times C^N$ such that, for any Δ', the distribution of $(M^{\Delta'}(.), X^{\Delta'}(.))$ is $\alpha\circ(f_M^{\Delta'}, f_X^{\Delta'})^{-1}$, the distribution of $(M(.), X(.))$ is $\alpha\circ(f_M, f_X)^{-1}$, and moreover, for α-almost every $a \in A$, $(f_M^{\Delta'}(., a), f_X^{\Delta'}(., a))$ converges to $(f_M(., a), f_X(., a)$ as Δ' goes to zero. By the change-of-variables formula (Hildenbrand (1974), p. 50), (C.4) is equivalent to the inequality

$$\begin{aligned}&\lim_{\bar\Delta\to 0}\sup_{\Delta'\leq\bar\Delta}[-\int \exp\{-r[s^{\Delta'}\left(\sum_{i=1}^{N}[f_{M_i}^{\Delta'}(1,a) + f_{X_i}^{\Delta'}(1,a)]\right) - \Gamma(f_M^{\Delta'}(.,a))]\}\, d\alpha(a)] \\ &\leq\ -\int \exp\{-r[s\left(\sum_{i=1}^{N}[f_{M_i}(1,a) + f_{X_i}(1,a)]\right) - \Gamma(f_M(.,a))]\}\, d\alpha(a).\end{aligned} \tag{C.5}$$

By Fatou's lemma, the left-hand side of (C.5) is no greater than

$$-\int \exp\{-r \lim_{\bar{\Delta}\to\infty} \sup_{\Delta'\leq\bar{\Delta}} [s^{\Delta'}\left(\sum_{i=1}^{N}[f_{M_i}^{\Delta'}(1,a)+f_{X_i}^{\Delta'}(1,a)]\right) - \Gamma(f_M^{\Delta'}(.,a))]\}\, d\alpha(a)]. \quad \text{(C.6)}$$

By construction of $s(.)$, for any $a \in A$, for which $(f_M^{\Delta'}(.,a), f_X^{\Delta'}(.,a))$ converges to $(f_M(.,a),$ $f_X(.,a))$,

$$\begin{aligned} \lim_{\bar{\Delta}\to 0} \sup_{\Delta'\leq\bar{\Delta}} s^{\Delta'}\left(\sum_{i=1}^{N}[f_{M_i}^{\Delta'}(1,a)+f_{X_i}^{\Delta'}(1,a)]\right) &\leq s\left(\lim_{\Delta'\to 0}\sum_{i=1}^{N}[f_{M_i}^{\Delta'}(1,a)+f_{X_i}^{\Delta'}(1,a)]\right) \\ &= s\left(\sum_{i=1}^{N}[f_{M_i}(1,a)+f_{X_i}(1,a)]\right). \quad \text{(C.7)} \end{aligned}$$

By Lemma C.1, for any such $a \in A$,

$$\lim_{\bar{\Delta}\to 0} \sup_{\Delta'\leq\bar{\Delta}} [-\Gamma(f_M^{\Delta'}(.,a))] \leq -\Gamma(\lim_{\Delta'\to 0} f_M^{\Delta'}(.,a)) = -\Gamma(f_M(.,a)). \quad \text{(C.8)}$$

Therefore, (C.6) is no greater than the right-hand side of (C.5). This proves (C.5) and hence (C.4). *Q.E.D.*

In the **second step of the proof of Proposition 4**, we show that the agent's payoff from the limit pair $(M(.), X(.))$ in the continuous-time model with incentive scheme $s(.)$ is no better than his *maximal* payoff in the continuous-time model with this incentive scheme. For this purpose we will show that the pair $(M(.), X(.))$ can be interpreted as the result of the agent's choosing a mixed strategy, i.e., as the result of a prior randomization over admissible pure strategies in the continuous-time model.

Proposition C.3 *Let $W \in \mathcal{M}(C^N)$ be the distribution of the Brownian motion $X(.)$ with zero drift, variance-covariance matrix $\mathbf{\Sigma}$, and initial condition $X(0) = 1$, and let $\mathcal{F}(C^N, C^N_{\hat{K}})$ be the space of functions from C^N to $C^N_{\hat{K}}$ that are adapted to the filtration on C^N that results from augmenting the filtration generated by the N-dimensional Brownian motion by the addition of null sets. There exists a measure $P \in \mathcal{M}(\mathcal{F}(C^N, C^N_{\hat{K}}))$ such that*

$$\Phi = (P \times W) \circ \gamma^{-1}, \quad \text{(C.9)}$$

where $\Phi \in \mathcal{M}(C^N_{\hat{K}} \times C^N)$ is the distribution of the limit $(M(.), X(.))$ in Proposition C.1 and $\gamma : \mathcal{F}(C^N, C^N_{\hat{K}}) \times C^N \to C^N_{\hat{K}} \times C^N$ is defined by the formula

$$\gamma(g, B) = (g(B), B). \quad \text{(C.10)}$$

Proposition C.3 is essentially a version of Kuhn's Theorem for the continuous-time agency problem. A measure $\Phi \in \mathcal{M}(C_{\hat{K}}^N \times C^N)$ that exhibits the conditional-independence property established in Proposition C.1 can be interpreted as a behaviour strategy in the sense that it allows for distinct randomizations at different information sets. Specifically, if we think of a pair $(F, Y) \in C_{\hat{K}}^N \times C^N$ as a possible *outcome* resulting from the agent's behaviour, then Φ is a mixture over such outcomes; moreover, for any t, the conditional distribution over continuations $\{(M(\tau), X(\tau)\}_{\tau > t}$ given the history up to and including t corresponds to a mixture over the possible outcomes of these continuations, and the "sequence" of such conditional distributions that is inherent in the specification of the measure Φ provides an analogue to the distinct randomizations at different informations sets in the conventional specification of a behaviour strategy in a discrete-time model. The proposition asserts that such a "behaviour strategy" can actually be interpreted as the result of a prior randomization over pure strategies, where the latter are defined as suitably measurable functions from the space C^N of time paths of cumulative deviations to the space $C_{\hat{K}}^N$ of time paths of cumulative controls.

Proof of Proposition C.3: Given the measure $\Phi \in \mathcal{M}(C_{\hat{K}}^N \times C^N)$, let $\Phi(.|.)$ be a regular conditional distribution on $C_{\hat{K}}^N$ given $\{C_{\hat{K}}^N\} \times \mathcal{B}(C^N)$. Formally, we treat $\Phi(. \mid .)$ as a measurable function from C^N into $\mathcal{M}(C_{\hat{K}}^N)$. By Lusin's Theorem (see, e.g., Halmos, P. (1950), *Measure Theory*, Princeton: Van Nostrand, p. 243), for $n = 1, 2, ...$, there exists a compact set $C_n^N \subset C^N$ such that $W(C_n^N) \geq 1 - \frac{1}{n}$ and, moreover, the restriction of the function $\Phi(. \mid .)$ to the set C_n^N is continuous.

For any n, let $\mathcal{C}(C_n^N, C_{\hat{K}}^N)$ be the space of continuous functions from C_n^N into $C_{\hat{K}}^N$, adapted so that for any $g \in \mathcal{C}(C_n^N, C_{\hat{K}}^N)$, and any $t \in [0, 1]$ and any $B \in C^N$ the value of the image $g(.)$ at t, $g(t; .)$, satisfies $g(t; B) = g(t; \hat{B})$ for all B and $\hat{B} \in C_n^N$ with $B(t') = \hat{B}(t')$ for all $t' \leq t$. The space $\mathcal{C}(C_n^N, C_{\hat{K}}^N)$ will be endowed with the topology of uniform convergence.

Given a countable dense set of points $B_n^1, B_n^2, ...$ in C_n^N, consider the sequence of subsets $\mathcal{P}_n^1, \mathcal{P}_n^2, ...$ of $\mathcal{M}(\mathcal{C}(C_n^N, C_{\hat{K}}^N))$ such that

$$\mathcal{P}_n^1 = \{P \in \mathcal{M}(\mathcal{C}(C_n^N, C_{\hat{K}}^N)) | \text{ for any Borel set } Y \in \mathcal{B}(C_{\hat{K}}^N),$$
$$\Phi(Y \mid B_n^1) = P(\{g \in \mathcal{C}(C_n^N, C_{\hat{K}}^N) | \; g(B_n^1) \in Y\})\}, \quad \text{(C.11)}$$

and, for $i = 2, 3, ...,$

$$\mathcal{P}_n^i = \{P \in \mathcal{P}_n^{i-1} | \text{ for any Borel set } Y \in \mathcal{B}(C_{\hat{K}}^N)$$
$$\Phi(Y \mid B_n^i) = P(\{g \in \mathcal{C}(C_n^N, C_{\hat{K}}^N) | \ g(B_n^i) \in Y\})\}. \tag{C.12}$$

Clearly, for any i, the set $\mathcal{P}_n^i$ is a nonempty and closed subset of $\mathcal{M}(\mathcal{C}(C_n^N, C_{\hat{K}}^N))$, and one has $\mathcal{P}_n^{i+1} \subset \mathcal{P}_n^i$. It follows that the sequence $\{\mathcal{P}_n^i\}$ has a limit $\mathcal{P}_n$, which is again a nonempty closed subset of $\mathcal{M}(\mathcal{C}(C_n^N, C_{\hat{K}}^N))$.

We claim that any measure $P \in \mathcal{P}_n$ satisfies

$$\Phi(Y \mid B) = P(\{g \in \mathcal{C}(C_n^N, C_{\hat{K}}^N) | \ g(B) \in Y\})\} \tag{C.13}$$

for all sets $Y \in \mathcal{B}(C_{\hat{K}}^N)$ and all $B \in C_n^N$. To verify this claim, fix $B_n \in C_n^N$, and let $\{B_n^{i^k}\}$ be any subsequence of the sequence $\{B_n^i\}$ such that $B_n = \lim_{k\to\infty} B_n^{i^k}$. Given that the restriction of $\Phi(. \mid .)$ to C_n^N is continuous, we have

$$\lim_{k\to\infty} \int_{C_{\hat{K}}^N} f(M) d\Phi(M| \ B_n^{i^k}) = \int_{C_{\hat{K}}^N} f(M) d\Phi(M| \ B_n) \tag{C.14}$$

for any bounded continuous function f of $C_{\hat{K}}^N$ into $I\!R$. Given that $P \in \mathcal{P}_n = \cap_{i=1}^{\infty} \mathcal{P}_n^i$, the measure P satisfies (C.13) for $B = B_n^{i^1}, B_n^{i^2}, ...$ This implies that

$$\int_{C_{\hat{K}}^N} f(M) d\Phi(M| \ B_n^{i^k}) = \int_{\mathcal{C}(C_n^N, C_{\hat{K}}^N)} f(g(B_n^{i^k})) dP(g) \tag{C.15}$$

for all k. Given that the elements of $\mathcal{C}(C_n^N, C_{\hat{K}}^N)$ are continuous, convergence of $B_n^{i^k}$ to B_n implies convergence of $f(g(B_n^{i^k}))$ to $f(g(B_n))$ for any $g \in \mathcal{C}(C_n^N, C_{\hat{K}}^N)$ and any bounded continuous $f : C_{\hat{K}}^N \to I\!R$. By Lebesgue's bounded-convergence theorem, it follows that

$$\lim_{k\to\infty} \int_{\mathcal{C}(C_n^N, C_{\hat{K}}^N)} f(g(B_n^{i^k})) dP(g) = \int_{\mathcal{C}(C_n^N, C_{\hat{K}}^N)} f(g(B_n)) dP(g) \tag{C.16}$$

for any bounded continuous function f on $C_{\hat{K}}^N$. Upon combining (C.14), (C.15), and (C.16), one concludes that

$$\int_{C_{\hat{K}}^N} f(M) d\Phi(M| \ B_n) = \int_{\mathcal{C}(C_n^N, C_{\hat{K}}^N)} f(g(B_n)) dP(g)$$

for any bounded continuous function f on $C_{\hat{K}}^N$. The validity of (C.13) for $B = B_n$ follows immediately.

Recalling that $\mathcal{C}(C_n^N, C_{\hat{K}}^N)$ is endowed with the topology of uniform convergence, we can use the Borel σ-algebra $\mathcal{B}(\mathcal{C}(C_n^N, C_{\hat{K}}^N)$ to specify on the underlying function space $\mathcal{F}(C^N, C_{\hat{K}}^N)$ itself a σ-algebra $\mathcal{A}_n$ of cylinder sets taking the form $A = \{g \in \mathcal{F}(C^N, C_{\hat{K}}^N)|\ (g|C_n^N)\ \in Y\}$ for some $Y \in \mathcal{B}(\mathcal{C}(C_n^N, C_{\hat{K}}^N)$; here $(g|C_n^N)$ denotes the restriction of g to C_n^N. The formula

$$Q_P^n(\{g \in \mathcal{F}(C^N, C_{\hat{K}}^N)|\ (g|C_n^N)\ \in Y\}) = P(Y) \tag{C.17}$$

etsablishes a natural bijection between measures P on $(\mathcal{C}(C_n^N, C_{\hat{K}}^N), \mathcal{B}(\mathcal{C}(C_n^N, C_{\hat{K}}^N))$ and measures Q_P^n on $(\mathcal{F}(C^N, C_{\hat{K}}^N), \mathcal{A}_n)$.

Now consider the class $\mathcal{A} = \cup_{n=1}^{\infty}\mathcal{A}_n$ of all cylinder subsets of $\mathcal{F}(C^N, C_{\hat{K}}^N)$, and let $\mathcal{Q}$ be the set of countably additive set functions on $(\mathcal{F}(C^N, C_{\hat{K}}^N), \mathcal{A})$. For any n, we define $\mathcal{Q}^n$ as the set of measures $Q \in \mathcal{Q}$ such that the restriction of Q to $\mathcal{A}_n$ satisfies (C.17) for some $P \in \mathcal{P}_n$.

We may assume that the sequence $\{C_n^N\}$ of compact subsets of C^N in the application of Lusin's Theorem is nondecreasing. This implies that the sequence $\{\mathcal{A}_n\}$ of σ-algebras on $\mathcal{F}(C^N, C_{\hat{K}}^N)$ is nondecreasing, and in turn that the sequence $\{\mathcal{Q}^n\}$ of sets of measures on $(\mathcal{F}(C^N, C_{\hat{K}}^N), \mathcal{A})$ is nonincreasing. It follows that the limit $\mathcal{Q}^* = \cap_{n=1}^{\infty}\mathcal{Q}^n$ is a well defined nonempty subset of $\mathcal{Q}$.

Now let P be any element of $\mathcal{Q}^*$. By Kolmogorov's Extension Theorem (see, e.g., Gihman, I. and A.W. Skorokhod (1972), *Theory of Stochastic Processes*, Vol. I, Heidelberg: Springer, p.46), there exists a unique extension of P to the σ-algebra generated by $\mathcal{A}$. In an abuse of notation, we refer to this extension also as P. Extending the set function P still further, we introduce the set

$$\mathcal{F}_0 = \{g \in \mathcal{F}(C^N, C_{\hat{K}}^N)|g(B) = 0 \text{ for all } B \in [C^N \setminus \cup_{n=1}^{\infty} C_n^N]\}$$

of those functions in $\mathcal{F}(C^N, C_{\hat{K}}^N)$ that assign the null function to all B outside the sets $C_1^N, C_2^N, ...$, and we define, for any $A \in \mathcal{A}$,

$$P(A \cap \mathcal{F}_0) = P(A)$$

and

$$P(A \cap [\mathcal{F}(C^N, C_{\hat{K}}^N)\backslash\mathcal{F}_0]) = 0.$$

With this specification of a measure P on $\mathcal{F}(C^N, C^N_{\hat{K}})$, it is now easy to see that the distribution of the pair $(g(B), B) \in C^N_{\hat{K}} \times C^N$ that is induced by the product measure $P \times W$ is exactly the same as the original measure Φ on $C^N_{\hat{K}} \times C^N$. *Q.E.D.*

Corollary C.1 *Let $g^* \in \mathcal{F}(C^N, C^N_{\hat{K}})$ be an optimal cumulative-control strategy of the agent in the continuous-time model with incentive scheme $s(.)$, and let $(M^*(\cdot), X^*(\cdot))$ be the associated cumulative-control and cumulative-deviations process. Then the limit $(M(.), X(.))$ in Proposition C.1 satisfies*

$$\begin{aligned} &-E\exp\{-r[s\left(\sum_{i=1}^{N}(M_i(1)+X_i(1))\right)-\Gamma(M(.))]\} \\ &\quad\le -E\exp\{-r[s\left(\sum_{i=1}^{N}(M_i^*(1)+X_i(1))\right)-\Gamma(M^*(.))]\}. \end{aligned} \tag{C.18}$$

<u>Proof:</u> By the definition of $(M^*(\cdot), X^*(\cdot))$ and g^*, we trivially have

$$\begin{aligned} &-E\exp\{-r[s\left(\sum_{i=1}^{N}(M_i^*(1)+X_i(1))\right)-\Gamma(M^*(.))]\} \\ &= -\int \exp\{-r[s\left(\sum_{i=1}^{N}(g_i^*(1,B)+B_i(1))\right)-\Gamma(g^*(.,B))]\}\, dW(B), \end{aligned} \tag{C.19}$$

where $B \in C^N$ is any realization and $W \in \mathcal{M}(C^N)$ is the distribution of the Brownian motion $X(\cdot)$.

The optimality of the strategy g^* for the agent in the continuous-time model with incentive scheme $s(\cdot)$ implies that

$$\begin{aligned} &-\int \exp\{-r[s\left(\sum_{i=1}^{N}(g_i^*(1,B)+B_i(1))\right)-\Gamma(g^*(.,B))]\}\, dW(B) \\ &\quad\ge -\int \exp\{-r[s\left(\sum_{i=1}^{N}(g_i(1,B)+B_i(1))\right)-\Gamma(g(.,B))]\}\, dW(B) \end{aligned} \tag{C.20}$$

for $g \in \mathcal{F}(C^N, C^N_{\hat{K}})$, i.e., for any admissible pure strategy of the agent in the continuous-time model. From (C.20), we immediately obtain

$$-\int \exp\{-r[s\left(\sum_{i=1}^{N}(g_i^*(1,B)+B_i(1))\right)-\Gamma(g^*(.,B))]\}\, dW(B) \tag{C.21}$$

$$\ge\ -\int\int \exp\{-r[s\left(\sum_{i=1}^{N}(g_i(1,B)+B_i(1))\right)-\Gamma(g(.,B))]\}\, d(P\times W)(g,B),$$

where $P \in \mathcal{M}(\mathcal{F}(C^N, C^N_{\hat{K}}))$ is given by Proposition C.3. Since Proposition C.3 in turn implies

$$\begin{aligned} &-E \exp\{-r[s\left(\sum_{i=1}^{N}(M_i(1) + X_i(1))\right) - \Gamma(M(.))]\} \\ &= -\int\int \exp\{-r[s\left(\sum_{i=1}^{N}(g_i(1, B) + B_i(1))\right) - \Gamma(g(., B))]\} \, d(P \times W)(g, B), \end{aligned}$$

the validity of (C.18) follows immediately. *Q.E.D.*

In the **third step of the proof of Proposition 4**, we argue that the agent's maximal payoff in the continuous-time model with the incentive scheme $s(.)$ is not significantly better than his maximal payoffs in the discrete-time models with incentive schemes $s^{\Delta'}(.)$ when Δ' is small. This is the point of

Proposition C.4 *Under the assumptions of Proposition 4, if* $\{(M^{\Delta'}(\cdot), X^{\Delta'}(\cdot))\}$ *converges in distribution to* $(M(\cdot), X(\cdot))$, *then*

$$\begin{aligned} &\lim_{\bar{\Delta}\to 0} \inf_{\Delta' \leq \bar{\Delta}} [-E \exp\{-r[s^{\Delta'}(z^{\Delta'}) - \Gamma(M^{\Delta'}(\cdot))]\}] \\ &\geq -E \exp\{-r[s\left(\sum_{i=1}^{N}(M_i^*(1) + X_i(1))\right) - \Gamma(M^*(\cdot))]\} \end{aligned} \tag{C.22}$$

where $(M^*(\cdot), X(\cdot))$ *is the cumulative-control and cumulative-deviations process that is generated by an optimal cumulative-control strategy* $g^* \in \mathcal{F}(C^N, C^N_{\hat{K}})$ *of the agent in the continuous-time model with incentive scheme* $s(.)$.

To establish this result, we shall exhibit a sequence of cumulative-control strategies for the discrete-time models whose payoffs for small period lengths are not significantly worse than the payoff from choosing the optimal cumulative-control process g^* at the incentive scheme $s(\cdot)$ in the continuous-time model. Trivially, this will imply that for small period lengths the payoffs of the *optimal* control processses in the different discrete-time models are also not significantly worse than the agent's payoff from choosing the optimal control process $\mu^*(.)$ at the incentive scheme $s(.)$ in the continuous-time model.

The discrete-time cumulative-control strategies that we shall exhibit can be seen as approximations to the optimal cumulative-control process g^* at the incentive scheme $s(\cdot)$ in the continuous-time model. Complications arise because, apart from the need to adapt control processes to the time structures of the discrete-time models, we need to take account of possible discontinuities of the cost function $\Gamma(\cdot)$ and the incentive scheme $s(.)$ as well as the strategy g^*. The discontinuities in $\Gamma(\cdot)$ and g^* are taken care of by the following preliminary lemmas. In (C.23) below, as in the text, $\bar{E}$ is the maximum of $\sum_{i=1}^N \mu_i$ on the set $\hat{K}$, and $\mu(\bar{E})$ is the vector in $\hat{K}$ which minimizes $\hat{c}(\mu)$ under the constraint $\sum_{i=1}^N \mu_i = \bar{E}$.

Lemma C.2 *For any function $g \in \mathcal{F}(C^N, C^N_{\hat{K}})$ and any $h > 0$, define a new function function $\hat{g}(h,g)$ so that for any $B \in C^N$, $\hat{g}(\cdot, B; h, g) \in C^N_{\hat{K}}$ is given as*

$$\hat{g}(t, B; h, g) = \mu(\bar{E})\, t \quad \text{if} \quad t \leq h \tag{C.23}$$

and

$$\begin{aligned} \hat{g}(t, B; h, g) = {} & (1 - h(\frac{t}{h} - \left[\frac{t}{h}\right]) g(h(\left[\frac{t}{h}\right] - 1), B) \\ & + h(\frac{t}{h} - \left[\frac{t}{h}\right]) g(h \left[\frac{t}{h}\right], B), \quad \text{if} \quad t > h, \end{aligned} \tag{C.24}$$

where $\left[\frac{t}{h}\right]$ is the largest integer not exceeding $\frac{t}{h}$. Then $\hat{g}(h,g) \in \mathcal{F}(C^N, C^N_{\hat{K}})$, and, for any $B \in C^N$,

$$\lim_{h \to 0} \hat{g}(\cdot, B; h, g) = g(\cdot, B), \text{ uniformly in } B, \tag{C.25}$$

and

$$\lim_{h \to 0} \Gamma(\hat{g}(\cdot, B; h, g)) = \Gamma(g(B)). \tag{C.26}$$

Proof: Since $g(\cdot, B) \in C^N_{\hat{K}}$ for all B, (C.23) and (C.24) imply that for any $h \in (0, 1]$ and any $B \in C^N$, the function $\hat{g}(\cdot, B; h, g)$ has a Radon-Nikodym derivative $f_{\hat{g}(\cdot, B; hg)}$ satisfying

$$f_{\hat{g}(\cdot, B; hg)}(t) = \mu(\bar{E}), \quad \text{if} \quad t \leq h, \tag{C.27}$$

and

$$\begin{aligned} f_{\hat{g}(\cdot, B; h, g)}(t) &= \frac{g(h\left[\frac{t}{h}\right], B) - g(h(\left[\frac{t}{h}\right] - 1), B)}{h} \\ &= \frac{1}{h} \int_{h(\left[\frac{t}{h}\right] - 1)}^{h\left[\frac{t}{h}\right]} f_{g(\cdot, B)}(t')\, dt', \quad \text{if} \quad t > h. \end{aligned} \tag{C.28}$$

Since $\mu(\bar{E}) \in \hat{K}$ and, for any t', $f_{g(\cdot,B)}(t') \in \hat{K}$, and $\hat{K}$ is convex, it follows that $f_{\hat{g}(\cdot,B;hg)}(t) \in \hat{K}$ for all t, and hence that $\hat{g}(\cdot, B; h, g) \in C^N_{\hat{K}}$. (C.23) and (C.24) also imply that the map $B \to \hat{g}(\cdot, B; h, g)$ has the same measurability properties as the map $B \to g(\cdot, B)$. Therefore $\hat{g}(h, g) \in \mathcal{F}(C^N, C^N_{\hat{K}})$.

Since $f_{g(\cdot,B)}(t') \in \hat{K}$ for all B, the difference $g(h\left[\frac{t}{h}\right], B) - g(h(\left[\frac{t}{h}\right] - 1), B)$ on the right-hand side of (C.24) is bounded, uniformly in t, B, and h. For any $t \in (0, 1]$ then, (C.24) implies (C.25) because

$$\lim_{h\to 0} h(\left[\frac{t}{h}\right] - 1) = t$$

and the functions $g(\cdot, B)$ are equicontinuous. Finally, we note that (C.2), (C.27), and (C.28) yield

$$\begin{aligned}\Gamma(\hat{g}(\cdot, B; h, g)) &= \int_0^1 \hat{c}(f_{\hat{g}(\cdot,B;hg)}(t))dt \\ &= \hat{c}(\mu(\bar{E}))h + \sum_{i=1}^{[\frac{1}{h}]-1} h\hat{c}\left(\frac{\int_{(i-1)h}^{ih} f_{g(\cdot,B)}(t)\, dt}{h}\right)\end{aligned}$$

for any $h > 0$ and any $B \in C^N$. By the convexity of $\hat{c}$, it follows that

$$\begin{aligned}\Gamma(\hat{g}(\cdot, B; h, g)) &\leq \hat{c}(\mu(\bar{E}))h + \sum_{i=1}^{[\frac{1}{h}]-1} h \int_{(i-1)h}^{ih} \hat{c}(f_{g(\cdot,B)}(t))\, dt \\ &\leq \hat{c}(\mu(\bar{E}))h + \Gamma(g(\cdot, B)) \end{aligned} \tag{C.29}$$

and hence that $\lim_{\bar{h}\to 0} \sup_{h\leq\bar{h}} \Gamma(\hat{g}(\cdot, B; h, g)) \leq \Gamma(g(\cdot, B))$. By Lemma C.1 and (C.25), we also have $\lim_{\bar{h}\to 0} \inf_{h\leq\bar{h}} \Gamma(\hat{g}(\cdot, B; h, g)) \geq \Gamma(g(\cdot, B))$. (C.26) follows immediately. *Q.E.D.*

Lemma C.3 *Let $g \in \mathcal{F}(C^N, C^N_{\hat{K}})$ be such that the function*

$$B \to f_{g(\cdot,B)}(\cdot) \tag{C.30}$$

of C^N into $L_1([0, 1], \hat{K})$ that is defined by the Radon-Nikodym derivatives of the images of g is continuous. Then the functions $B \to g(\cdot, B)$ of C^N into $C^N_{\hat{K}}$ and $B \to \Gamma(g(\cdot, B))$ of C^N into $I\!R$ are continuous, and the convergence in (C.26) is uniform over any compact subset of C^N.

<u>Proof:</u> Given the continuity of the function (C.30) and the boundedness of $\hat{c}$ on $\hat{K}$, continuity of the functions $B \to g(\cdot, B)$ and $B \to \Gamma(g(\cdot, B))$ follows from Lebesgue's

bounded-convergence theorem. To complete the proof, consider any sequences $\{h^k\}$ converging to zero and $\{B^k\}$ converging to $B^* \in C^N$. Then from (C.29), we have $\Gamma(\hat{g}(\cdot, B^k; h^k, g)) \leq \hat{c}(\mu(\bar{E}))h^k + \Gamma(g(\cdot, B^k))$ for all k, and hence

$$\lim_{\bar{k}\to\infty} \sup_{k\geq\bar{k}} \Gamma(\hat{g}(\cdot, B^k; h^k, g)) \;\leq\; \lim_{\bar{k}\to\infty} \sup_{k\geq\bar{k}} \Gamma(g(\cdot, B^k)) \;.$$

Given that B^k converges to B^* and the map $B \to \Gamma(g(\cdot, B))$ is continuous, it follows that $\lim_{\bar{k}\to\infty} \sup_{k\geq\bar{k}} \Gamma(\hat{g}(\cdot, B^k; h^k, g)) = \Gamma(g(\cdot, B^*))$. By Lemma C.1 and the continuity of the map $B \to g(\cdot, B)$, we also have $\lim_{\bar{k}\to\infty} \inf_{k\geq\bar{k}} \Gamma(\hat{g}(\cdot, B^k; h^k, g)) \geq \Gamma(g(\cdot, B^*))$. Thus $\lim_{k\to\infty} \Gamma(\hat{g}(\cdot, B^k; h^k, g)) = \Gamma(g(\cdot, B^*))$, and the convergence in (C.26) must be uniform over compacta. *Q.E.D.*

Lemma C.4 *For any $g \in \mathcal{F}(C^N, C^N_{\hat{K}})$ and $k = 1, 2, ...,$ there exist functions $g^k(g) \in \mathcal{F}(C^N, C^N_{\hat{K}})$, such that (i) for W - almost every $B \in C^N$, $g^k(\cdot, B; g)$ and $\Gamma(g^k(\cdot, B; g))$ converge to $g(\cdot, B)$ and $\Gamma(g(\cdot, B))$ as k goes out of bounds, and (ii) for any k, the function $B \to f_{g^k(\cdot,B;g)}(\cdot)$ of C^N into $L_1([0, 1], \hat{K})$ that is defined by the Radon-Nikodym derivatives of the images of $g^k(g)$ is continuous.*

Proof: Using Lemma C.2, for the given $g \in \mathcal{F}(C^N, C^N_{\hat{K}})$ and $k = 1, 2, \ldots$, let h^k be such that for all $B \in C$ and all $t \in [0, 1]$,

$$\left|\hat{g}(t, B; h^k, g) - g(t, B)\right| \;\leq\; \frac{1}{k} \;.$$

Note that for each k, the Radon-Nikodym derivatives $f_{\hat{g}(\cdot,B;h^k,g)}(\cdot)$ of $\hat{g}(\cdot, B; h^k, g)$ take the form

$$f_{\hat{g}(\cdot,B;h^k,g)}(t) \;=\; d^k_{[t/h_k]}(B) \;,$$

where $d^k_j, j = 0, 1, \ldots, 1/h_k - 1$, are suitably adapted functions on C^N. By Lusin's Theorem (see, e.g., Halmos, 1950, p.243), for each k and j, there exists a compact subset C^N_{kj} of C^N, with $W(C^N_{kj}) \geq 1 - h^k/2^k$, such that the restriction of the function d^k_j to the set C^N_{kj} is continuous. By Tietze's extension theorem, for each k and j, there exists a continuous function $\overline{d}^k_j$ of C^N into $\hat{K}$ such that $\overline{d}^k_j(B) = d^k_j(B)$ for all $B \in C^N_{kj}$. Without loss of generality, we may assume that for each k and j, $\overline{d}^k_j$ shares the measurability properties of d_j^k so $\overline{d}_j^k(B)$ is independent of the behaviour of $B(t)$ for $t > jh_k$. If we define $g^k(g)$ inductively by setting

$$g^k(t, B; g) \ = \ t\overline{d}_0^k(B) \quad \text{if} \quad t \le h_k$$

and

$$g^k(t, B; g) \ = \ g^k([t/h_k]h_k, B; g) + (t - h_k[t/h_k])\overline{d}_{[t/h_k]}^k(B) \quad \text{if} \quad t \in (h_k, 1] \ ,$$

we therefore have $g^k(g) \in \mathcal{F}(C^N, C_{\hat{K}}^N)$. For any $B \in C$, the Radon-Nikodym derivative of $g^k(B)$ satisfies:

$$f_{g^k(\cdot,B;g)}(t) \ = \ \overline{d}_{[t/h_k]}^k(B) \ ,$$

so that indeed the function $B \ \rightarrow f_{g^k(\cdot,B;g)}(\cdot)$ of C^N into $L_1([0,1], \hat{K})$ is continuous. If we write $C_k^N := \bigcap_{j=0}^{[1/h_k]-1} C_{kj}^N$, then for $B \in C_k^N$, we also have

$$f_{g^k(\cdot,B;g)}(t) \ = \ f_{\hat{g}(\cdot,B;h^k,g)}(t)$$

for all t, hence $g^k(\cdot, B; g) = \hat{g}(\cdot, B; h^k, g)$.

For $r = 1, 2, \ldots$, let $C^{Nr} = \bigcap_{k \ge r} C_k^N$, and $C^{N*} = \lim_{r\to\infty} C^r$. By elementary set theory we have

$$C^{Nr} \ = \ C^N \setminus \bigcup_{k\ge r}[C^N \setminus C_k^N] \ \text{ and } \ C_k^N \ = \ C^N \setminus \bigcup_{j=0}^{[1/h_k]-1} [C^N \setminus C_{kj}^N] \ ,$$

hence,

$$W(C^{Nr}) \ \ge \ 1 - \sum_{k=r}^{\infty} W(C^N \setminus C_k^N) \ = \ 1 - \sum_{k=r}^{\infty}(1 - W(C_k^N))$$

and

$$W(C_k^N) \ \ge \ 1 - \sum_{j=0}^{[1/h_k]-1} W(C^N \setminus C_{kj}^N) \ \ge \ 1 - \sum_{j=0}^{[1/h_k]-1} (1 - W(C_{kj}^N)) \ = \ 1 - \frac{1}{2^k}$$

because $W(C_{kj}^N) \ge 1 - h^k/2^k$ for $j = 0, \ldots, [1/h_k] - 1$. It follows that for any r,

$$W(C^{Nr}) \ \ge \ 1 - \sum_{k=r}^{\infty}\left(\frac{1}{2^k}\right) \ = \ 1 - \frac{1}{2^{r-1}} \ ,$$

and hence that $W(C^{N*}) = 1$.

For any $B \in C^{N*}$ though, we have $B \in C^{Nr}$ for some r, and hence $B \in C_k^N$ for any sufficiently large k. But then $B \in C^{N*}$ implies $\omega_k(B) = \omega^{h_k}(B)$ for any sufficiently large

k and hence, by Lemma C.2, $\lim_{k\to\infty} g^k(\cdot, B; g) = g(\cdot, B)$ and $\lim_{k\to\infty} \Gamma(g^k(\cdot, B; g)) = \Gamma(g(\cdot, B))$. The subset of C^N for which $g^k(\cdot, B; g)$ and $\Gamma(g^k(\cdot, B; g))$ converge to $g(\cdot, B)$ and $\Gamma(g(\cdot, B))$ contains C^{N*} and hence has W-measure equal to one. *Q.E.D.*

We now turn to the specification of cumulative-control strategies for the discrete-time models that will approximate the payoff from the optimal strategy g^* at the incentive scheme $s(\cdot)$ in the continuous-time model. Let $g^k(g^*)$, $k = 1, 2, ...,$ be the continuous approximations to the strategy g^* that are given by Lemma C.4. For any k and any $\varepsilon > 0$, define $g^{k\varepsilon}(g^*)$ so that

$$g^{k\varepsilon}(t, B) = (1-\varepsilon)g^k(t, B) + \varepsilon t\mu(\bar{E}) \tag{C.31}$$

for any $t \in [0, 1]$ and any $B \in C^N$. For any k, $\varepsilon > 0$, and $\Delta' > 0$, use Lemma C.2 to define discrete-time approximations $g^{\Delta' k} = \hat{g}(\Delta', g^k(g^*))$ and $g^{\Delta' k\varepsilon} = \hat{g}(\Delta', g^{k\varepsilon}(g^*))$ of the cumulative-control strategies $g^k(g^*)$ and $g^{k\varepsilon}(g^*)$, and note that, trivially,

$$g^{\Delta' k\varepsilon}(t, B) = (1-\varepsilon)g^{\Delta' k}(t, B) + \varepsilon t\mu(\bar{E}) \tag{C.32}$$

for all $t \in [0, 1]$ and all $B \in C^N$.

For any k, ε, and Δ', $g^{\Delta' k\varepsilon}$ is an admissible cumulative-control strategy for the discrete-time model with period length Δ'. To see this, note that the Radon-Nikodym derivative $f_{g^{\Delta' k\varepsilon}(.,B)}$of $g^{\Delta' k\varepsilon}(., B)$ satisfies

$$f_{g^{\Delta' k\varepsilon}(.,B)}(t) = (1-\varepsilon)\frac{g^k((\left[\frac{t}{\Delta'}\right]\Delta', B; g^*) - g^k((\left[\frac{t}{\Delta'}\right]-1)\Delta', B; g^*)}{\Delta'} + \varepsilon\mu(\bar{E})$$

for almost all $t \in [0, 1]$. For $t \leq \Delta'$, we have

$$f_{g^{\Delta' k\varepsilon}(.,B)}(t) = \mu(\bar{E});$$

and for $\tau = 1, 2, ..., \frac{1}{\Delta'}$, and $t \in (\tau\Delta', (\tau+1)\Delta']$,

$$f_{g^{\Delta' k\varepsilon}(.,B)}(t) = (1-\varepsilon)\frac{\int_{(\tau-1)\Delta'}^{\tau\Delta'} f_{g^k(.,B;g^*)}(t')dt'}{\Delta'} + \varepsilon\mu(\bar{E}), \tag{C.33}$$

indicating that $f_{g^{\Delta' k\varepsilon}(.,B)}(t)$ is constant on $(\tau\Delta', (\tau+1)\Delta']$, that $f_{g^{\Delta' k\varepsilon}(.,B)}(t)$ is a convex combination of elements of $\hat{K}$ and hence itself an element of $\hat{K}$, and that $f_{g^{\Delta' k\varepsilon}(.,B)}(t)$

depends on information about B only to the extent that such information is available as of $\tau\Delta' = \left[\frac{t}{\Delta'}\right]\Delta'$.

Given the cumulative-control strategy $g^{\Delta' k\varepsilon}(.,.)$, an admissible control process $\mu^{\Delta' k\varepsilon}(\cdot)$ and associated process $X^{\Delta' k\varepsilon}(\cdot)$ of cumulative deviations from the mean are defined by a recursion argument. For $t \leq \Delta'$, $\mu^{\Delta' k\varepsilon}(t) = \mu(\bar{E})$, and, in accordance with (19) and (20) in the main text, $X_i^{\Delta' k\varepsilon}(t) = tk_i\Delta'^{\frac{1}{2}}(\tilde{A}_i^{\Delta',1} - p_i^{\Delta'}(\mu(\bar{E})))$. For $\tau = 1, 2, ..., \frac{1}{\Delta'}$, and $t \in (\tau\Delta', (\tau+1)\Delta']$,

$$\mu^{\Delta' k\varepsilon}(t) = f_{g^{\Delta' k\varepsilon}(.,.)}(t), \tag{C.34}$$

and

$$X_i^{\Delta' k\varepsilon}(t) = X_i^{\Delta' k\varepsilon}(\left[\frac{t}{\Delta'}\right]\Delta') + (t - \left[\frac{t}{\Delta'}\right]\Delta')k_i\Delta'^{\frac{1}{2}}(\tilde{A}_i^{\Delta',1} - p_i^{\Delta'}(\mu^{\Delta' k\varepsilon}(t))). \tag{C.35}$$

If we write $M^{\Delta' k\varepsilon}(\cdot)$ for the associated cumulative-control process and $\Phi^{\Delta' k\varepsilon}$ for the joint distribution of the pair $(M^{\Delta' k\varepsilon}(\cdot), X^{\Delta' k\varepsilon}(\cdot))$, we obtain the agent's expected payoff from choosing the control process $\mu^{\Delta' k\varepsilon}(\cdot)$ in the discrete-time model with period length Δ' and incentive scheme $s^{\Delta'}(.)$ as

$$-\int_{C_{\hat{K}}^N \times C^N} \exp\{-r[s^{\Delta'}\left(\sum_{i=1}^N (M_i(1) + B_i(1))\right) - \Gamma(M(\cdot))]\}\, d\Phi^{\Delta' k\varepsilon}(M, B). \tag{C.36}$$

To prove Proposition C.4, we will show that if Δ' is small, then for suitably chosen k and ε, this payoff is close to the expected payoff (C.19) from the optimal strategy g^* in the continuous-time model with incentive scheme $s(.)$. Formally, we shall claim that for any given $\eta > 0$ and any sufficiently small $\Delta' > 0$, there exist k and ε such that

$$\begin{aligned} &-\int_{C_{\hat{K}}^N \times C^N} \exp\{-r[s^{\Delta'}\left(\sum_{i=1}^N (M_i(1) + B_i(1))\right) - \Gamma(M(\cdot))]\}\, d\Phi^{\Delta' k\varepsilon}(M, B) \\ \geq\ & -(1+\eta)\int_{C_{\hat{K}}^N \times C^N} \exp\{-r[s\left(\sum_{i=1}^N (M_i(1) + B_i(1))\right) - \Gamma(M^*(\cdot))]\} d\Phi^*(M, B), \end{aligned} \tag{C.37}$$

where Φ^* is the joint distribution of $M^*(\cdot)$ and $X(\cdot)$. Since $\eta > 0$ is arbitrary and, trivially, for any Δ'

$$\begin{aligned} &-E\exp\{-r[s^{\Delta'}(z^{\Delta'}) - \Gamma(M^{\Delta'}(.))]\} \\ \geq\ & -\int_{C_{\hat{K}}^N \times C^N} \exp\{-r[s^{\Delta'}\left(\sum_{i=1}^N (M_i(1) + B_i(1))\right) - \Gamma(M(.))]\}\, d\Phi^{\Delta' k\varepsilon}(M, B), \end{aligned} \tag{C.38}$$

this will be enough to establish (C.22) and prove Proposition C.4.

As a first step towards proving (C.37), we show that for any $\varepsilon > 0$ and any k, the distributions $\Phi^{\Delta' k\varepsilon}$ converge to the distribution $\Phi^{k\varepsilon}$ of the pair $(g^{k\varepsilon}(B; g^*), B)$ that is induced by the measure W on C^N. Indeed, controlling also for the behaviour of $\Gamma(M)$, we obtain the following slightly stronger result:

Lemma C.5 *For any $\varepsilon > 0$ and any k, as Δ' converges to zero, the distributions $\bar{\Phi}^{\Delta' k\varepsilon}$ of the triples $(M, \Gamma(M), B)$ that are induced by $\Phi^{\Delta' k\varepsilon}$ converge to the distribution $\bar{\Phi}^{k\varepsilon}$ of the triple $(g^{k\varepsilon}(B; g^*), \Gamma(g^{k\varepsilon}(B; g^*)), B)$ that is induced by the measure W on C^N.*

Proof: We need to show that

$$\lim_{\Delta' \to 0} \int h(M, \Gamma(M), B)\, d\Phi^{\Delta' k\varepsilon}(M, B) = \int h(g^{k\varepsilon}(B; g^*), \Gamma(g^{k\varepsilon}(B; g^*)), B)\, dW(B) \quad \text{(C.39)}$$

for all bounded, continuous functions $h : C^N_{\hat{K}} \times I\!R \times C^N$. By construction, for any Δ', k, and ε, we have

$$\int h(M, \Gamma(M), B)\, d\Phi^{\Delta' k\varepsilon}(M, B) = \int h(g^{\Delta' k\varepsilon}(B; g^*), \Gamma(g^{\Delta' k\varepsilon}(B; g^*)), B)\, d\Psi^{\Delta' k\varepsilon}(B), \quad \text{(C.40)}$$

where $\Psi^{\Delta' k\varepsilon}$ is the distribution of the process $X^{\Delta' k\varepsilon}(\cdot)$, i.e., the marginal distribution on C^N that is induced by $\Phi^{\Delta' k\varepsilon}$. By Proposition B.1 in Appendix B, as Δ' converges to zero, the distributions $\Psi^{\Delta' k\varepsilon}$ converge to W. By Skorokhod's Theorem (Hildenbrand, 1974, p. 50), it follows that there exists a measure space $(A, \mathcal{A}, \alpha)$ and random variables $\tilde{B}^{\Delta' k\varepsilon}$, $\tilde{B}^{k\varepsilon}$ on $(A, \mathcal{A}, \alpha)$ such that for any Δ', k, and ε, $\Psi^{\Delta' k\varepsilon} = \alpha \circ (\tilde{B}^{\Delta' k\varepsilon}(.))^{-1}$ and $W = \alpha \circ (\tilde{B}^{k\varepsilon}(.))^{-1}$, and moreover, as Δ' converges to zero, $\tilde{B}^{\Delta' k\varepsilon}(a)$ converges to $\tilde{B}^{k\varepsilon}(a)$ for α-almost all $a \in A$. Using the change-of-variables formula, we can write (C.40) in the form

$$\begin{aligned} &\int h(M, \Gamma(M), B)\, d\Phi^{\Delta' k\varepsilon}(M, B) \\ &= \int h(g^{\Delta' k\varepsilon}(\tilde{B}^{\Delta' k\varepsilon}(a); g^*), \Gamma(g^{\Delta' k\varepsilon}(\tilde{B}^{\Delta' k\varepsilon}(a); g^*)), \tilde{B}^{\Delta' k\varepsilon}(a))\, d\alpha(a). \end{aligned} \quad \text{(C.41)}$$

Observe that for any $\varepsilon > 0$ and k, $g^{k\varepsilon}(g^*)$ satisfies the conditions of Lemma C.3. By Lemmas C.2 and C.3, it follows that the convergence of $\tilde{B}^{\Delta' k\varepsilon}(a)$ to $\tilde{B}^{k\varepsilon}(a)$ implies the convergence of $g^{\Delta' k\varepsilon}(\tilde{B}^{\Delta' k\varepsilon}(a); g^*) = \hat{g}(\tilde{B}^{\Delta' k\varepsilon}(a); \Delta', g^{k\varepsilon}(g^*))$ to $g^{k\varepsilon}(\tilde{B}^{k\varepsilon}(a); g^*)$ and of $\Gamma(g^{\Delta' k\varepsilon}(\tilde{B}^{\Delta' k\varepsilon}(a); g^*)) = \Gamma(\hat{g}(\tilde{B}^{\Delta' k\varepsilon}(a); \Delta', g^{k\varepsilon}(g^*)))$ to $\Gamma(g^{k\varepsilon}(\tilde{B}^{k\varepsilon}(a); g^*))$. With h continuous and bounded, it follows that

$$\begin{aligned}&\lim_{\Delta'\to 0} h(g^{\Delta' k\varepsilon}(\tilde{B}^{\Delta' k\varepsilon}(a); g^*), \Gamma(g^{\Delta' k\varepsilon}(\tilde{B}^{\Delta' k\varepsilon}(a); g^*)), \tilde{B}^{\Delta' k\varepsilon}(a)) \\ &= \quad h(g^{k\varepsilon}(\tilde{B}^{k\varepsilon}(a); g^*), \Gamma(g^{k\varepsilon}(\tilde{B}^{k\varepsilon}(a); g^*)), \tilde{B}^{k\varepsilon}(a))\end{aligned}$$

for α-almost all $a \in A$ and hence, by Lebesgue's bounded-convergence theorem, that

$$\begin{aligned}&\lim_{\Delta'\to 0} \int h(M, \Gamma(M), B)\, d\Phi^{\Delta' k\varepsilon}(M, B) \\ &= \lim_{\Delta'\to 0} \int h(g^{\Delta' k\varepsilon}(\tilde{B}^{\Delta' k\varepsilon}(a); g^*), \Gamma(g^{\Delta' k\varepsilon}(\tilde{B}^{\Delta' k\varepsilon}(a); g^*)), \tilde{B}^{\Delta' k\varepsilon}(a))\, d\alpha(a) \\ &= \int h(g^{k\varepsilon}(\tilde{B}^{k\varepsilon}(a); g^*), \Gamma(g^{k\varepsilon}(\tilde{B}^{k\varepsilon}(a); g^*)), \tilde{B}^{k\varepsilon}(a))\, d\alpha(a).\end{aligned}$$

Given that $W = \alpha \circ (\tilde{B}^{k\varepsilon}(.))^{-1}$, (C.39) follows by another application of the change-of-variables formula. *Q.E.D.*

Whereas in Lemma C.5 the continuous approximations of the strategy g^* were kept fixed, the following lemma shows that in fact the conclusion of the lemma remains valid even if we let k go out of bounds as Δ' goes to zero. Notice that as k goes out of bounds, the functions $g^{k\varepsilon}(g^*)$ converge to the function $g^{*\varepsilon}$ satisfying

$$g^{*\varepsilon}(t, B) = (1 - \varepsilon) g^*(t, B) + \varepsilon t \mu(\bar{E}) \tag{C.42}$$

for any $t \in [0, 1]$ and $B \in C^N$.

Lemma C.6 *For any $\varepsilon > 0$, there exists a sequence $\{k^{\Delta'}\}$ such that, as Δ' converges to zero, the distributions $\bar{\Phi}^{\Delta' k^{\Delta'} \varepsilon}$ of the triples $(M, \Gamma(M), B)$ that are induced by $\Phi^{\Delta' k^{\Delta'} \varepsilon}$ converge to the distribution $\bar{\Phi}^{\varepsilon}$ of the triple $(g^{*\varepsilon}(B), \Gamma(g^{*\varepsilon}(B)), B)$ that is induced by the measure W on C^N.*

Proof: Let ρ denote the Prohorov metric on $\mathcal{M}(C^N_{\hat{K}} \times I\!R \times C^N)$. Lemma C.5 implies

$$\lim_{\Delta'\to 0} \rho(\bar{\Phi}^{\Delta' k\varepsilon}, \bar{\Phi}^{k\varepsilon}) = 0 \tag{C.43}$$

for all k and ε. Also, Lemma C.4, in combination with Theorem 5.5, p.34, of Billlingsley (1968), implies

$$\lim_{k\to\infty} \rho(\bar{\Phi}^{k\varepsilon}, \bar{\Phi}^{\varepsilon}) = 0 \tag{C.44}$$

for all ε. Fix $\varepsilon > 0$. Then for each integer n, there exists k_n such that $\rho(\bar{\Phi}^{k_n\varepsilon}, \bar{\Phi}^{\varepsilon}) \leq \frac{1}{n}$. By (C.43), for each integer n, there also exists $\Delta'_n > 0$ such that $\rho(\bar{\Phi}^{\Delta'_n k_n \varepsilon}, \bar{\Phi}^{k_n\varepsilon}) \leq \frac{1}{n}$. For any $\Delta' > 0$, define

$$n^*(\Delta') := \sup\{n | \Delta'_n > \Delta'\},$$

and

$$k^{\Delta'} = k_{n^*(\Delta')}.$$

Then, by construction,

$$\rho(\bar{\Phi}^{\Delta' k^{\Delta'}\varepsilon}, \bar{\Phi}^{\varepsilon}) \leq \rho(\bar{\Phi}^{\Delta' k_{n^*(\Delta')}\varepsilon}, \bar{\Phi}^{k_{n^*(\Delta')}\varepsilon}) + \rho(\bar{\Phi}^{k_{n^*(\Delta')}\varepsilon}, \bar{\Phi}^{\varepsilon}) \leq \frac{2}{n^*(\Delta')}, \tag{C.45}$$

so we have $\lim_{\Delta'\to 0} \rho(\bar{\Phi}^{\Delta' k^{\Delta'}\varepsilon}, \bar{\Phi}^{\varepsilon}) = 0$ if $\lim_{\Delta'\to 0} n^*(\Delta') = \infty$. Suppose to the contrary that $\lim_{\Delta'\to 0} n^*(\Delta') \neq \infty$. Since $n^*(\Delta')$ is obviously nondecreasing in Δ', there exists $\bar{n} = \lim_{\Delta'\to 0} n^*(\Delta')$. Consider $\Delta'_{\bar{n}+1}$. Since $\bar{n}+1 > n^*(\Delta')$ for all $\Delta' > 0$, we must have $\Delta'_{\bar{n}+1} = 0$, contrary to the specification of $\Delta'_{\bar{n}+1}$. The assumption that $\lim_{\Delta'\to 0} n^*(\Delta') \neq \infty$ thus leads to a contradiction and must be false. By (C.45), it follows that $\lim_{\Delta'\to 0} \rho(\bar{\Phi}^{\Delta' k^{\Delta'}\varepsilon}, \bar{\Phi}^{\varepsilon}) = 0$. *Q.E.D.*

Using Lemma C.6, in (C.37) set $k = k^{\Delta'}$, and note that, by the definition of $\bar{\Phi}^{\Delta' k^{\Delta'}\varepsilon}$,

$$\begin{aligned} &-\int_{C^N_{\hat{K}}\times C^N} \exp\{-r[s^{\Delta'}\left(\sum_{i=1}^{N}(M_i(1)+B_i(1))\right) - \Gamma(M(\cdot))]\}\, d\Phi^{\Delta' k\varepsilon}(M,B) \\ &= \; -\int_{C^N_{\hat{K}}\times I\!R\times C^N} \exp\{-r[s^{\Delta'}\left(\sum_{i=1}^{N}(M_i(1)+B_i(1))\right) - \Gamma]\}\, d\bar{\Phi}^{\Delta' k\varepsilon}(M,\Gamma,B). \end{aligned} \tag{C.46}$$

By the definition of Φ^*, we also have

$$\begin{aligned} &-(1+\eta)\int_{C^N_{\hat{K}}\times C^N} \exp\{-r[s\left(\sum_{i=1}^{N}(M_i(1)+B_i(1))\right) - \Gamma(M^*(\cdot))]\} d\Phi^*(M,B) \\ &= \; -(1+\eta)\int \exp\{-r[s\left(\sum_{i=1}^{N}(g^*_i(1,B)+B_i(1))\right) - \Gamma(g^*(.,B))]\}\, dW(B), \end{aligned} \tag{C.47}$$

so (C.37) is equivalent to the requirement that

$$\begin{aligned} &-\int_{C^N_{\hat{K}}\times I\!R\times C^N} \exp\{-r[s^{\Delta'}\left(\sum_{i=1}^{N}(M_i(1)+B_i(1))\right) - \Gamma]\}\, d\bar{\Phi}^{\Delta' k^{\Delta'}\varepsilon}(M,\Gamma,B) \\ &\geq \; -(1+\eta)\int \exp\{-r[s\left(\sum_{i=1}^{N}(g^*_i(1,B)+B_i(1))\right) - \Gamma(g^*(.,B))]\}\, dW(B). \end{aligned} \tag{C.48}$$

By another application of Skorokhod's Theorem, there exists a measure space $(A, \mathcal{A}, \alpha)$ and random variables $(\tilde{M}^{\Delta'\varepsilon}, \tilde{\Gamma}^{\Delta'\varepsilon}, \tilde{B}^{\Delta'\varepsilon})$, $(\tilde{M}^{\varepsilon}, \tilde{\Gamma}^{\varepsilon}, \tilde{B}^{\varepsilon})$, such that for any Δ' and ε, $\bar{\Phi}^{\Delta' k^{\Delta'}\varepsilon} = \alpha \circ (\tilde{M}^{\Delta'\varepsilon}, \tilde{\Gamma}^{\Delta'\varepsilon}, \tilde{B}^{\Delta'\varepsilon})^{-1}$ and $\bar{\Phi}^{\varepsilon} = \alpha \circ (\tilde{M}^{\varepsilon}, \tilde{\Gamma}^{\varepsilon}, \tilde{B}^{\varepsilon})^{-1}$, and moreover,

$$\lim_{\Delta' \to 0} (\tilde{M}^{\Delta'\varepsilon}(a), \tilde{\Gamma}^{\Delta'\varepsilon}(a), \tilde{B}^{\Delta'\varepsilon}(a)) = (\tilde{M}^{\varepsilon}(a), \tilde{\Gamma}^{\varepsilon}(a), \tilde{B}^{\varepsilon}(a)) \tag{C.49}$$

for α-almost all $a \in A$. Another application of the change-of-variables formula yields

$$\begin{aligned} & -\int_{C^N_{\hat{K}} \times I\!R \times C^N} \exp\{-r[s^{\Delta'}\left(\sum_{i=1}^{N}(M_i(1) + B_i(1))\right) - \Gamma]\} \, d\bar{\Phi}^{\Delta' k^{\Delta'}\varepsilon}(M, \Gamma, B) \quad \text{(C.50)} \\ = & -\int_A \exp\{-r[s^{\Delta'}\left(\sum_{i=1}^{N}(\tilde{M}_i^{\Delta'\varepsilon}(1;a) + \tilde{B}_i^{\Delta'\varepsilon}(1;a))\right) - \tilde{\Gamma}^{\Delta'\varepsilon}(a)]\} \, d\alpha(a) \end{aligned}$$

Lemma C.7 *For α-almost every $a \in A$,*

$$\begin{aligned} & \lim_{\bar{\Delta} \to 0} \sup_{\Delta' \le \bar{\Delta}} \exp\{-r[s^{\Delta'}\left(\sum_{i=1}^{N}(\tilde{M}_i^{\Delta'\varepsilon}(1;a) + \tilde{B}_i^{\Delta'\varepsilon}(1;a))\right) - \tilde{\Gamma}^{\Delta'\varepsilon}(a)]\} \\ \le & \exp\{-r[s\left(\sum_{i=1}^{N}(g_i^*(1, \tilde{B}^{\varepsilon}(a)) + \tilde{B}_i^{\varepsilon}(1;a))\right) - \tilde{\Gamma}^{\varepsilon}(a)]\}. \quad \text{(C.51)} \end{aligned}$$

Proof: To establish (C.51), we first note that (C.49), the definition of $\bar{\Phi}^{\varepsilon} = \alpha \circ (\tilde{M}^{\varepsilon}, \tilde{\Gamma}^{\varepsilon}, \tilde{B}^{\varepsilon})^{-1}$, and (C.42) imply

$$\begin{aligned} \lim_{\Delta' \to 0} \sum_{i=1}^{N}(\tilde{M}_i^{\Delta'\varepsilon}(1;a) + \tilde{B}_i^{\Delta'\varepsilon}(1;a)) &= \sum_{i=1}^{N}(\tilde{M}_i^{\varepsilon}(1;a)) + \tilde{B}_i^{\varepsilon}(1;a)) \quad \text{(C.52)} \\ &= \sum_{i=1}^{N} g_i^{*\varepsilon}(1, \tilde{B}^{\varepsilon}(a)) + \sum_{i=1}^{N} \tilde{B}_i^{\varepsilon}(1;a) \\ &= \varepsilon\bar{E} + (1-\varepsilon)\sum_{i=1}^{N} g_i^*(1, \tilde{B}^{\varepsilon}(a)) + \sum_{i=1}^{N} \tilde{B}_i^{\varepsilon}(1;a) \end{aligned}$$

for α-almost every $a \in A$. By the definition of the incentive scheme $s(.)$, it follows that

$$\begin{aligned} & \lim_{\Delta' \to 0} s^{\Delta'}\left(\sum_{i=1}^{N}(\tilde{M}_i^{\Delta'\varepsilon}(1;a) + \tilde{B}_i^{\Delta'\varepsilon}(1;a))\right) \\ = & \; s\left(\varepsilon\bar{E} + (1-\varepsilon)\sum_{i=1}^{N} g_i^*(1, \tilde{B}^{\varepsilon}(a)) + \sum_{i=1}^{N} \tilde{B}_i^{\varepsilon}(1;a)\right) \quad \text{(C.53)} \end{aligned}$$

whenever $\varepsilon\bar{E} + (1-\varepsilon)\sum_{i=1}^{N} g_i^*(1, \tilde{B}^{\varepsilon}(a)) + \sum_{i=1}^{N} \tilde{B}_i^{\varepsilon}(1;a)$ is a continuity point of $s(.)$. By the definition of $s(.)$ in combination with the monotonicity of $s^{\Delta'}(.)$ and $s(.)$, we also have

$$\begin{aligned} & \lim_{\bar{\Delta} \to 0} \inf_{\Delta' \le \bar{\Delta}} s^{\Delta'}\left(\sum_{i=1}^{N}(\tilde{M}_i^{\Delta'\varepsilon}(1;a) + \tilde{B}_i^{\Delta'\varepsilon}(1;a))\right) \\ \ge & \; s\left(\sum_{i=1}^{N} g_i^*(1, \tilde{B}^{\varepsilon}(a)) + \sum_{i=1}^{N} \tilde{B}_i^{\varepsilon}(1;a)\right) \quad \text{(C.54)} \end{aligned}$$

whenever $\sum_{i=1}^N g_i^*(1, \tilde{B}^\varepsilon(a)) < \bar{E}$. Upon combining (C.53) and (C.54), we see that (C.54) holds for any $a \in A$ which satisfies (C.49) and for which $\bar{E} + \sum_{i=1}^N \tilde{B}_i^\varepsilon(1; a)$ is a continuity point of $s(.)$. Since $\bar{\Phi}^\varepsilon = \alpha \circ (\tilde{M}^\varepsilon, \tilde{\Gamma}^\varepsilon, \tilde{B}^\varepsilon)^{-1}$ implies $\alpha \circ (\tilde{B}^\varepsilon)^{-1} = W$, under the measure α, $\bar{E} + \sum_{i=1}^N \tilde{B}_i^\varepsilon(1; a)$ has a normal distribution; because this distribution assigns mesaure zero to any countable set and because the monotone function $s(.)$ has at most countably many points of discontinuity, it follows that the set of $a \in A$ for which (C.54) holds is the intersection of two sets of full measure and must itself have full measure. Thus (C.54) holds for α-almost all $a \in A$. Since (C.49) also implies $\lim_{\Delta' \to 0} \tilde{\Gamma}^{\Delta' \varepsilon}(a) = \tilde{\Gamma}^\varepsilon(a)$ for α-almost all $a \in A$, the validity of (C.51) follows immediately. *Q.E.D.*

Lemma C.8 *For any $\varepsilon > 0$ and any null sequence $\{\Delta'\}$, the corresponding sequence of integrands in (C.50) is uniformly integrable.*

Proof: By the definition of $\bar{\Phi}^{\Delta' k^{\Delta'} \varepsilon}$, the integrals in (C.50) can also be written as

$$-\int_{C^N} \exp\{-r[s^{\Delta'}\left(\sum_{i=1}^N (g_i^{\Delta' k^{\Delta'} \varepsilon}(1, B) + B_i(1))\right) - \Gamma(g^{\Delta' k^{\Delta'} \varepsilon}(B)]\} \, d\Psi^{\Delta' k^{\Delta'} \varepsilon}(B),$$

where $\Psi^{\Delta' k^{\Delta'} \varepsilon}$ is the marginal distribution on C^N that is induced by $\bar{\Phi}^{\Delta' k^{\Delta'} \varepsilon}$. Given that the integrands are nonnegative-valued, it suffices to show that there exist measurable functions $h^{\Delta'}(.)$ of C^N into $I\!R_+$ such that, for all Δ',

$$\exp\{-r[s^{\Delta'}\left(\sum_{i=1}^N (g_i^{\Delta' k^{\Delta'} \varepsilon}(1, B) + B_i(1))\right) - \Gamma(g^{\Delta' k^{\Delta'} \varepsilon}(B)]\} \leq h^{\Delta'}(B) \qquad \text{(C.55)}$$

for almost all $B \in C^N$, and moreover, the sequence $\{h^{\Delta'}(.)\}$ is uniformly integrable. With $h^{\Delta'}(B) \geq 0$ for all Δ' and all B, uniform integrability of the sequence $\{h^{\Delta'}(.)\}$ is equivalent to the requirements that, as Δ' goes to zero, $h^{\Delta'}(.)$ converge in distribution to $h(.)$ and $\int_{C^N} h^{\Delta'}(B) \, d\Psi^{\Delta' k^{\Delta'} \varepsilon}(B)$ converge to $\int_{C^N} h(B) \, dW(B)$ (Hildenbrand (1974), p. 52).

Construction of the majorizing functions $h^{\Delta'}(.)$ involves three arguments. First, for any $y \in I\!R$, standard analysis yields

$$\begin{aligned} \exp(y) &= \sum_{u=0}^\infty \frac{y^u}{u!} \leq \sum_{v=0}^\infty \frac{y^{2v}}{(2v)!} + y \sum_{v=0}^\infty \frac{y^{2v}}{(2v)!} \leq \sum_{v=0}^\infty \frac{y^{2v}}{(2v)!} + (1 + y^2) \sum_{v=0}^\infty \frac{y^{2v}}{(2v)!} \\ &\leq 2 + \sum_{v=1}^\infty \frac{y^{2v}}{(2v)!}(2 + 2v - 1) = 2 + 2 \sum_{v=1}^\infty \frac{y^{2v}}{(2v-1)!}. \end{aligned} \qquad \text{(C.56)}$$

Second, for any n, the cost term in (C.55) satisfies

$$\Gamma(g^{\Delta' k^{\Delta'} \varepsilon}(1, B)) \leq \bar{c}, \tag{C.57}$$

where $\bar{c} := \max_{\mu \in \hat{K}} \hat{c}(\mu)$. Third, for any Δ', the incentive payment in (C.55) satisfies

$$\begin{aligned} & s^{\Delta'} \left(\sum_{i=1}^{N} (g_i^{\Delta' k^{\Delta'} \varepsilon}(1, B) + B_i(1)) \right) \\ \geq \; & S + \sum_{\tau=1}^{1/\Delta'} \sum_{i=1}^{N} (B_i(\tau\Delta'; a) - B_i((\tau - 1)\Delta'; a)) \; \hat{c}_i(\mu^{\Delta' \tau}) \end{aligned} \tag{C.58}$$

with probability one for some constant S; for any Δ' the marginal-cost terms $\hat{c}_i(\mu^{\Delta'\tau})$ in (C.58) correspond to the control process $\mu^{\Delta'}(\cdot)$ that the incentive scheme $s^{\Delta'}(.)$ serves to implement.

Before proving (C.58), we note that if we apply (C.56) - (C.58) jointly to (C.55), we obtain

$$\begin{aligned} & \exp\{-r[s^{\Delta'} \left(\sum_{i=1}^{N} (g_i^{\Delta' k^{\Delta'} \varepsilon}(1, B) + B_i(1) \right) - \Gamma(g^{\Delta' k^{\Delta'} \varepsilon}(1, B))]\} \\ \leq \; & 2e^{-r(A - \bar{c})} \{1 + \sum_{v=1}^{\infty} \frac{\left(\sum_{i=1}^{N} \sum_{\tau=1}^{1/\Delta'} (B_i(\tau\Delta') - B_i((\tau - 1)\Delta') \; \hat{c}_i(\mu^{\Delta'\tau}) \right)^{2v}}{(2v - 1)!} \} \\ \leq \; & h^{\Delta'}(B) \end{aligned}$$

for $\Psi^{\Delta' k^{\Delta'} \varepsilon}$-almost all $B \in C^N$, where

$$\begin{aligned} h^{\Delta'}(B) \; := \; & 2e^{-r(A - \bar{c})} \{1 + \\ & \sum_{v=1}^{\infty} \frac{\gamma^{2v}}{(2v - 1)!} \left(\sum_{i=1}^{N} \sum_{\tau=1}^{1/\Delta'} (B_i(\tau\Delta') - B_i((\tau - 1)\Delta') \right)^{2v} \}, \end{aligned}$$

with $\gamma := \max_{\mu \in \hat{K}} \max_i \hat{c}_i(\mu)$. If we define $h(B) := 2e^{-r(A - \bar{c})}(1 + \gamma^2\sigma^2)$, then Proposition B.1 in Appendix B implies that $h^{\Delta'}(.)$ converges in distribution to $h(.)$ and $\int_{C^N} h^{\Delta'}(B) \; d\Psi^{\Delta' k^{\Delta'} \varepsilon}(B)$ converges to $\int_{C^N} h(B) \; dW(B)$ and, hence, that the sequence $\{h^{\Delta'}(.)\}$ is uniformly integrable.

To establish (C.58), we recall that, by assumption, for any $\Delta' > 0$, the incentive scheme $s^{\Delta'}(.)$ serves to implement the control process $\mu^{\Delta'}(\cdot)$ and, moreover, this process takes values in the interior of $\hat{K}$. By Theorem 4 of Holmström and Milgrom (1987) and

(27) in the main text, this implies that if cumulative aggregate output is equal to

$$z^{\Delta' k^{\Delta'}\varepsilon} = \sum_{i=1}^{N} Z_i^{\Delta' k^{\Delta'}\varepsilon}(1) = \sum_{i=0}^{N} k_i(\Delta')^{\frac{1}{2}} \sum_{\tau=1}^{1/\Delta'} (\tilde{A}_i^{\Delta' k^{\Delta'}\varepsilon\tau} - \hat{p}_i), \tag{C.59}$$

then the incentive payment to the agent is equal to

$$s^{\Delta'}(z^{\Delta' k^{\Delta'}\varepsilon}) = \sum_{i=1}^{N} \sum_{\tau=1}^{1/\Delta'} \tilde{A}_i^{\Delta' k^{\Delta'}\varepsilon\tau} s_i^{\Delta'\tau}, \tag{C.60}$$

where, for any τ and $i = 1, 2, ..., N$,

$$\begin{aligned} s_i^{\Delta'\tau} &= \Delta'\hat{c}(\mu^{\Delta'\tau}) - \frac{1}{r}\ln\left(1 - r\hat{c}_i(\mu^{\Delta'\tau})k_i(\Delta')^{\frac{1}{2}} + r\sum_{j=1}^{N} p_j^{\Delta'}(\mu^{\Delta'\tau})\hat{c}_j(\mu^{\Delta'\tau})k_j(\Delta')^{\frac{1}{2}}\right) \\ &\geq \Delta'\underline{c} + \hat{c}_i(\mu^{\Delta'\tau})k_i(\Delta')^{\frac{1}{2}} - \sum_{j=1}^{N} p_j^{\Delta'}(\mu^{\Delta'\tau})\hat{c}_j(\mu^{\Delta'\tau})k_j(\Delta')^{\frac{1}{2}}, \end{aligned} \tag{C.61}$$

and therefore

$$\begin{aligned} \sum_{i=1}^{N} \tilde{A}_i^{\Delta' k^{\Delta'}\varepsilon\tau} s_i^{\Delta'\tau} &\geq \Delta'\underline{c} + \sum_{i=1}^{N}(\tilde{A}_i^{\Delta' k^{\Delta'}\varepsilon\tau} - p_i^{\Delta'}(\mu^{\Delta'\tau}))\hat{c}_i(\mu^{\Delta'\tau})k_i(\Delta')^{\frac{1}{2}} \\ &= \Delta'\underline{c} + \sum_{i=1}^{N}(Z_i^{\Delta' k^{\Delta'}\varepsilon\tau} - Z_i^{\Delta' k^{\Delta'}\varepsilon,\tau-1})\hat{c}_i(\mu^{\Delta'\tau}) + \sum_{i=1}^{N}(\hat{p}_i - p_i^{\Delta'}(\mu^{\Delta'\tau}))\,\hat{c}_i(\mu^{\Delta'\tau})k_i(\Delta')^{\frac{1}{2}} \\ &= \Delta'\underline{c} + \sum_{i=1}^{N}(X_i^{\Delta' k^{\Delta'}\varepsilon\tau} - X_i^{\Delta' k^{\Delta'}\varepsilon,\tau-1})\hat{c}_i(\mu^{\Delta'\tau}) + \sum_{i=1}^{N} \Delta'(\mu_i^{\Delta' k^{\Delta'}\varepsilon\tau} - \mu_i^{\Delta'\tau})\,\hat{c}_i(\mu^{\Delta'\tau}) \\ &\geq \Delta' A + \sum_{i=1}^{N}(X_i^{\Delta' k^{\Delta'}\varepsilon\tau} - X_i^{\Delta' k^{\Delta'}\varepsilon,\tau-1})\,\hat{c}_i(\mu^{\Delta'\tau}), \end{aligned} \tag{C.62}$$

where $A = \underline{c} + \min_{\mu,\mu'\in\hat{K}} \sum_{i=1}^{N}(\mu_i' - \mu_i)\,\hat{c}_i(\mu)$. Given that under the control process $\mu^{\Delta' k^{\Delta'}\varepsilon}(\cdot)$ the random variable $z^{\Delta' k^{\Delta'}\varepsilon}$ takes the form (C.59) with probability one, (C.60) and (C.61) imply that the incentive payment $s^{\Delta'}(z^{\Delta' k^{\Delta'}\varepsilon})$ must satisfy the inequality

$$s^{\Delta'}(z^{\Delta' k^{\Delta'}\varepsilon}) \;\geq\; A + \sum_{\tau=1}^{1/\Delta'} \sum_{i=1}^{N}(X_i^{\Delta' k^{\Delta'}\varepsilon\tau} - X_i^{\Delta' k^{\Delta'}\varepsilon,\tau-1})\hat{c}_i(\mu^{\Delta'\tau}) \tag{C.63}$$

with probability one. This means that (C.58) holds for any Δ' and $\Psi^{\Delta' k^{\Delta'}\varepsilon}$-almost all $B \in C^N$. *Q.E.D.*

Proof of Proposition C.4: Suppose that the Proposition is false. Then there exists $\eta > 0$, and there exists a null sequence $\{\Delta'\}$ such that

$$\begin{aligned} &-E\exp\{-r[s^{\Delta'}(z^{\Delta'}) - \Gamma(M^{\Delta'}(\cdot))]\} \\ &\quad\leq -(1+2\eta)E\exp\{-r[s\left(\sum_{i=1}^{N}(M_i^*(1) + X_i(1))\right) - \Gamma(M^*(\cdot))]\} \end{aligned} \tag{C.64}$$

for all Δ' along the sequence. By the optimality the strategy $\mu^{\Delta'}(\cdot)$ in the discrete-time model with period length Δ' and incentive scheme $s^{\Delta'}(\cdot)$, it follows that

$$-\int_{C_{\hat{K}}^N \times C^N} \exp\{-r[s^{\Delta'}\left(\sum_{i=1}^N (M_i(1)+B_i(1))\right) - \Gamma(M(\cdot))]\} \, d\Phi^{\Delta' k^{\Delta'} \varepsilon}(M,B) \tag{C.65}$$
$$= -\int_{C_{\hat{K}}^N \times I\!R \times C^N} \exp\{-r[s^{\Delta'}\left(\sum_{i=1}^N (M_i(1)+B_i(1))\right) - \Gamma]\} \, d\bar{\Phi}^{\Delta' k^{\Delta'} \varepsilon}(M,\Gamma,B)$$
$$\leq -(1+2\eta)\int_{C^N} \exp\{-r[s\left(\sum_{i=1}^N (g_i^*(1,B)+B_i(1))\right) - \Gamma(g^*(.,B))]\} dW(B)$$

for any $\varepsilon > 0$ and associated sequence $\{k^{\Delta'}\}$ given by Lemma C.6. Then (C.50) yields

$$-\int_A \exp\{-r[s^{\Delta'}\left(\sum_{i=1}^N (\tilde{M}_i^{\Delta'\varepsilon}(1;a)+\tilde{B}_i^{\Delta'\varepsilon}(1;a))\right) - \tilde{\Gamma}^{\Delta'\varepsilon}(a)]\} \, d\alpha(a) \tag{C.66}$$
$$\leq -(1+2\eta)\int_{C^N} \exp\{-r[s\left(\sum_{i=1}^N (g_i^*(1,B)+B_i(1))\right) - \Gamma(g^*(.,B))]\} dW(B)$$

for any $\varepsilon > 0$ and any Δ' along the null sequence in question.

However, from Lemmas C.7 and C.8 in combination with the Generalized Lebesgue's Theorem (Hildenbrand (1974), p. 50), one also obtains

$$\lim_{\bar{\Delta}\to 0} \inf_{\Delta' \leq \bar{\Delta}} -\int_A \exp\{-r[s^{\Delta'}\left(\sum_{i=1}^N (\tilde{M}_i^{\Delta'\varepsilon}(1;a)+\tilde{B}_i^{\Delta'\varepsilon}(1;a))\right) - \tilde{\Gamma}^{\Delta'\varepsilon}(a)]\} \, d\alpha(a)$$
$$\geq -\int_A \exp\{-r[s\left(\sum_{i=1}^N (g_i^*(1,\tilde{B}^\varepsilon(a))+\tilde{B}_i^\varepsilon(1;a))\right) - \tilde{\Gamma}^\varepsilon(a)]\} \, d\alpha(a) \tag{C.67}$$

for any $\varepsilon > 0$. By the definition of $\bar{\Phi}^\varepsilon = \alpha \circ (\tilde{M}^\varepsilon, \tilde{\Gamma}^\varepsilon, \tilde{B}^\varepsilon)^{-1}$ and the change-of-variables formula, (C.67) can be rewritten as

$$\lim_{\bar{\Delta}\to 0} \inf_{\Delta' \leq \bar{\Delta}} -\int_A \exp\{-r[s^{\Delta'}\left(\sum_{i=1}^N (\tilde{M}_i^{\Delta'\varepsilon}(1;a)+\tilde{B}_i^{\Delta'\varepsilon}(1;a))\right) - \tilde{\Gamma}^{\Delta'\varepsilon}(a)]\} \, d\alpha(a)$$
$$\geq -\int_{C^N} \exp\{-r[s\left(\sum_{i=1}^N (g_i^{*\varepsilon}(1,B)+B_i(1))\right) - \Gamma(g^{*\varepsilon}(.,B))]\} dW(B). \tag{C.68}$$

By the definition (C.42) of $g^{*\varepsilon}$, in combination with the monotonicity of $s(.)$ and the convexity of $\Gamma(.)$, it follows that

$$\lim_{\bar{\Delta}\to 0} \inf_{\Delta' \leq \bar{\Delta}} -\int_A \exp\{-r[s^{\Delta'}\left(\sum_{i=1}^N (\tilde{M}_i^{\Delta'\varepsilon}(1;a)+\tilde{B}_i^{\Delta'\varepsilon}(1;a))\right) - \tilde{\Gamma}^{\Delta'\varepsilon}(a)]\} \, d\alpha(a)$$
$$\geq -\int_{C^N} \exp\{-r[s\left(\sum_{i=1}^N (g_i^*(1,B)+B_i(1))\right) - \varepsilon\hat{c}(\mu(\bar{E})) - (1-\varepsilon)\Gamma(g^*(.,B))]\} dW(B)$$

for any $\varepsilon > 0$. If $\varepsilon > 0$ is chosen so that

$$\exp\{r\varepsilon[\hat{c}(\mu(\bar{E})) - \min_{\mu \in \hat{K}} \hat{c}(\mu)]\} \leq 1 + \eta,$$

it follows that

$$\begin{aligned} &\lim_{\bar{\Delta} \to 0} \inf_{\Delta' \leq \bar{\Delta}} - \int_A \exp\{-r[s^{\Delta'}\left(\sum_{i=1}^N (\tilde{M}_i^{\Delta'\varepsilon}(1;a) + \tilde{B}_i^{\Delta'\varepsilon}(1;a))\right) - \tilde{\Gamma}^{\Delta'\varepsilon}(a)]\} \, d\alpha(a) \\ &\geq \; -(1+\eta) \int_{C^N} \exp\{-r[s\left(\sum_{i=1}^N (g_i^*(1,B) + B_i(1))\right) - \Gamma(g^*(.,B))]\} dW(B), \end{aligned} \tag{C.69}$$

which contradicts (C.66). The assumption that Proposition C.4 is false has thus led to a contradiction, which proves the proposition. *Q.E.D.*

Proof of Proposition 4: Given that

$$\begin{aligned} &\lim_{\bar{\Delta} \to 0} \inf_{\Delta' \leq \bar{\Delta}} [-E \exp\{-r[s^{\Delta'}(z^{\Delta'}) - \Gamma(M^{\Delta'}(.))]\}] \\ &\leq \lim_{\bar{\Delta} \to 0} \sup_{\Delta' \leq \bar{\Delta}} [-E \exp\{-r[s^{\Delta'}(z^{\Delta'}) - \Gamma(M^{\Delta'}(.))]\}], \end{aligned} \tag{C.70}$$

the three inequalities (C.4), (C.18), and (C.22) imply that

$$\begin{aligned} &\lim_{\Delta' \to 0} -E \exp\{-r[s^{\Delta'}(z^{\Delta'}) - \Gamma(M^{\Delta'}(.))]\} \\ &= \; -E \exp\{-r[s\left(\sum_{i=1}^N (M_i(1) + X_i(1))\right) - \Gamma(M(.))]\} \end{aligned} \tag{C.71}$$

and

$$\begin{aligned} &-E \exp\{-r[s\left(\sum_{i=1}^N (M_i(1) + X_i(1))\right) - \Gamma(M(.))]\} \\ &= \; -E \exp\{-r[s\left(\sum_{i=1}^N (M_i^*(1) + X_i(1))\right) - \Gamma(M^*(.))]\}, \end{aligned} \tag{C.72}$$

i.e., that the limit pair $(M(.), X(.))$ provides the agent with the maximal payoff that he can obtain in the continuous-time model with incentive scheme $s(.)$. Given that, by the results of Schättler and Sung (1993, Theorems 4.1 and 4.2), with strict convexity of the cost function $\hat{c}$, the solution to the agent's problem in the continuous-time model is unique, this implies that the measure P in Proposition C.3 is degenerate and assigns all mass to the strategy g^*. Any limit $(M(.), X(.))$ of (a subsequence) of the sequence

$\{(M^{\Delta'}(\cdot), X^{\Delta'}(\cdot))\}$ must therefore satisfy $M(\cdot) = M^*(\cdot)$ almost surely. This implies that the sequence $\{(M^{\Delta'}(\cdot), X^{\Delta'}(\cdot))\}$ itself is converging to $(M^*(\cdot), X(\cdot))$.

To complete the proof, we note that, by (C.71), (C.5) can be rewritten as

$$\begin{aligned} &\lim_{\Delta'\to 0}[-\int \exp\{-r[s^{\Delta'}\left(\sum_{i=1}^{N}[f^{\Delta'}_{M_i}(1,a)+f^{\Delta'}_{X_i}(1,a)]\right)-\Gamma(f^{\Delta'}_{M}(.,a))]\}\,d\alpha(a)] \\ &= -\int \exp\{-r[s\left(\sum_{i=1}^{N}[f_{M_i}(1,a)+f_{X_i}(1,a)]\right)-\Gamma(f_{M}(.,a))]\}\,d\alpha(a). \quad \text{(C.73)}\end{aligned}$$

Thus, all the inequalities in the proof of Proposition C.2 must in fact be equations, i.e., we must have

$$\begin{aligned} &\lim_{\Delta'\to 0}[-\int \exp\{-r[s^{\Delta'}\left(\sum_{i=1}^{N}[f^{\Delta'}_{M_i}(1,a)+f^{\Delta'}_{X_i}(1,a)]\right)-\Gamma(f^{\Delta'}_{M}(.,a))]\}\,d\alpha(a)] \\ &= -\int \exp\{-r\lim_{\Delta'\to 0}[s^{\Delta'}\left(\sum_{i=1}^{N}[f^{\Delta'}_{M_i}(1,a)+f^{\Delta'}_{X_i}(1,a)]\right)-\Gamma(f^{\Delta'}_{M}(.,a))]\}\,d\alpha(a)] \text{(C.74)}\end{aligned}$$

as well as

$$\lim_{\Delta'\to 0} s^{\Delta'}\left(\sum_{i=1}^{N}[f^{\Delta'}_{M_i}(1,a)+f^{\Delta'}_{X_i}(1,a)]\right) = s\left(\sum_{i=1}^{N}[f_{M_i}(1,a)+f_{X_i}(1,a)]\right) \quad \text{(C.75)}$$

and

$$\lim_{\Delta'\to\infty}[-\Gamma(f^{\Delta'}_{M}(.,a))] = -\Gamma(f_M(.,a)) \quad \text{(C.76)}$$

for α-almost all $a \in A$.

In view of (C.75), convergence of expected incentive payments is guaranteed if the sequence $\{s^{\Delta'}\left(\sum_{i=1}^{N}[f^{\Delta'}_{M_i}(1,.)+f^{\Delta'}_{X_i}(1,.)]\right)\} = \{s^{\Delta'}\left(z^{\Delta'}\right)\}$ is uniformly integrable. The argument is similar to the one given in Lemma C.8: For any Δ', $s^{\Delta'}(.)$ satisfies

$$s^{\Delta'}\left(z^{\Delta'}\right) = \sum_{i=0}^{N}\sum_{\tau=1}^{1/\Delta'} \tilde{A}_i^{\Delta'\tau} s_i^{\Delta'\tau}, \quad \text{(C.77)}$$

where

$$z^{\Delta'} = \sum_{i=1}^{N} Z_i^{\Delta'}(1) = \sum_{i=1}^{N} k_i(\Delta')^{\frac{1}{2}} \sum_{\tau=1}^{1/\Delta'}(\tilde{A}_i^{\Delta'\tau}-\hat{p}_i), \quad \text{(C.78)}$$

and, for any τ and $i = 1, 2, ..., N$,

$$s_i^{\Delta'\tau} = \Delta'\hat{c}(\mu^{\Delta'\tau}) - \frac{1}{r}\ln\left(1 - r\hat{c}_i(\mu^{\Delta'\tau})k_i(\Delta')^{\frac{1}{2}} + r\sum_{j=1}^{N} p_j^{\Delta'}(\mu^{\Delta'\tau})\hat{c}_j(\mu^{\Delta'\tau})k_j(\Delta')^{\frac{1}{2}}\right), \quad \text{(C.79)}$$

and

$$s_0^{\Delta'\tau} = \Delta'\hat{c}(\mu^{\Delta'\tau}) - \frac{1}{r}\ln\left(1 + r\sum_{j=1}^{N} p_j^{\Delta'}(\mu^{\Delta'\tau})\hat{c}_j(\mu^{\Delta'\tau})k_j(\Delta')^{\frac{1}{2}}\right) \tag{C.80}$$

Upon using Taylor expansions, as before, we can write (C.77) as

$$\begin{aligned} s^{\Delta'}\left(z^{\Delta'}\right) &= \Delta'\sum_{\tau=1}^{1/\Delta'}\hat{c}(\mu^{\Delta'\tau}) + \sum_{i=1}^{N}\sum_{\tau=1}^{1/\Delta'}\hat{c}_i(\mu^{\Delta'\tau})[X_i^{\Delta'}(\tau\Delta') - X_i^{\Delta'}((\tau-1)\Delta')] \\ &\quad + \frac{r}{2}\sum_{i=1}^{N}\sum_{\tau=1}^{1/\Delta'}\Delta'\tilde{A}_i^{\Delta'\tau}\left[\hat{c}_i(\mu^{\Delta'\tau})k_i - \sum_{j=1}^{N}\hat{p}_j\hat{c}_j(\mu^{\Delta'\tau})k_j\right]^2 + O((\Delta')^{\frac{1}{2}}), \end{aligned} \tag{C.81}$$

with higher-order terms being uniformly small. Uniform integrability follows immediately. This completes the proof of Proposition 4. *Q.E.D.*