Uncertainty, Investment and Productivity with Relational Contracts

James M. Malcomson*
University of Oxford, U.K.

This version: April 19, 2018

Preliminary: please do not quote

*I thank Matthias Fahn, Michael Powell, Marta Troya Martinez and John Vickers for helpful comments on this paper, and the ESRC Centre SKOPE at the University of Oxford for its financial support of this research. This work was supported by Economic and Social Research Council grant number RES-557-28-5003.
Abstract

Recent literature on the effects of uncertainty on investment have focused on the option value from irreversibility of investment. However, Bloom et al. (2018) find that, with that model, a negative aggregate shock to total factor productivity, not just an increase in its uncertainty, is also required to capture the drop in investment and productivity in recessions. This paper shows that, in the presence of relational contracts, an increase in uncertainty alone reduces investment and productivity in the long run even if the parties are otherwise risk neutral. It first develops the theory to show that this is a general property of relational contract models. Then, to assess the practical significance, it uses a calibrated model with parameters based on those in Bloom et al. (2018) to show that this model can generate effects on investment of the magnitude of the negative aggregate shock in that paper purely with an increase in uncertainty. This raises the possibility that real business cycle models could be freed from negative technology shocks that some regard as controversial.

Keywords: Relational contracts, risk, investment, general capital, specific capital

JEL classification: C73, D82, D86
1 Introduction

Relational contracts, arrangements between parties for which the ongoing relationship between them plays an essential role in determining outcomes, are widely seen as an insightful way to view economic relationships in a variety of areas. See Malcomson (1999) for applications to employment and Malcomson (2013) for applications to supply relationships. This paper investigates the insights relational contracts offer to understanding the impact of uncertainty on investment. It shows, among other things, that greater uncertainty then results in lower general investment and productivity in the long run even with risk neutral parties. This raises the possibility that real business cycle models could be freed from the negative aggregate technology shocks that Bloom et al. (2018) find necessary to capture investment and productivity over the business cycle but describe (p. 2) as “controversial, as it suggests that recessions are times of technological regress.”

The financial crisis of 2007 caused a significant fall in the level of investment. Figure 1 shows the time series for gross fixed capital formation in the US. One reason for this fall and similar falls in previous recessions discussed in the literature is the increase in uncertainty engendered by the crisis. See Bloom (2014) for a survey. The explanation put forward in this literature is that, although firms are risk neutral and so unconcerned about the extent of uncertainty in the long run, the irreversibility of investment gives rise to a real option value to delaying investment that increases with uncertainty, as demonstrated by Dixit and Pindyck (1994). Bloom (2009) provides a calibration and simulation that nicely captures the rapid drop and rebound in output and employment for the three years following the 2007 crisis.

The option value model is essentially one of adjustment costs. With risk neutral firms, investment returns to the same long-run path even if the increased uncertainty persists. With the calibration in Bloom (2009), nearly all the adjustment is completed after 36 months so, at the time he was writing, it would have been plausible for investment not to have returned to its long-run path following the 2007 crisis. But a glance at Figure 1 shows that, although investment has continued to increase from 2009 to 2016, it has remained below where it would have been had it continued on the trend for the period 2002 to 2007 and has not returned to its previous trend level even 9 years after the crisis. It is hard to imagine that this can be solely the result of risk-neutral firms adjusting to a change in the level of uncertainty. It looks more like a shift downwards in the long-run path. In a general equilibrium setting, Bloom et al. (2018) find that a negative aggregate shock to total factor productivity, not just an increase in its uncertainty, is required to capture the drop in investment and productivity in recessions.

The present paper shows that, in the presence of relational contracts, greater uncertainty affects long-run equilibrium investment even when the parties are jointly risk neutral. In the standard model of relational incentive contracts between a principal
and an agent, for example MacLeod and Malcomson (1989), the agent chooses a level of performance each period (typically referred to as effort in the literature) that affects the payoff to the principal. But agent performance is not contractible in the sense that payment can be conditioned on it in an enforceable legal agreement. This may be either because it cannot be verified by third parties in court or because it is too complicated to describe in a legally enforceable way to be worthwhile. Whatever the reason, conventional performance-related payments are not available. Instead, incentives for performance are provided by a combination of bonuses related to output that, while not legally enforceable, the principal nevertheless finds it worthwhile to pay and the payoff from continuation of the relationship. A central question in the literature concerns what payment and performance the parties will deliver even though they are not legally obliged to do so. To that standard model, the present paper adds two things. The first is that the value of the agent’s performance to the principal is uncertain because it is affected by an iid shock each period. The second is that the parties may invest in capital that enhances the productivity of the relationship.

The main theoretical results in the paper are the following. The non-contractibility of performance, and the consequent use of a relational contract, reduces the agent’s performance below first best when productivity shocks are unusually favourable. Under plausible conditions, an increase in risk as measured by second-order stochastic dominance results in this applying to somewhat less favourable shocks as well, and thus also reduces the value of a relational contract, even though both principal and agent are risk neutral. So more risky projects become less attractive when incentives
for performance have to be provided by a relational contract, even with risk-neutral parties, and additional risk may thus deter the parties from engaging in a project. In that respect, the presence of risk makes the parties act in a way that appears risk averse. Use of a relational contract may also exaggerate the effect of uncertainty on investment in capital relative to what would happen with contractible performance. But the direction of that effect depends on whether the capital is general or specific in the sense of Becker (1975). With investments that are general in the sense of Becker (1975), increased uncertainty reduces investment. With investments that are specific in the sense of Becker (1975), the impact is less severe and may even increase investment, which is interesting in the light of the observation in Bloom (2014) that uncertainty appears to stimulate some types of investments.

To assess the empirical magnitudes of these theoretical effects, the paper uses a production framework similar to Bloom et al. (2018) but with the option value of delaying investment replaced by the constraints arising from a relational contract. With parameters based on those in Bloom et al. (2018), this formulation generates effects on general investment of the magnitude of the negative aggregate shock in that paper purely with greater uncertainty. It is this that raises the possibility that real business cycle models could be freed from unattractive negative aggregate technology shocks.

The paper is organized as follows. The next section discusses related literature. Section 3 sets out the model and the assumptions used for the analysis. Section 4 analyses optimal effort in a relational contract in this setting. Section 5 studies the effect of risk on the returns to the relationship for given capital stock. Section 6 studies the effect of risk on optimal investment in capital. Section 7 develops a version of the model suitable for calibration and provides results from that. Section 8 contains concluding remarks. Proofs of results are in Appendix A. Other appendices provide details of derivations and calculations for the empirical exercise.

2 Related literature

The literature on the option value associated with uncertainty and irreversible investment stems from Dixit and Pindyck (1994). The theoretical dynamics applied to aggregate investment are set out in Abel and Eberly (1996). The empirical implementations most closely related to the present paper are in Bloom et al. (2007), Bloom (2009) and Bloom et al. (2018). Many other contributions on capital adjustment costs are discussed in the survey by Bloom (2014). Among the papers not discussed there is Bond et al. (2011).

In the relational contracts literature, Malcomson (2015b) analyses the effect of productivity shocks, and Li and Matouschek (2013) the effect of shocks to the opportunity cost to the principal, but the models used there have no investment. Malcomson (2015a) studies investment but without productivity shocks, as does Garicano and
Rayo (2017) for investment in human capital that takes the form of knowledge transfer. Fahn et al. (2017a) and Fahn et al. (2017b) consider the implications of relational contracts for firm capital structure. Englmaier and Fahn (2017) consider the implications of investments into liquidity-generating capital for the ability of firms to meet their financial commitments in an uncertain relational contract setting. But none of these investigate the impact of uncertainty on fixed capital formation of the type represented in Figure 1. Hence the first task of the present paper is to develop the necessary theory in a relational contract setting. The paper then explores a formulation suitable for calibration and uses it to investigate the magnitude of the effect of increased uncertainty on capital investment and productivity.

3 The model

The model used here is essentially that of MacLeod and Malcomson (1989) with the two additions specified in the Introduction. First, to allow for risk, the productivity of the agent’s effort is subject to an iid shock each period. Second, the parties may make a capital investment that enhances the productivity of the relationship.

A principal uses an agent to perform a specific task each period. Both are risk neutral and discount the future with the same discount factor \( \delta \in (0, 1) \). The relationship between the two can, in principle, continue indefinitely. Output from the match in period \( t \) is \( y(e_t, K, \theta_t) \), where \( e_t \in [0, \bar{e}] \) is the employee’s effort at \( t \), \( K \in [0, \bar{K}] \) is the capital stock, and \( \theta_t \in [\theta_l, \theta_r] \) is an iid random variable that affects productivity, is distributed \( F(\theta, \sigma) \), with \( dF(\theta, \sigma) > 0 \) for all \( \theta \in [\theta_l, \theta_r] \) and \( \sigma \) a parameter that determines its riskiness, and is observed by both parties at the start of period \( t \), so there is no asymmetric information.\(^1\) Effort \( e_t \) in period \( t \) is chosen at cost \( c(e_t) \) to the agent after \( \theta_t \) is revealed and so can be conditioned on the shock. Neither output \( y(e_t, K, \theta_t) \) nor effort \( e_t \) is contractible in the sense that payment can be conditioned on performance in a formal legal agreement. This may be either because they cannot be verified by third parties or because it is too complicated to describe them in a legally enforceable way to be worthwhile. Effort can be thought of as anything unverifiable the agent may do that affects the payoff to the principal. In the context of employment, it could be literal effort. In the context of a supply chain, it could be the quality of the intermediate products supplied.

Modelling capital adjustment costs in the relational contract framework is technically complicated. Because the concern here is with long-run equilibrium properties, capital adjustment costs are not included in the model. But, for reasons given in the adjustment cost literature, it is implausible that capital can be fully adjusted to shocks in the short run, so here investment in capital is assumed to take place at the beginning of period.

\(^1\)For the calibrated version of the model, a trend in the mean of \( \theta \) can be taken account of through an adjustment to the discount factor \( \delta \), with the calibrated version applying to detrended data.
of the relationship, before any shock has been revealed, at a one-off cost $C(K)$. This cost is to be thought of as the present discounted cost of using capital $K$, including replacement investment.\footnote{For the calibration that follows, $C(K)$ is taken to be linear so, even if capital can be added each period, it is not optimal to build it up slowly over time.}

**Assumption 1** The functions $y$, $c$ and $C$ have the following properties:

1. For all $(K, \theta) \in [0, \bar{K}] \times [\underline{\theta}, \bar{\theta}]$, $y(0, K, \theta) = 0$ and, for $e \in (0, \bar{e})$, $y(e, K, \theta)$ is: (1) three times differentiable in all its arguments with $y_1(e, K, \theta) > 0$, $y_2(e, K, \theta) > 0$, $y_3(e, K, \theta) > 0$, $y_{11}(e, K, \theta) \leq 0$, $y_{12}(e, K, \theta) > 0$ and $y_{13}(e, K, \theta) > 0$, and (2) concave in $K$ for given $(e, \theta)$;

2. $c(0) = 0$ and, for all $e \in [0, \bar{e}]$, $c(e)$ is twice differentiable, with $c'(e) > 0$ and $c''(e) \geq 0$;

3. $y(e, K, \theta) - c(e)$ is strictly increasing in $e$ for $e = 0$ and for $e > 0$ sufficiently small, strictly decreasing in $e$ for $e = \bar{e}$, and strictly concave in $e$ for all $(K, \theta) \in [0, \bar{K}] \times [\underline{\theta}, \bar{\theta}]$;

4. $C(K)$ is convex and strictly increasing, with $C(0) = 0$ and $\lim_{K \to 0^+} C(K) \geq 0$, with $y(e, K, \theta) - C(K)$ strictly increasing in $K$ for $K = 0$ and for $K > 0$ sufficiently small, strictly decreasing in $K$ for $K = \bar{K}$, and strictly concave in $K$ for all $(K, \theta) \in [0, \bar{K}] \times [\underline{\theta}, \bar{\theta}]$.

The payoff to the principal in period $t$ of the match is $y(e_t, K, \theta) - W_t$, where $W_t$ is the payment to the agent in period $t$. The payoff to the agent is $W_t - c(e_t)$. Because both are risk neutral, the joint payoff to them both from being matched in a period is just the sum of their individual payoffs. For a period in which the principal and agent are not matched, the principal’s payoff is $\bar{u}(K, \sigma) \geq 0$ and the agent’s payoff $u(K, \sigma) \geq 0$, with $\bar{u}(K, \sigma) := u(K, \sigma) + \bar{v}(K, \sigma)$, the inequalities holding for all $K \in [0, \bar{K}]$. Conditional on $(e, K, \theta)$, the joint payoff to principal and agent from being matched in a period is $s(e, K, \theta) := y(e, K, \theta) - c(e)$. As a benchmark, if effort were contractible it would be set at the first-best level for given $(K, \theta)$, denoted $e^*(K, \theta)$, that maximizes this joint payoff and is given by

$$y_1(e^*(K, \theta), K, \theta) = c'(e^*(K, \theta)).$$

(1)

Capital is productive, so it may be optimal to have $K > 0$. For this, there are two benchmarks. First, if capital stock could be chosen in each period $t$ after $\theta_t$ had been revealed and effort were contractible, capital would be chosen at the first-best level given by

$$K_{FB}^*(\theta) \in \arg\max_{K \in [0, \bar{K}]} y(e^*(K, \theta), K, \theta) - c(e^*(K, \theta)) - C(K).$$

(2)
But if, as assumed, investment in capital has to be decided at the beginning of the relationship before any shock has been revealed with effort being contractible, capital stock would be chosen to satisfy

$$K^* (\sigma) \in \arg \max_{K \in [0, \bar{K}]} \frac{1}{1 - \delta} \int_{\theta} \left[ y(e^* (K, \theta), K, \theta) - c(e^* (K, \theta)) \right] dF(\theta, \sigma) - C (K).$$  \hspace{1cm} (3)

If output $y(e_t, K, \theta_t)$ or effort $e_t$ were contractible, it would in this context be straightforward to agree a contract that would deliver effort schedule $e^* (K, \theta)$ and investment $K^* (\sigma)$. However, with neither contractible, effort above the minimum level $e = 0$ can be sustained only by a relational contract. Following the analysis in MacLeod and Malcolmson (1989), payment $W_t$ to the agent in the relational contract has two components, a fixed component $w_t$ that is guaranteed independent of performance in period $t$ and a bonus component $b_t$ conditional on performance in period $t$. The bonus cannot be legally enforceable because performance is not contractible, so the relational contract must be such that it is in the principal’s interest to pay it. The next section analyses optimal effort with a relational contract for given investment in capital.

4 Optimal effort

A relational contract that is carried out even though not legally enforceable is referred to in the literature as self-enforcing. The set of self-enforcing contracts is largest when the punishments for deviation are, as in Abreu (1988), the most severe available, which here corresponds to the deviating party receiving the future payoff that would result from the relationship ending. Because the parties are risk neutral and use the same discount factor, they can redistribute the joint gains from their relationship in any way they choose by an upfront payment at the start of the relationship. It is thus optimal for them to select an equilibrium contract that maximises those joint gains at the start of the relationship. Moreover, by an argument in Levin (2003, Theorem 2) that applies to the model here, if an optimal contract exists, there are stationary contracts that are optimal. An optimal stationary contract depends only on current payoff-relevant information, so optimal effort and payment functions have the form $e_t = e(K, \theta_t)$, $w_t = w(K, \theta_t)$ and $b_t = b(e_t, K, \theta_t)$. The first result specifies the effort schedules that can be sustained by a relational contract.

**Proposition 1** An effort schedule $e(K, \theta)$ that generates expected joint payoff $S(K, \sigma)$ each period with capital stock $K$ can be implemented by a stationary contract if and only if

$$\frac{\delta}{1 - \delta} [S(K, \sigma) - s(K, \sigma)] \geq c(e(K, \theta)), \quad \text{for all } \theta \in [\theta, \bar{\theta}].$$  \hspace{1cm} (4)

In Proposition 1, $S(K, \sigma)$ is the expected joint payoff from one period of the relation-
ship, given investment in capital $K$ at the start, before the realization for that period of the shock $\theta$ from a distribution parameterized by $\sigma$. Thus the left-hand side of (4) is the joint payoff gain to both parties from continuing the relationship in future periods. So (4) states that, for effort $e(K, \theta)$ to be implementable, this joint gain must be at least as great as the cost to the agent of delivering effort $e(K, \theta)$ in a period. This is a standard result for relational contracts with risk-neutral parties, originally derived in MacLeod and Malcomson (1989). The essential intuition can be seen as follows. If the agent were to receive no bonus but instead all the joint payoff gain from continuing the relationship in the future, it would be worth incurring up to that amount in cost of effort to keep the relationship going. A positive bonus reduces by the amount of the bonus the amount of the future joint gain the agent must receive to be worth incurring given effort. But, because payment of the bonus cannot be enforced legally, the principal must receive future gain of at least the amount of the bonus to make it worthwhile paying that bonus to keep the relationship going. So the total of the future joint payoff gain required to sustain the relationship is unaffected by having a positive bonus. All the size of the bonus does is to affect the amounts of the minimum future joint gain that must go to the principal and to the agent. Thus, provided (4) is satisfied, there are different combinations of the fixed payment $w_t$ and the bonus $b_t$ that will induce effort schedule $e(K, \theta)$. These affect how the joint gain from the relationship is distributed between principal and agent. But no combination of these can induce effort schedule $e(K, \theta)$ if (4) is not satisfied. The following corollary is a straightforward consequence of Proposition 1.

**Corollary 1** Suppose, for given $S(K, \sigma)$, first-best effort $e^*(K, \theta)$ does not satisfy (4) for some $\theta' \in [\underline{\theta}, \bar{\theta}]$. Then first-best effort $e^*(K, \theta)$ does not satisfy (4) for any $\theta \in [\underline{\theta}, \bar{\theta}]$ with $\theta > \theta'$.

Corollary 1 follows from Proposition 1 because first-best effort given by (1) is increasing in $\theta$ and, hence, so is $c(e^*(K, \theta))$. But, because $\theta$ is an iid shock, the left-hand side of (4) is independent of $\theta$ because it is an expectation over the next period’s $\theta$. So, with the left-hand side of (4) independent of $\theta$ and the right-hand side increasing in $\theta$ when effort is first best, the result in the corollary follows directly. That provides an important step in determining the optimal effort schedule. The following assumption, used for the rest of the paper, avoids the uninteresting case in which no relational contract can sustain positive effort.

**Assumption 2** There exists an effort schedule with $e(K, \theta) \in (0, \bar{e})$ for all $\theta \in [\underline{\theta}, \bar{\theta}]$ for which (4) is satisfied with strict inequality for all $K \in (0, \bar{K}]$.

**Proposition 2** If a relationship is potentially mutually beneficial, an optimal stationary effort schedule $e(K, \theta)$ for given $K$ takes one of three forms$^3$:

$^3$It can be shown that the results in Proposition 2 hold even if only the principal observes $\theta$. In
1. First best: if

\[
\frac{\delta}{1-\delta} \int_{\tilde{\theta}}^{\bar{\theta}} [s(e^*(K, \theta), K, \theta) - \xi(K, \sigma)] dF(\theta, \sigma) \geq c\left(e^*(K, \tilde{\theta})\right),
\]

then \(e(K, \theta) = e^*(K, \theta)\) for all \(\theta \in [\tilde{\theta}, \bar{\theta}]\).

2. Full pooling: if

\[
\frac{\delta}{1-\delta} \int_{\tilde{\theta}}^{\bar{\theta}} [s(e^*(K, \theta), K, \theta) - \xi(K, \sigma)] dF(\theta, \sigma) \leq c\left(e^*(K, \tilde{\theta})\right),
\]

then \(e(K, \theta) = \bar{e}(K, \sigma)\) for all \(\theta \in [\tilde{\theta}, \bar{\theta}]\), where \(\bar{e}(K, \sigma)\) is given by

\[
\frac{\delta}{1-\delta} \int_{\tilde{\theta}}^{\bar{\theta}} [s(\bar{e}(K, \sigma), K, \theta) - \xi(K, \sigma)] dF(\theta, \sigma) = c(\bar{e}(K, \sigma)).
\] (5)

3. Partial pooling: otherwise, there exists \(\hat{\theta}(K, \sigma) \in (\tilde{\theta}, \bar{\theta})\) such that

\[
e(K, \theta) = \begin{cases} 
\ e^*(K, \theta), & \text{for } \theta \in [\tilde{\theta}, \hat{\theta}(K, \sigma)], \\
\ e^*(K, \hat{\theta}(K, \sigma)), & \text{for } \theta \in [\hat{\theta}(K, \sigma), \bar{\theta}],
\end{cases}
\]

with \(\hat{\theta}(K, \sigma)\) the highest value of \(\tilde{\theta} \in (\tilde{\theta}, \bar{\theta})\) satisfying

\[
\frac{\delta}{1-\delta} \left\{ \int_{\tilde{\theta}}^{\hat{\theta}} [s(e^*(K, \theta), K, \theta) - \xi(K, \sigma)] dF(\theta, \sigma) \\
\quad + \int_{\hat{\theta}}^{\bar{\theta}} [s(e^*K, \tilde{\theta}), K, \theta) - \xi(K, \sigma)] dF(\theta, \sigma) \right\} - c\left(e^*K, \hat{\theta}\right) = 0. \] (6)

Proposition 2 establishes that optimal effort is first-best for all \(\theta\), or the same for all \(\theta\), or the same for all \(\theta\) above some cutoff value \(\hat{\theta}(K, \sigma)\) and first-best for all \(\theta\) below \(\hat{\theta}(K, \sigma)\). The reasoning behind this is as follows. By Proposition 1, (4) is a necessary and sufficient condition for an effort schedule to be implementable. Thus optimal effort maximizes the joint payoff gain subject to that constraint. By definition, first-best effort maximizes the joint payoff when the constraint (4) is not imposed. It also maximizes the contribution effort for that \(\theta\) makes to \(S(K, \sigma)\), which weakens the constraint (4) for all other \(\theta\). So, if first-best effort satisfies (4) for some \(\theta\), it remains optimal for that \(\theta\) when (4) is added as a constraint. When this property holds for all \(\theta\), the result that case, the principal will reveal \(\theta\) truthfully provided the expected payoff from doing so is non-decreasing in \(\theta\) and this property is satisfied by the optimal effort in the proposition. To show that formally, however, complicates the exposition and so is not done here.
in Part 1 applies. In that case, the outcome is exactly the same as if effort were contractible. Of course, (4) may not be satisfied for all $\theta$ with the first-best effort schedule. Indeed, it may not be satisfied for any $\theta$. In that case, effort will have to be below first-best for all $\theta$. For each $\theta$ individually, it is then optimal to have effort as high as possible while satisfying the constraint (4) so that it is as close to first-best as possible. Moreover, that does not conflict with getting the highest effort possible for any other $\theta$ because, with effort below first-best, increasing effort for some $\theta$ increases the joint payoff $s(e, K, \theta)$, which relaxes the constraint (4) for all $\theta$. Because the left-hand side of (4) is independent of $\theta$, having (4) hold with equality for all $\theta$ implies the same effort for all $\theta$, so there is full pooling as in Part 2 of the proposition. To have the constraint (4) bind for all $\theta$ then implies that $e(K, \theta) = \tilde{e}(K, \sigma)$ for all $\theta$, where $\tilde{e}(K, \sigma)$ is given by (5).

The remaining possibility is that (4) is satisfied by first-best effort for some, but not all, $\theta$. It follows from Corollary 1 that, if it is satisfied by first-best effort for some types, it will also be satisfied for all lower types. As in Part 2, for all types for which (4) is not satisfied for first best effort, it is optimal to have effort at the highest level that satisfies (4) and that level is the same for all such types. That implies that there is a critical type $\tilde{\theta}(K, \sigma)$ such that effort is at the first-best level for all $\theta \leq \tilde{\theta}(K, \sigma)$ and is independent of $\theta$ for all $\theta > \tilde{\theta}(K, \sigma)$ at $e^*(\tilde{\theta}(K, \sigma), K)$. That gives the form of the optimal schedule in Part 3 of the proposition. (Because it is optimal to have as many types as possible with first-best effort, $\tilde{\theta}(K, \sigma)$ is the highest $\theta$ that satisfies (6) if there is more than one.) It follows from (6) that $\tilde{\theta}(K, \sigma)$ satisfies

\[
\frac{\delta}{1 - \delta} \left\{ \int_{\tilde{\theta}}^{\theta} s(e^*(K, \theta), K, \theta) dF(\theta, \sigma) - s(K, \sigma) 
- \int_{\tilde{\theta}(K, \sigma)}^{\theta} \left[ s(e^*(K, \theta), K, \theta) - s\left(\left( e^*\left( K, \tilde{\theta}(K, \sigma) \right) \right), K, \theta) \right] dF(\theta, \sigma) \right\} = c\left( e^*(K, \tilde{\theta}(K, \sigma)) \right).
\]

(7)

The second integral in this is strictly positive because $e^*(K, \theta)$ uniquely maximizes $s(e, K, \theta)$.

Whenever either Part 2 or Part 3 of Proposition 2 applies, output from the match (and so also productivity for the single agent) is lower than if effort were contractible.

5 The effect of risk on effort

This section considers the effect of the riskiness of the distribution $F(\theta, \sigma)$, parameterized by $\sigma$, on the optimal effort schedule derived in Section 4. A conventional way to measure riskiness is in terms of second-order stochastic dominance for distributions.
To analyse the cases in which effort under the relational contract is not first-best for all \( \theta \), it is useful to use a somewhat modified formulation of second-order stochastic dominance.

**Definition 1** \( F(\theta, \sigma_L) \) dominates \( F(\theta, \sigma_H) \) in the second-order stochastic sense for \( \bar{\theta} \in [\underline{\theta}, \bar{\theta}] \) if

\[
\int_{\underline{\theta}}^{\bar{\theta}} \left[ F(x, \sigma_L) - F(x, \sigma_H) \right] dx \geq 0, \quad \text{for all } \theta \in [\underline{\theta}, \bar{\theta}],
\]

with strict inequality for a set of values of \( x \in [\underline{\theta}, \bar{\theta}] \) with positive probability.

This definition corresponds to the standard definition when \( F(\theta, \sigma_L) \) has the same mean as \( F(\theta, \sigma_H) \) and \( \bar{\theta} = \theta \) (in which case the integral in Definition 1 equals zero for \( \theta = \underline{\theta} \), see Laffont (1989, p. 25).) It then corresponds to a mean-preserving spread in the sense of Rothschild and Stiglitz (1970). The reason for the modification is apparent from (7), the constraint that arises when optimal effort involves partial pooling. The first integral in (7) is just the joint payoff with contractible effort. So to evaluate the difference in payoff when effort is not contractible, the second integral needs to be evaluated and it is the lower limit of that integral that is variable, as in the integral in Definition 1.

To consider the effect on the optimal relational contract of changes in riskiness, the following result is useful. It is a straightforward adaptation to Definition 1 of the result for the standard definition of second-order stochastic dominance that a risk-averse agent prefers \( F(\theta, \sigma_L) \) to \( F(\theta, \sigma_H) \) if the former stochastically dominates the latter in the second-order sense. The proof is a straightforward adaptation of the standard proof for the case \( \bar{\theta} = \theta \) in Laffont (1989, p. 32-33).

**Lemma 1** Suppose a twice-differentiable function \( g(\theta) \) is decreasing and concave on \( [\underline{\theta}, \bar{\theta}] \) and \( g(\bar{\theta}) = 0 \) if \( \bar{\theta} > \theta \). Then, if \( F(\theta, \sigma_L) \) dominates \( F(\theta, \sigma_H) \) in the second-order stochastic sense for \( \bar{\theta} \in [\underline{\theta}, \bar{\theta}] \) and the two distributions have the same mean,

\[
\int_{\underline{\theta}}^{\bar{\theta}} g(\theta) dF(\theta, \sigma_L) \geq \int_{\underline{\theta}}^{\bar{\theta}} g(\theta) dF(\theta, \sigma_H),
\]

with the inequality strict if \( g(\theta) \) is strictly concave on \( [\underline{\theta}, \bar{\theta}] \).

With relational contracts, it is important to distinguish between idiosyncratic risk that affects only the relationship under consideration and systemic risk that also affects the payoffs if the parties separate. To define these formally, let \( \hat{S}(K, \sigma) \) denote the expected joint payoff each period before the realization for that period of the shock \( \theta \) from adopting the optimal effort schedule in Proposition 2.
Definition 2 Risk is idiosyncratic if $\hat{s}(K, \sigma)$ is independent of $\sigma$. Risk is systemic if $\hat{S}(K, \sigma) - s(K, \sigma)$ is independent of $\sigma$.

There are obviously partially idiosyncratic intermediates between these two extremes in which neither $s(K, \sigma)$ nor $\hat{S}(K, \sigma) - s(K, \sigma)$ is independent of $\sigma$ but the essential insights are illustrated by those two cases.

To apply the result in Lemma 1, define

$$h(e, K, \theta) := s(e, K, \theta) - s^*(e(K, \theta), K, \theta).$$

This is the difference in the joint payoff for some arbitrary effort $e$ and that for first-best effort for given $K$ and $\theta$, the negative of the integrand in the second integral in (7). Lemma 1 can thus be applied straightforwardly to that integral whenever $h(e, K, \theta)$ is decreasing and concave in $\theta$ because certainly $h(e, K, \theta)$ is zero for $e = e^*(K, \theta)$. Because $e^*(K, \theta)$ maximises $s(e, K, \theta)$, by the envelope theorem

$$h_3(e, K, \theta) = s_3(e, K, \theta) - s_3(e^*(K, \theta), K, \theta) < 0$$

for $e < e^*(K, \theta)$ (9) because $s_{13} = y_{13} > 0$. Thus, when $h(e, K, \theta)$ is also concave in $\theta$, Lemma 1 can be used to compare the second integral in (7) under $\sigma_H$ with that under $\sigma_L$. Since the first integral in (7) is just the effect when effort is contractible, this provides a measure of the difference in the response to risk when effort is not contractible from when it is.

Proposition 3 Consider $F(\theta, \sigma_L)$ that dominates $F(\theta, \sigma_H)$ in the second-order stochastic sense for $\hat{\theta} \in [\theta, \bar{\theta}]$. For systemic risk, $\hat{\theta}(K, \sigma_H) = \hat{\theta}(K, \sigma_L)$. For idiosyncratic risk, consider the following two cases:

1. Suppose the optimal effort schedule exhibits full pooling (that is, the case in Part 2 of Proposition 2) for both $F(\theta, \sigma_H)$ and $F(\theta, \sigma_L)$ and $\hat{\theta} = \theta$. Then for idiosyncratic risk and given $K$, if $s(\hat{e}(K, \sigma_H), K, \theta)$ is strictly concave in $\theta$ for $\theta \in [\theta, \bar{\theta}]$, effort is lower for all $\theta$, and the joint payoff is also lower, for $\sigma_H$ than for $\sigma_L$ and there is a larger difference for $\sigma_H$ than for $\sigma_L$ between the joint payoff when effort is contractible and when it is not.

2. Suppose the optimal effort schedule exhibits partial pooling (that is, the case in Part 3 of Proposition 2) for both $F(\theta, \sigma_H)$ and $F(\theta, \sigma_L)$ that have the same mean and $\hat{\theta} = \hat{\theta}(K, \sigma_H)$. Then for idiosyncratic risk and given $K$, if

$$\int_{\theta}^{\bar{\theta}} s(e^*(K, \theta), K, \theta) dF(\theta, \sigma_L) \geq \int_{\theta}^{\bar{\theta}} s(e^*(K, \theta), K, \theta) dF(\theta, \sigma_H)$$

and $h(e^*(K, \hat{\theta}(K, \sigma_H)), K, \theta)$ is strictly concave in $\theta$ for $\theta \in [\hat{\theta}(K, \sigma_H), \bar{\theta}]$, $\hat{\theta}(K, \sigma_H) < \hat{\theta}(K, \sigma_L)$, there is a larger difference for $\sigma_H$ than for $\sigma_L$ between the joint payoff when effort is contractible and when it is not, and the joint payoff is lower for $\sigma_H$ than for $\sigma_L$. 

11
The essence of Proposition 3 is that, while systemic risk has no impact on optimal effort for given capital stock, that is not the case with idiosyncratic risk even when both parties to the relational contract are individually risk neutral. The constraints required to make the contract self-enforcing result in their receiving a higher joint payoff from a less risky distribution. In this respect, it is as if they were risk averse for given capital $K$. Because $\hat{\theta}(K, \sigma_H) < \hat{\theta}(K, \sigma_L)$, effort is at first-best level for a larger interval of $\theta$ for $\sigma_L$ than $\sigma_H$. Moreover, because first-best effort is strictly increasing in $\theta$, $e^*(\hat{\theta}(K, \sigma_L), K) > e^*(\hat{\theta}(K, \sigma_H), K)$, so those $\theta$ for which effort is below first-best under $\sigma_L$ have higher effort than under $\sigma_H$. Thus all $\theta$ with effort below first-best under $\sigma_H$ have strictly higher effort under $\sigma_L$.

Part 2 of Proposition 3 depends on $h(e^*(K, \hat{\theta}(K, \sigma_H)), K, \theta)$ being strictly concave in $\theta$ for $\theta \in [\hat{\theta}(K, \sigma_H), \hat{\bar{\theta}}]$. The following lemma gives conditions for this.

**Lemma 2** $h(e, K, \theta)$ is strictly concave in $\theta$ for $\theta \in [\hat{\theta}(K, \sigma_H), \hat{\bar{\theta}}]$ for $e$ sufficiently close to $e^*(K, \theta)$.

### 6 Optimal capital

The two previous section considered optimal effort, and the effect of riskiness on that and on the joint payoff from the relational contract, for given capital stock. This section analyses optimal capital stock.

Optimal capital when effort is contractible is given by (3). Optimal capital in the case of non-contractible effort and partial pooling (Part 3 in Proposition 2) is the solution to

$$
\max_{K \in [0, \bar{K}]} \frac{1}{1 - \delta} \left\{ \int_{\hat{\theta}}^{\hat{\bar{\theta}}} s(e^*(K, \theta), K, \theta) dF(\theta, \sigma) \right. \\
\left. + \int_{\hat{\theta}(K, \sigma)}^{\hat{\bar{\theta}}} s \left( e^*(K, \hat{\theta}(K, \sigma)), K, \theta \right) dF(\theta, \sigma) \right\} - C(K)
$$

subject to $\hat{\theta}(K, \sigma)$ the highest value of $\tilde{\theta}$ satisfying (6). (11)

For the case of full pooling (Part 2 in Proposition 2), the corresponding maximand is the same except with $\hat{\theta}(K, \sigma) = \hat{\bar{\theta}}$ and $e^*(.)$ replaced by $\hat{e}(.)$ defined by (5). Whichever applies, denote optimal investment by $\hat{K}(\sigma)$. In the case of no pooling (Part 1 in Proposition 2), the corresponding maximum is the same but with $\hat{\theta}(K, \sigma) = \hat{\bar{\theta}}$, which is exactly the same as the maximand in (3) so optimal capital is $K^*(\sigma)$, just as when effort is contractible.

Becker (1975) made the distinction between investment in general capital that is equally valuable with a different agent and investment in specific capital that is valuable only with a specific agent. In the context of relational contracts, this distinction
has the additional importance, noted by Klein and Leffler (1981), that investment in specific capital has the effect of relaxing the implementability constraint (4). For the model here, the distinction is captured by the following assumption.

**Definition 3** Capital is general if $S(K, \sigma) - \bar{S}(K, \sigma)$ is independent of $K$. Capital is specific if $\bar{S}(K, \sigma)$ is independent of $K$.

There are, of course, intermediates between these two extremes in which $\bar{S}_1(K, \sigma)$ lies between 0 and $S_1(K, \sigma)$. The present paper follows the literature in focussing attention on the extreme cases.

### 6.1 General capital

When capital is general, the optimal capital stock when effort is not contractible, denoted $\hat{K}_G(\sigma)$, is lower than when effort is contractible whenever there is some pooling, as in the following proposition.

**Proposition 4** Suppose capital is general and the optimal effort schedule exhibits some pooling (Part 2 or Part 3 of Proposition 2 applies). Then $\hat{K}_G(\sigma) < K^*(\sigma)$.

This result is perhaps not surprising. When effort is not contractible to the extent that makes a difference to optimal effort, it is constrained below the first-best schedule for given capital stock for some values of $\theta$. Thus, given the complementarity of effort and capital in producing output, the marginal product of capital averaged over $\theta$ is lower than when effort is contractible and investment is less worthwhile.

The next result concerns the effect of riskiness on the level of general capital.

**Proposition 5** Suppose capital is general, the optimal effort schedule exhibits some pooling and the conditions of Parts 1 and 2 of Proposition 3 apply for full and partial pooling of efforts respectively. Suppose also $h_2(e^*(K, \hat{\theta}(K, \sigma_H)), K, \theta)$ is decreasing and strictly concave in $\theta$ for $\theta \in [\hat{\theta}(K, \sigma_H), \tilde{\theta}]$ in the case of partial pooling and $h_2(\bar{e}(K, \sigma_H), K, \theta)$ is decreasing and strictly concave in $\theta$ for $\theta \in [\tilde{\theta}, \hat{\theta}]$ in the case of full pooling. Then, with either systemic or idiosyncratic risk, $K^*(\sigma_H) - \hat{K}_G(\sigma_H) > K^*(\sigma_L) - \hat{K}_G(\sigma_L)$.

This result establishes that, under the conditions specified, a more risky distribution increases the difference between the optimal amount of general capital when effort is contractible and when effort is not contractible. This applies even to increases in systemic risk which, by Proposition 3, do not affect $\hat{\theta}(K, \sigma)$ for given $K$. The reason is that the presence of the second integral in the maximand in (11) makes the objective function more concave than when that effort choice is unconstrained. This effect is exacerbated with increases in idiosyncratic risk that, again by Proposition 3, also lower $\hat{\theta}(K, \sigma)$ so that the set of $\theta$ for which effort is first best is reduced.
Proposition 5 depends on \( h_2(e, K, \theta) \) being decreasing and strictly concave in \( \theta \) for \( e \leq e^*(K, \theta) \). The following lemma gives a condition on \( y(e, K, \theta) \) that is sufficient for the first of these.

**Lemma 3** A sufficient condition for \( h_2(e, K, \theta) \) to be decreasing in \( \theta \) for \( e \leq e^*(K, \theta) \) is \( y_{123}(e, K, \theta) \geq 0 \).

Conditions for \( h_2(e, K, \theta) \) to be strictly concave in \( \theta \) are better considered for specific examples.

### 6.2 Specific capital

The most important distinction between general and specific capital in the context of relational contracts is that an increase in specific capital weakens the incentive constraint for self-enforcement. That can be seen from (4). For any given effort schedule, additional specific capital increases \( S(K, \sigma) \) while, by definition of its being specific, leaves \( s(K, \sigma) \) unchanged. Then, as shown in Malcomson (2015a), the optimal level of capital is higher when capital is specific than when it is general whenever that constraint is binding. The formal result, with \( \hat{K}_S(\sigma) \) denoting the optimal level of capital when it is specific and effort is non-contractible, is as follows.

**Proposition 6** For effort non-contractible and the optimal effort schedule exhibiting some pooling (Part 2 or Part 3 of Proposition 2 applies), \( \hat{K}_S(\sigma) > \hat{K}_G(\sigma) \) and, for \( y_{12} \) sufficiently small, \( \hat{K}_S(\sigma) > K^*(\sigma) \).

Proposition 6 establishes not only the result that, with non-contractible effort, the optimal level of specific capital is higher than that of general capital but also that, under certain circumstances, it is higher even than the optimal level when effort is contractible. This occurs when the reduced return that results from the non-contractibility of effort is more than offset by the relaxation of the implementability constraint. The next result concerns the effects of the level of riskiness on the level of specific capital.

**Proposition 7** Suppose the optimal effort schedule involves either full or partial pooling (that is, Part 2 or Part 3 of Proposition 2 applies) for both distribution \( F(\theta, \sigma_H) \) and distribution \( F(\theta, \sigma_L) \) that dominates \( F(\theta, \sigma_H) \) in the second-order stochastic sense for \( \theta \) for full pooling and for \( \hat{\theta}(K, \sigma_H) \) for partial pooling and, in the latter case, has the same mean and the distributions satisfy (10) for given \( K \). Then with either systemic or idiosyncratic risk, for \( y_{133}(e, K, \theta) \geq 0 \) for all \( (e, K, \theta) \) (with strict inequality for full pooling), \( \hat{K}_S(\sigma_H) - \hat{K}_G(\sigma_H) > \hat{K}_S(\sigma_L) - \hat{K}_G(\sigma_L) \) and, for \( y_{12} \) sufficiently small in addition, \( \hat{K}_S(\sigma_H) - K^*(\sigma_H) > \hat{K}_S(\sigma_L) - K^*(\sigma_L) \).

Proposition 7 gives conditions under which a more risky distribution results in a greater difference between the optimal levels of specific and general capital. It may
even result in a greater difference between optimal specific capital when effort is not contractible than when effort is contractible. As with Proposition 6, the reason is that additional specific capital relaxes the implementability constraint in a relational contract. When a change in $\sigma$ reduces the joint payoff $S(K, \sigma)$ as in Proposition 3, it tightens the constraint (4) and so increases the return to specific capital that relaxes that constraint. It thus results in an increase in specific capital. Note that the condition $y_{133}(e, K, \theta) \geq 0$ is sufficient but not necessary. Moreover, it holds for the functional forms in the following example.

**Example 1** For $y(e, K, \theta) = \theta \hat{y}(e, K)$,

$$y_3(e, K, \theta) = \hat{y}(e, K); \quad y_{33}(e, K, \theta) = 0; \quad y_{133}(e, K, \theta) = 0.$$  

So, for this example, the condition $y_{133}(e, K, \theta) \geq 0$ is satisfied.

The overall conclusions from this section are as follows. When capital is general, the optimal amount of capital within a relational contract when effort is non-contractible is less than when effort is contractible. Moreover, increased riskiness increases the difference. But when capital is specific and the relational contract restricts effort for some outcomes below first best, it relaxes the implementability constraint and so will be higher than if capital were general and, in some cases, higher even than if effort were contractible. Moreover, the difference can actually be greater with a more risky distribution than with a less risky one.

### 7 A calibrated model

To illustrate the quantitative significance of the theoretical results derived in the previous sections, consider the specification

$$y(e, K, \theta) = \theta^\gamma K^\alpha e^\beta, \quad \alpha, \beta, \gamma > 0, \quad \alpha + \beta \leq 1;$$  

(12)

$$c(e) = ce^n, \quad c > 0, n \geq 1;$$  

(13)

$$C(K) = CK^k, \quad C > 0, k \geq 1.$$  

(14)

The revenue function in (12) is that in Bloom et al. (2018) except that the labour hours input is replaced by the agent’s effort. Bloom (2009) shows how the exponents $\gamma, \alpha$ and $\beta$ relate to those in an underlying Cobb-Douglas production function and an isoelastic demand curve for the product. Bloom et al. (2018) use a labour cost function that is linear in labour hours with the addition of adjustment costs. Here that is replaced by the effort cost function (13). They use a capital cost function that is also linear with the addition of adjustment costs. The linear part of this is generalized in (14) to allow $k > 1$. Because the model here is stationary and capital is chosen before shocks are revealed, there is no role for capital adjustment costs.
For this specification, optimal choice of effort conditional on $K$ and $\theta$ is

$$e^* (K, \theta) = \left( \frac{\beta}{nc} \theta^\gamma K^K \right)^{1/n}.$$  \hfill (15)

(Derivations for this section are given in Appendix B.) Then

$$s (e^* (K, \theta), K, \theta) = \left( 1 - \frac{\beta}{n} \right) \left( \frac{\beta}{nc} \right)^{\beta/n} \theta^{\gamma/n} K^{-\beta/n},$$  \hfill (16)

which is concave in $\theta$ for $\gamma \leq 1 - \beta/n$, and

$$s (e^* (K, \hat{\theta}), K, \theta) = \left( \frac{\beta}{nc} \right)^{\beta/n} \hat{\theta}^{-\gamma/n} K^{-\beta/n} \left( \frac{\theta}{\hat{\theta}} \right)^\gamma \frac{\beta}{n}.$$  \hfill (17)

Expected revenue per agent (labour productivity in the case of employment) is

$$K^{\alpha - \beta/n} \left( \frac{\beta}{nc} \right)^{\beta/n} \left[ \int_\hat{\theta}^{\theta} \theta^{-\beta/n} dF (\theta, \sigma) + \hat{\theta}^{\gamma/n} \int_0^{\hat{\theta}} \theta^\gamma dF (\theta, \sigma) \right].$$  \hfill (18)

For his model without relational contracts, Bloom (2009, footnote 17) argues, following a discussion in Abel and Eberly (1996), that it is appropriate to calibrate the model to avoid long-run effects of uncertainty reducing or increasing output, which corresponds here to setting $\gamma = 1 - \beta/n$ in (16). That makes the parties jointly risk-neutral in the absence of a relational contract, which is a good benchmark for determining the effect of having a relational contract, and so adopted here.

For the calibrations, it fits with Bloom et al. (2018) to take $\theta$ as log-normally distributed with risk parameterized by setting $\sigma$ equal to the standard deviation of $\ln \theta$. (For more on this, see below.) For this distribution, full pooling is never an optimal relational contract because it would require $\hat{e} (K, \sigma) \leq e^* (K, 0) = 0$, so continuing the relationship would not be worthwhile. The same applies to an optimal relational contract with first-best effort for all $\theta$ as long as the future joint payoff is finite because then, with first-best effort given by (15), (4) cannot be satisfied as $\theta$ goes to infinity. Thus partial pooling is the only case to consider. Rather than solve directly for optimal capital, $\hat{K}^G (\sigma)$ and $\hat{K}^S (\sigma)$, it is convenient to solve for the optimal cutoff values $\hat{\theta} (K, \sigma)$ below which effort is first-best when capital is chosen optimally. Denote these cutoff values by $\hat{\theta}^G (\sigma)$ and $\hat{\theta}^S (\sigma)$ for general and specific capital respectively.

When capital is general, $S (K, \sigma) - s (K, \sigma)$ in Proposition 1 is by definition independent of $K$, so it is convenient to define

$$\hat{S} (\sigma) = \frac{\delta}{1 - \delta} [S (K, \sigma) - s (K, \sigma)].$$  \hfill (19)
It can then be shown that the first-order condition for optimal capital yields that

\[
\hat{S}(\sigma)^{1-\left(1-\frac{\beta}{\bar{n}}\right)\frac{k}{\bar{\alpha}}} \hat{\theta}^G(\sigma)^{\left(1-\frac{\beta}{\bar{n}}\right)\frac{k}{\bar{\alpha}}-1} \left\{ E(\theta \mid \sigma) - \int_{\hat{\theta}^G(\sigma)}^{\hat{\theta}^G(\sigma)} \left[ \theta - \hat{\theta}^G(\sigma)^{\frac{\beta}{\bar{n}}} \theta^{1-\frac{\beta}{\bar{n}}} \right] dF(\theta, \sigma) \right\} (20)
\]

is independent of \( \sigma \). There exist values of the parameters \( c \) and \( C \) (neither of which enter (20)) for which this first-order condition corresponds to a maximum as long as the expression in (20) is positive and increasing in \( \hat{\theta}^G(\sigma) \). It can also be shown that

\[
\frac{\hat{K}^G(\sigma_L)}{\hat{K}^G(\sigma_H)} = \left[ \frac{\hat{S}(\sigma_L) \hat{\theta}^G(\sigma_H)}{\hat{S}(\sigma_H) \hat{\theta}^G(\sigma_L)} \right]^{\frac{1-\beta/n}{\bar{\alpha}}}. \tag{21}
\]

For systemic risk, \( \hat{S}(\sigma) \) is also independent of \( \sigma \). For given parameter values and distributions \( F(\theta, \sigma) \), one can then use (20) to solve for \( \hat{\theta}^G(\sigma) \) for any given value \( \hat{\theta}^G(\sigma_L) \), and (21) to solve for the ratio \( \hat{K}^G(\sigma_L) / \hat{K}^G(\sigma_H) \), independently of \( \hat{S}(\sigma) \). Systemic risk seems the natural interpretation of aggregate shocks such as the 2007 crisis, so this provides a way of assessing the model for general capital against deviations from trend in the data underlying Figure 1.

When capital is specific, \( \hat{S}(K, \sigma) \) is by definition independent of \( K \) and so will be written, \( \hat{S}(\sigma) \). It can then be shown that the first-order condition for optimal capital yields the result that the expression

\[
\left( \frac{\delta}{1-\delta} \hat{S}(\sigma) \right)^{1-\left(1-\frac{\beta}{\bar{n}}\right)\frac{k}{\bar{\alpha}}} \left\{ \frac{\delta}{1-\delta} \left(1 - \frac{\beta}{n}\right) E(\theta \mid \sigma) - \frac{\beta}{n} \hat{\theta}^S(\sigma) \right\} \left\{ E(\theta \mid \sigma) - \frac{\delta}{1-\delta} \int_{\hat{\theta}^S(\sigma)}^{\hat{\theta}^S(\sigma)} \left[ \left(1 - \frac{\beta}{n}\right) \theta - \hat{\theta}^S(\sigma)^{\frac{\beta}{\bar{n}}} \theta^{1-\frac{\beta}{\bar{n}}} + \frac{\beta}{n} \hat{\theta}^S(\sigma) \right] dF(\theta, \sigma) \right\} \left(1-\frac{\beta}{\bar{n}}\right)\frac{k}{\bar{\alpha}}-1 \right\}
\]

\[
-\frac{\delta}{1-\delta} \int_{\hat{\theta}^S(\sigma)}^{\hat{\theta}^S(\sigma)} \left[ \left(1 - \frac{\beta}{n}\right) \theta - \hat{\theta}^S(\sigma)^{\frac{\beta}{\bar{n}}} \theta^{1-\frac{\beta}{\bar{n}}} + \frac{\beta}{n} \hat{\theta}^S(\sigma) \right] dF(\theta, \sigma)
\]

\[
+ \frac{1}{\beta/n} \int_{\hat{\theta}^S(\sigma)}^{\hat{\theta}^S(\sigma)} \left[ \frac{\delta}{1-\delta} \left(1 - \frac{\beta}{n}\right) E(\theta \mid \sigma) \right] \left\{ E(\theta \mid \sigma) - \frac{\delta}{1-\delta} \int_{\hat{\theta}^S(\sigma)}^{\hat{\theta}^S(\sigma)} \left[ \left(1 - \frac{\beta}{n}\right) \theta - \hat{\theta}^S(\sigma)^{\frac{\beta}{\bar{n}}} \theta^{1-\frac{\beta}{\bar{n}}} + \frac{\beta}{n} \hat{\theta}^S(\sigma) \right] dF(\theta, \sigma) \right\} \frac{1}{1-\delta} \int_{\hat{\theta}^S(\sigma)}^{\hat{\theta}^S(\sigma)} \left( \hat{\theta}^S(\sigma)^{\frac{\beta}{\bar{n}}-1} \theta^{1-\frac{\beta}{\bar{n}}-1} \right) dF(\theta, \sigma)
\]

\[
- \frac{\beta}{n} \hat{\theta}^S(\sigma) \left\{ \frac{\delta}{1-\delta} \int_{\hat{\theta}^S(\sigma)}^{\hat{\theta}^S(\sigma)} \left( \hat{\theta}^S(\sigma)^{\frac{\beta}{\bar{n}}-1} \theta^{1-\frac{\beta}{\bar{n}}-1} \right) dF(\theta, \sigma) \right\}
\]
is independent of $\sigma$. There exist values of the parameters $c$ and $C$ (neither of which enter (22)) for which this first-order condition corresponds to a maximum as long as the expression in (22) is positive and decreasing in $\hat{\theta}^S(\sigma)$. It can also be shown that

$$\hat{K}^S(\sigma) = \left(\frac{\delta}{1-\delta}\hat{\theta}^S(\sigma)\right)^{\frac{1-\beta/n}{\alpha}} \left(\frac{\beta}{n\hat{c}}\right) \frac{\delta}{1-\delta} \left[1 - \left(1 - \frac{\beta}{n}\right)E(\theta \mid \sigma) - \frac{\beta}{n}\hat{\theta}^S(\sigma)\right]$$

$$- \frac{\delta}{1-\delta} \int_{\hat{\theta}^S(\sigma)}^{\tilde{\theta}} \left[\left(1 - \frac{\beta}{n}\right)\theta - \hat{\theta}^S(\sigma)\left(\frac{\beta}{n}\theta^{1-\frac{\beta}{n}} + \frac{\beta}{n}\hat{\theta}^S(\sigma)\right)\right] dF(\theta, \sigma) \right]^{\frac{1-\beta/n}{\alpha}}. \quad (23)$$

When capital is specific, the simplest case is a change in idiosyncratic risk because then $\xi(\sigma)$ is independent of $\sigma$. Then, for given parameter values and distributions $F(\theta, \sigma_H)$ and $F(\theta, \sigma_L)$, one can use (22) to solve for $\hat{\theta}^S(\sigma_H)$ for given $\hat{\theta}^S(\sigma_L)$, and (23) to solve for the ratio $\hat{K}^S(\sigma_L) / \hat{K}^S(\sigma_H)$, independently of $\xi(\sigma)$.

Equations (21) and (23) are for differences in capital, whereas the data in Figure 1 are for gross investment. In the comparison of stationary equilibria considered here, differences in capital result in differences in investment for two reasons. The first is in investment to replace depreciation of the different capital stocks in continuing firms. The second is in investment by new firms that replace those that go out of business and whose capital becomes valueless. In Bloom et al. (2018), depreciation is assumed to be a constant proportion of capital, so the percentage change in investment to replace depreciation as between two stationary equilibria is the same as the percentage change in capital. The same applies to investment by new firms if these are a constant proportion of the total.

The formulation in this section is calibrated using parameters based on those in Bloom et al. (2018). As already noted, Bloom et al. (2018) uses a Cobb-Douglas revenue function like that in (12) but with $e$ being the total labour input measured in hours. The long-run costs of both capital and labour (that is, excluding adjustment costs) are the going market rates, which correspond to $n = k = 1$ in (13) and (14). The parameters $\alpha$ and $\beta$ are calibrated from factor shares adjusted for an isoelastic product demand function with a 33% markup. It is, however, worth noting that results calculated from (20)–(23) depend on $\beta$ and $n$ only through their ratio $\beta/n$, so they are robust to changes in those parameters that leave their ratio unchanged. Because Bloom et al. (2018) are concerned with business cycles, total factor productivity in their model (corresponding to $\theta^\gamma$) is time varying. It is the product of separate aggregate ($A$) and firm ($Z$) components that switch between two regimes, low risk ($L$) and high risk ($H$), all of which follow autoregressive processes with normally distributed logs of innovations. The counterpart to these processes in the stationary long-run framework used here is that $\theta$ is log-normally distributed, with the low risk and high risk regimes the two different long-run equilibria corresponding to $\sigma_L$ and $\sigma_H$ respectively in the theoretical

18
model of the previous sections. However, the estimated standard deviations in Bloom et al. (2018) apply to \( \theta^*_G \) and so need to be adjusted correspondingly. Moreover, for the purpose of showing the effects of changes in long-run uncertainty, the aggregate and idiosyncratic components are combined using the standard formula for the product of log-normal distributions to give the low risk and high risk parameters for \( \theta^*_G \) of \( \sigma_L \) and \( \sigma_H \) respectively. The resulting low and high risk distributions for \( \theta \) are specified to have a mean of 1 to ensure that the parties are jointly risk neutral in the absence of the relational contract constraint (which corresponds to the expression in (16) being linear in \( \theta \) with the calibration \( \gamma = 1 - \beta/n \)). The resulting parameter values are given in Table 1. (Details of calculations for this section are given in Appendix C.) It is worth noting that the aggregate uncertainty change is essentially swamped by the idiosyncratic uncertainty change because \( \sigma_H/\sigma_L \approx \sigma_Z^H/\sigma_Z^L \).

To fully calibrate the relational contract model requires specifications for \( \delta \) and \( \xi(K, \sigma) \). The former is straightforward but the latter has no counterpart in Bloom et al. (2018). There also seems no obvious way to derive it from data. But, given the other parameters, each combination of \( \delta \) and \( \xi(K, \sigma) \) implies a cutoff \( \hat{\theta}^i(\sigma) \), for \( i \in \{ G, S \} \), for which effort for \( \theta > \hat{\theta}^i(\sigma) \) is constrained to \( e^*(K, \hat{\theta}^i(\sigma)) \), that is below first best. So, to illustrate the effect of a change in risk, the effects of the change are calculated for different possible values of \( \hat{\theta}^i(\sigma_L) \), which are to be interpreted as values on the domain of a log-normal distribution with mean 1.

Consider first when all capital is general. Then \( \hat{S}(\sigma) \) is independent of \( \sigma \) when risk is systemic, in which case the effect on the ratio of the left-hand side of (20) is independent of \( \hat{S}(\sigma) \). Table 2 gives the effect of the change from \( \sigma_L \) to \( \sigma_H \) in Table 1 for the specified values of \( \hat{\theta}^G(\sigma_L) \) on the assumption that all risk is systemic, with \( \hat{\theta}^G(\sigma_H) \) calculated from (20), and the capital changes from (21), for the specified \( \hat{\theta}^G(\sigma_L) \). The productivity changes are the expected changes in revenue per agent given the distribution of \( \theta \) for the specified \( \hat{\theta}^G(\sigma_L) \), calculated from (18). It can, moreover, be shown that the impact of an increase in idiosyncratic risk on general capital is greater in magnitude
Table 2: Effect of increase in systemic risk with general capital for given $\hat{\theta}^G (\sigma_L)$

<table>
<thead>
<tr>
<th>$\hat{\theta}^G (\sigma_L)$</th>
<th>0.50</th>
<th>0.75</th>
<th>1.00</th>
<th>1.25</th>
<th>1.50</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\theta}^G (\sigma_H)$</td>
<td>0.51</td>
<td>0.78</td>
<td>1.06</td>
<td>1.30</td>
<td>1.53</td>
</tr>
<tr>
<td>Capital change (%)</td>
<td>-1.65</td>
<td>-3.46</td>
<td>-5.95</td>
<td>-4.14</td>
<td>-2.23</td>
</tr>
<tr>
<td>Productivity change (%)</td>
<td>-3.27</td>
<td>-6.79</td>
<td>-11.55</td>
<td>-8.11</td>
<td>-4.41</td>
</tr>
</tbody>
</table>

(that is, with negative numbers more negative) than that of an increase in systemic risk. Thus the numbers in Table 1 can be taken as lower bounds on the magnitudes of the effects on capital if some risk is idiosyncratic.

In Table 2, the effect of the change in risk is larger for $\hat{\theta}^G (\sigma_L)$ nearer 1, the mean of the distribution of $\theta$. For $\hat{\theta}^G (\sigma_L)$ equal to that mean, capital falls by almost 6% and productivity by more than 11%. But even for $\hat{\theta}^G (\sigma_L)$ at 50% above or below the mean, the effects are substantial. Recall that these effects would be precisely zero in the absence of the incentive compatibility constraint for a relational contract (which corresponds to the limit as $\hat{\theta}^G (\sigma_L)$ goes to infinity). Recall also that the numbers in Table 2 are lower bounds on the magnitudes of the effects on general capital. So they correspond to a lower bound on the effects of an increase in risk that would have no effect in the absence of a relational contract. This is potentially significant in view of the conclusion in Bloom et al. (2018) that, even with adjustment costs, a negative aggregate total factor productivity shock seems necessary to match the data on recessions in the absence of relational contracts. Bloom et al. (2018, p. 21) use a -2% exogenous first moment aggregate shock alongside the second moment shock in a numerical experiment to fit their model to the data. With the calibration used here, $E (\theta \mid \sigma_H)$ 2% lower than $E (\theta \mid \sigma_L)$ would result in an approximately 4% reduction in general capital in the absence of relational contracts, which is well within the range in Table 2 that the increase in the second moment alone would generate with a relational contract. Thus, in the presence of relational contracts, one does not need a negative first moment shock to total factor productivity to generate the negative consequences of the negative first moment shock required, in the absence of relational contracts, of a size required to fit the data.

Specific capital may well be only a small proportion of total capital, so the effects of an increase in risk may not in this case be expected to be apparent in aggregate data. But numerical calculations serve to illustrate the theoretical results. With idiosyncratic risk, for which $\hat{s} (\sigma)$ is independent of $\sigma$, the parameter values in Table 1 and a lognormal distribution of $\theta$ with mean 1, only values of $\hat{\theta}^S (\sigma_L)$ lying between approximately 3.24 and 77.4 can be optimal for any positive values of $c$ and $C$. These are all well above the mean value of $\theta$. That lower values of $\hat{\theta}^S (\sigma_L)$, which imply a smaller interval of $\theta$ for which effort is first-best conditional on $K$, are never optimal reflects the result that additional specific capital relaxes the implementability constraint (4). This effect is magnified by the assumption in the calculation that all capital is specific. With
Table 3: Effect of increase in idiosyncratic risk with specific capital for given $\hat{\theta}^S (\sigma_L)$

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\theta}^S (\sigma_L)$</td>
<td>75.00</td>
</tr>
<tr>
<td>$\hat{\theta}^S (\sigma_H)$</td>
<td>14.23</td>
</tr>
<tr>
<td>Capital change (%)</td>
<td>2434.2</td>
</tr>
</tbody>
</table>

a lower factor share for specific capital than implied by the value of $\alpha$ used in the calculations, it would be smaller. The same applies to the effect of an increase in risk, for which results are given in Table 3. For that reason, the only effects given in Table 3 are chosen to illustrate the theoretical result that, under relational contracts, an increase in uncertainty may increase optimal specific investment when the parties are jointly risk neutral. Effects of the magnitude given there are unlikely to be observed in practice. But this result is interesting in the light of the comment in Bloom (2014, p. 153) that uncertainty appears to stimulate some investments.

8 Conclusion

This paper has analysed the effect of changes in the riskiness of the distribution of output for given agent performance in a relational contract in which agent performance is not contractible and, in particular, the effect on investment in capital. Unlike when performance is contractible, the effect on investment depends on whether capital is general or specific in the sense of Becker (1975). Investment in general capital is lower with a relational contract than with contractible performance and, under plausible conditions, is more adversely affected by an increase in uncertainty. Investment in specific capital is less adversely affected by reliance on a relational contract and may even increase with an increase in uncertainty. The reason is that, as noted by Klein and Leffler (1981), specific capital relaxes the incentive constraint in a relational contract because it increases the payoff to the parties staying together relative to the payoff from separating and it is this difference which constrains the performance that can be sustained by a relational contract.

An implication of these theoretical results is that, in the presence of relational contracts, an increase in uncertainty alone reduces long-run equilibrium investment in general capital and productivity when it would not otherwise do so — even, that is, if the parties are risk neutral. This is potentially important given that relational contracts are widely seen as an insightful way to view economic relationships in a variety of contexts.

To assess the potential empirical significance, the paper uses a calibrated production framework similar to that in Bloom et al. (2018) but with the option value of delaying investment replaced by the constraints arising from a relational contract to illustrate the magnitude of the negative effect of increased uncertainty on general capital. Bloom et al. (2018) find that, without relational contracts, a negative first moment
shock to aggregate total factor productivity seems necessary to match the data on recessions even with allowance for multiple adjustment costs. They comment (p. 2) that “(t)he reliance on negative technology shocks has proven to be controversial, as it suggests that recessions are times of technological regress.” The calibration exercise here shows that the same impact as their negative first moment shock is well within the range that a higher second moment alone could plausibly deliver as an equilibrium effect in the presence of relational contracts. This is suggestive that recessions of the magnitudes in the data might arise solely from a second moment shock to aggregate total factor productivity, with no negative first moment shock. To verify that obviously requires a relational contract model that, unlike the one used here, takes proper account of dynamics — a far from straightforward task. But if verified in that way, it would open up the attractive possibility of freeing real business cycle models from the implication that recessions are, on average, times of overall technological regress.

Appendix A  Proofs

Proof of Proposition 1. To simplify notation in the proof, drop the arguments \((K, \sigma)\) because these are predetermined at the time effort decisions are made and let \(u(\theta)\) denote the payoff to the agent in each period of a stationary relational contract in which \(\theta\) is realized. Then, if both parties stick to the contract,

\[
u(\theta) = w(\theta) + b(e(\theta), \theta) - c(e(\theta)).\]

(24)

So, having observed \(\theta\), the payoff to the agent from sticking to the contract at every date \(t\), conditional on the principal doing so, is

\[
u(\theta) + \frac{\delta}{1 - \delta} E_{\theta'} u(\theta')\]

(25)

because \(\theta\) is an iid draw each period. For the principal, let \(v(\theta)\) denote the payoff in each period of a stationary relational contract in which \(\theta\) is realized. Then, if both parties stick to the contract,

\[
v(\theta) = y(e(\theta), \theta) - w(\theta) - b(e(\theta), \theta).\]

(26)

So, having observed \(\theta\), the payoff to the principal from sticking to the contract at every date \(t\), conditional on the agent doing so, is

\[
v(\theta) + \frac{\delta}{1 - \delta} E_{\theta'} v(\theta')\]

(27)

Necessity. By setting \(e_t = 0\) and quitting the following period, the agent can, even
with no bonus, guarantee a payoff of

\[ w(\theta) + \frac{\delta}{1 - \delta} u. \]  

(28)

A necessary condition for the agent to stick to the contract for all \( \theta \) is that the payoff in (25) is no less than that in (28) for all \( \theta \). With the use of (24), that condition can be written

\[ \frac{\delta}{1 - \delta} \left[ E_{\theta'} u(\theta') - u \right] \geq c(e(\theta)) - b(e(\theta), \theta), \text{ for all } \theta \in [\theta, \bar{\theta}]. \]  

(29)

Once the agent has incurred effort \( e(\theta) \), the principal can, by setting \( b_t = 0 \) and quitting the following period, guarantee a payoff of

\[ y(e(\theta), \theta) - w(\theta) + \frac{\delta}{1 - \delta} v. \]  

(30)

A necessary condition for the principal to stick to the contract for all \( \theta \) is that the payoff in (27) is no less than that in (30) for all \( \theta \). With the use of (26), that condition can be written

\[ \frac{\delta}{1 - \delta} \left[ E_{\theta'} v(\theta') - v \right] \geq b(e(\theta), \theta), \text{ for all } \theta \in [\theta, \bar{\theta}]. \]  

(31)

As both (29) and (31) are necessary, so is their sum, which gives the combined necessary condition

\[ \frac{\delta}{1 - \delta} \left[ E_{\theta'} u(\theta') + E_{\theta'} v(\theta') - (u + v) \right] \geq c(e(\theta)), \text{ for all } \theta \in [\theta, \bar{\theta}]. \]  

(32)

But, by definition, \( S(K, r) = E_{\theta'} u(\theta') + E_{\theta'} v(\theta') \) and \( s(K, r) = u + v \). So that (32) is necessary implies that (4) is necessary.

**Sufficiency.** For (4) and hence (32) satisfied, there certainly exists a \( b(e(\theta), \theta) \in [0, c(e(\theta))] \) for each \( \theta \) such that (29) and (31) are both satisfied. For such \( b(e(\theta), \theta) \), the individual rationality conditions

\[ E_{\theta} u(\theta) - u \geq 0 \]  

(33)

\[ E_{\theta} v(\theta) - v \geq 0 \]  

(34)

are also satisfied. To establish that sticking to the contract is a best response for both parties, it remains only to show that the payoffs following deviation specified in (28) and (30), \( \delta u / (1 - \delta) \) and \( \delta v / (1 - \delta) \), are equilibrium payoffs. That is immediate if the strategies for the two parties specify that play following deviation by either is for both to end the relationship because, if the strategy for either is to end the relationship it is a best response of the other also to end the relationship given \( u, v \geq 0 \).
Proof of Corollary 1. If first-best effort \( e^\ast(K,\theta) \) is unattainable for \( \theta' \) for given \( S(K,\sigma) \), it must be because (4) is a binding constraint for \( \theta' \). The left-hand side of (4) is independent of \( \theta \). Moreover, \( e^\ast(K,\theta) \) is increasing in \( \theta \) and \( c(e) \) is increasing in \( e \). Thus the right-hand side of (4) is increasing in \( \theta \). So if (4) is not satisfied for \( \theta' \), it is not satisfied for \( \theta \in (\theta',\bar{\theta}] \). ■

Proof of Proposition 2. By Proposition 1, (4) is a necessary and sufficient condition for an effort schedule to be implementable. Thus an optimal effort schedule for given \( K \) is a solution to the problem

\[
\max_{e(K,\sigma) \in [0,\bar{e}]} \frac{1}{1-\delta} \int_{\theta}^{\bar{\theta}} \left[ s(eK,\bar{\theta}), K, \bar{\theta} \right] - s(K,\sigma) \right] dF \left( \bar{\theta}, \sigma \right) \quad \text{subject to}
\]

\[
\frac{\delta}{1-\delta} \int_{\theta}^{\bar{\theta}} \left[ s(eK,\bar{\theta}), K, \bar{\theta} \right] - s(K,\sigma) \right] dF \left( \bar{\theta}, \sigma \right) \geq c \left( e(K,\theta) \right), \quad \text{for all } \theta \in [\theta,\bar{\theta}].
\]

By Assumptions 1 and 2, optimal effort for all \( \theta \) is interior to \([0,\bar{e}]\). There are three cases to consider: (1) the constraint is not binding for any \( \theta \in [\theta,\bar{\theta}] \); (2) the constraint is binding for all \( \theta \in [\theta,\bar{\theta}] \); and (3) the constraint is not binding for some \( \theta \in [\theta,\bar{\theta}] \) but is binding for other \( \theta \in [\theta,\bar{\theta}] \). For any \( \bar{\theta} \in [\theta,\bar{\theta}] \) for which the constraint is not binding, it is immediate that it is optimal to set \( e(\bar{\theta}, K) = e^\ast(\bar{\theta}, K) \), the first-best effort for which \( s_1(e^\ast(K,\bar{\theta}), K, \bar{\theta}) = 0 \), because that both maximizes the objective function for \( \bar{\theta} \) and relaxes the constraint most for \( \theta \neq \bar{\theta} \). So, for case 1, in which the constraint does not bind for any \( \theta \) (for which a necessary and sufficient condition is that it does not bind for \( \bar{\theta} \) because \( e^\ast(\theta, K) \) is increasing in \( \theta \)), optimal effort is first best for all \( \theta \), as in Part 1 of the proposition. Next note that the integral in the constraint is independent of \( \theta \). So, for case 2, in which the constraint binds for all \( \theta \) (for which a necessary and, by Corollary 1, sufficient condition is that it binds for \( \theta \)), \( c(e(K,\theta)) \) and hence \( e(K,\theta) \) must be the same for all \( \theta \). In that case, optimal effort is the highest effort independent of \( \theta \) that satisfies the constraint, so \( e(K,\theta) = e(\bar{\theta}, K) \) given by (5), as in Part 2 of the proposition. The remaining case is case 3. By the argument just given, for the set of \( \theta \) for which the constraint does not bind, optimal effort must be first best and, for the set of \( \theta \) for which the constraint binds, \( e(K,\theta) \) must be independent of \( \theta \). But, from Corollary 1, if the latter set contains \( \theta' \), it also contains all \( \theta > \theta' \). That implies there is a cutoff type \( \tilde{\theta}(K,\sigma) \) such that, for all \( \theta \leq \tilde{\theta}(K,\sigma) \), effort is at the first-best level and, for all \( \theta > \tilde{\theta}(K,\sigma) \), effort is constant at \( e^\ast(\tilde{\theta}(K,\sigma), K) \). That gives the form of the optimal effort schedule in Part 3 of the proposition. ■
Proof of Lemma 1. Integration by parts gives

\[
\int_{\bar{\theta}}^{\theta} g(\theta) \, dF(\theta, \sigma_L) - \int_{\theta}^{\bar{\theta}} g(\theta) \, dF(\theta, \sigma_H) \\
= [g(\theta) F(\theta, \sigma_L)]_{\bar{\theta}}^{\theta} - \int_{\theta}^{\bar{\theta}} g'(\theta) \, F(\theta, \sigma_L) \, d\theta \\
- \left[ g(\theta) F(\theta, \sigma_H) \right]_{\bar{\theta}}^{\theta} - \int_{\theta}^{\bar{\theta}} g'(\theta) \, F(\theta, \sigma_H) \, d\theta \\
= - \int_{\theta}^{\bar{\theta}} [F(\theta, \sigma_L) - F(\theta, \sigma_H)] \, g'(\theta) \, d\theta,
\]

the last line following because the statement of the lemma specifies that \(g(\bar{\theta}) = 0\) if \(\bar{\theta} > \theta\) and both \(F(\bar{\theta}, \sigma_H)\) and \(F(\bar{\theta}, \sigma_L)\) equal 1. Integration of that last line by parts gives

\[
\int_{\theta}^{\bar{\theta}} g(\theta) \, dF(\theta, \sigma_L) - \int_{\theta}^{\bar{\theta}} g(\theta) \, dF(\theta, \sigma_H) \\
= - \left[ g'(\theta) \int_{\theta}^{\bar{\theta}} [F(x, \sigma_L) - F(x, \sigma_H)] \, dx \right]_{\bar{\theta}}^{\theta} \\
+ \int_{\theta}^{\bar{\theta}} \left[ \int_{\theta}^{\bar{\theta}} [F(x, \sigma_L) - F(x, \sigma_H)] \, dx \right] g''(\theta) \, d\theta. \tag{35}
\]

Now,

\[
\int_{\theta}^{\bar{\theta}} [F(x, \sigma_L) - F(x, \sigma_H)] \, dx = \int_{\theta}^{\bar{\theta}} [F(x, \sigma_L) - F(x, \sigma_H)] \, dx - \int_{\theta}^{\bar{\theta}} [F(x, \sigma_L) - F(x, \sigma_H)] \, dx
\]

and, with \(F(\theta, \sigma_L)\) and \(F(\theta, \sigma_H)\) having the same mean,

\[
\int_{\theta}^{\bar{\theta}} [F(x, \sigma_L) - F(x, \sigma_H)] \, dx = 0,
\]

see Laffont (1989). Use of these in (35) gives that

\[
\int_{\theta}^{\bar{\theta}} g(\theta) \, dF(\theta, \sigma_L) - \int_{\theta}^{\bar{\theta}} g(\theta) \, dF(\theta, \sigma_H) \\
= - g'(\bar{\theta}) \int_{\theta}^{\bar{\theta}} [F(x, \sigma_L) - F(x, \sigma_H)] \, dx \\
- \int_{\theta}^{\bar{\theta}} \left[ \int_{\theta}^{\bar{\theta}} [F(x, \sigma_L) - F(x, \sigma_H)] \, dx \right] g''(\theta) \, d\theta \geq 0,
\]

25
the sign following from \( g'(\theta) < 0, g''(\theta) \leq 0 \) and \( \int_\theta^\beta [F(x, \sigma_L) - F(x, \sigma_H)] \, dx \geq 0 \) for all \( \theta \in [\theta, \beta] \), with strict inequality for a set of values of \( x \in [\theta, \beta] \) with positive probability, as specified in Definition 1. If \( g''(\theta) < 0 \), the inequality is strict. ■

**Proof of Proposition 3.** When risk is systemic, the constraint (4) is by definition unaffected by changes in \( \sigma \) for given \( K \), so the optimal effort schedule in Proposition 2 is unaffected. Thus \( \hat{\theta}(K, \sigma_H) = \hat{\theta}(K, \sigma_L) \). For idiosyncratic risk, consider the two parts of the proposition separately.

**Part 1.** For optimal full pooling, the binding constraint is (5). Thus, by the standard result, for \( s(\bar{e}(K, \sigma_L), K, \theta) \) strictly concave in \( \theta \) for \( \theta \in [\hat{\theta}, \beta] \), \( F(\theta, \sigma_L) \) stochastically dominating \( F(\theta, \sigma_H) \) in the second-order sense for \( \bar{\theta} = \hat{\theta} \) implies

\[
\int_\hat{\theta}^\beta s(\bar{e}(K, \sigma_L), K, \theta) \, dF(\theta, \sigma_L) > \int_\hat{\theta}^\beta s(\bar{e}(K, \sigma_L), K, \theta) dF(\theta, \sigma_H).
\]

Thus the left-hand side of (5) is strictly negative for \( \sigma_H \) if \( \bar{e}(K, \sigma_H) = \bar{e}(K, \sigma_L) \). To satisfy that constraint requires \( \bar{e}(K, \sigma_H) < \bar{e}(K, \sigma_L) \) which, being further below first-best effort for every \( \theta \), implies a larger difference for \( \sigma_H \) than for \( \sigma_L \) between the joint payoff when effort is contractible and when it is not. It also implies a lower joint payoff for \( \sigma_H \) than for \( \sigma_L \) under an optimal relational contract for given \( K \).

**Part 2.** From (8), \( h(e^*(K, \theta), K, \theta) = 0 \) and, from (9), \( h(e, K, \theta) \) is decreasing in \( \theta \). Thus, by Lemma 1 for \( h(e^*(K, \hat{\theta}(K, \sigma_H)), K, \theta) \) strictly concave in \( \theta \) for \( \theta \in [\hat{\theta}(K, \sigma_H), \beta] \), \( F(\theta, \sigma_L) \) stochastically dominating \( F(\theta, \sigma_H) \) in the second-order sense for \( \bar{\theta} = \hat{\theta}(K, \sigma_H) \) implies

\[
\int_{\hat{\theta}(K, \sigma_H)}^\beta h(e^*(K, \hat{\theta}(K, \sigma_H)), K, \theta) \, dF(\theta, \sigma_L) > \int_{\hat{\theta}(K, \sigma_H)}^\beta h(e^*(K, \hat{\theta}(K, \sigma_H)), K, \theta) dF(\theta, \sigma_H).
\]

Since \( h \) is the negative of the integrand in the second integral in (7), the left-hand side of (7) is thus lower for \( F(\theta, \sigma_H) \) than for \( F(\theta, \sigma_L) \) for \( \hat{\theta}(K, \sigma) = \hat{\theta}(K, \sigma_H) \) when (10) holds. So the left-hand side of constraint (6) would be positive for \( F(\theta, \sigma_L) \) if \( \hat{\theta}(K, \sigma_L) = \hat{\theta}(K, \sigma_H) \). But, the left-hand side of constraint (6) is continuous in \( \bar{\theta} \) and negative for \( \bar{\theta} = \beta \) because otherwise optimal effort would not be in Case 2 of Proposition 2. Because, by definition, \( \hat{\theta}(K, \sigma_L) \) is the highest value of \( \hat{\theta} \) that satisfies (6) for \( F(\theta, \sigma_L) \), the left-hand side of (6) must be lower for all \( \bar{\theta} > \hat{\theta}(K, \sigma_L) \) than for \( \bar{\theta} = \hat{\theta}(K, \sigma_L) \) — otherwise, by continuity, there must be a \( \bar{\theta} > \hat{\theta}(K, \sigma_L) \) for which equality holds in (6), which is a contradiction. So it cannot be that \( \hat{\theta}(K, \sigma_L) \leq \hat{\theta}(K, \sigma_H) \). It must, therefore, be that \( \hat{\theta}(K, \sigma_H) < \hat{\theta}(K, \sigma_L) \). From (7), this implies a larger difference for \( \sigma_H \) than for \( \sigma_L \) between the joint payoff when effort is contractible and when it is not. With (10), it also implies a lower joint payoff for \( \sigma_H \) than for \( \sigma_L \) under an optimal relational contract for given \( K \). ■
Proof of Lemma 2. From (8),
\[
h_3(e, K, \theta) = \frac{d}{d\theta} [s(e, K, \theta) - s(e^*(K, \theta), K, \theta)] = s_3(e, K, \theta) - s_1(e^*(K, \theta), K, \theta) e_2^*(K, \theta) - s_3(e^*(K, \theta), K, \theta).
\]

So
\[
h_{33}(e, K, \theta) = \frac{d}{d\theta} [s_3(e, K, \theta) - s_1(e^*(K, \theta), K, \theta) e_2^*(K, \theta) - s_3(e^*(K, \theta), K, \theta)]
\]
\[
= s_{33}(e, K, \theta) - s_{11}(e^*(K, \theta), K, \theta) e_2^*(K, \theta) - s_{13}(e^*(K, \theta), K, \theta) e_2^*(K, \theta)
\]
\[
= s_{33}(e, K, \theta) - s_{11}(e^*(K, \theta), K, \theta) e_2^*(K, \theta) - s_{13}(e^*(K, \theta), K, \theta) e_2^*(K, \theta) - s_{33}(e^*(K, \theta), K, \theta)
\]
\[
= s_{33}(e, K, \theta) - s_{33}(e^*(K, \theta), K, \theta) - s_{11}(e^*(K, \theta), K, \theta) e_2^*(K, \theta) - 2s_{13}(e^*(K, \theta), K, \theta) e_2^*(K, \theta),
\]

the last equality making use of the definition of $e^*(K, \theta)$ in (1) which implies
\[
s_1(e^*(K, \theta), K, \theta) = 0, \quad \text{for all } \theta \in [\theta, \bar{\theta}].
\]

So $e_2^*(K, \theta)$ satisfies
\[
e_2^*(K, \theta) = -\frac{s_{13}(e^*(K, \theta), K, \theta)}{s_{11}(e^*(K, \theta), K, \theta)}.
\]

Thus
\[
h_{33}(e, K, \theta) = s_{33}(e, K, \theta) - s_{33}(e^*(K, \theta), K, \theta)
\]
\[
- s_{11}(e^*(K, \theta), K, \theta) \left( -\frac{s_{13}(e^*(K, \theta), K, \theta)}{s_{11}(e^*(K, \theta), K, \theta)} \right)^2
\]
\[
- 2s_{13}(e^*(K, \theta), K, \theta) \left( -\frac{s_{13}(e^*(K, \theta), K, \theta)}{s_{11}(e^*(K, \theta), K, \theta)} \right)
\]
\[
= s_{33}(e, K, \theta) - s_{33}(e^*(K, \theta), K, \theta)
\]
\[
- \frac{s_{13}(e^*(K, \theta), K, \theta)^2}{s_{11}(e^*(K, \theta), K, \theta)} + 2\frac{s_{13}(e^*(K, \theta), K, \theta)^2}{s_{11}(e^*(K, \theta), K, \theta)}
\]
\[
= s_{33}(e, K, \theta) - s_{33}(e^*(K, \theta), K, \theta) + \frac{s_{13}(e^*(K, \theta), K, \theta)^2}{s_{11}(e^*(K, \theta), K, \theta)}.
\]

The last term in this is negative because Assumption 1 requires $s(e, K, \theta)$ strictly concave in $e$ and $s_{13}(e, K, \theta) \neq 0$. Moreover, the first two terms go to zero as $e$ goes to $e^*(K, \theta)$ and then $h_{33}(e, K, \theta) < 0$. ■
Proof of Proposition 4. Start with the case of partial pooling of efforts (Case 3 of Proposition 2) and note that the integral terms in (6) are just $S(K, \sigma) - \bar{z}(K, \sigma)$. By the definition of general capital in Definition 3, this is independent of $K$, so total differentiation of (6) gives

$$-c' \left( e^* \left( K, \tilde{\theta} \right) \right) \left[ e_1^* \left( K, \tilde{\theta} \right) dK + e_2^* \left( K, \tilde{\theta} \right) d\tilde{\theta} \right] = 0.$$ 

With $\tilde{\theta} (K, \sigma)$ corresponding to some $\tilde{\theta}$ and $c' > 0$, this implies

$$\frac{d\tilde{\theta} (K, \sigma)}{dK} = -\frac{e_1^* (K, \tilde{\theta} (K, \sigma))}{e_2^* (K, \tilde{\theta} (K, \sigma))} < 0. \quad (36)$$

Thus, for general capital, the derivative of the maximand in (11) is, in view of (1),

$$\frac{1}{1 - \delta} \left\{ \int_{\tilde{\theta}} \bar{e}_2 \left( e^* (K, \theta), K, \theta \right) dF (\theta, \sigma) \right.$$ 

$$+ \int_{\tilde{\theta} (K, \sigma)} \left[ -s_2 (e^* (K, \theta), K, \theta) + s_2 (e^* (K, \tilde{\theta} (K, \sigma)), K, \theta) \right.$$ 

$$- s_1 (e^* (K, \tilde{\theta} (K, \sigma)), K, \theta) e_1^* (K, \tilde{\theta} (K, \sigma)) \right] dF (\theta, \sigma) \left\} - C' (K). \quad (37)$$

This must equal zero for $K = \hat{K}^G (\sigma)$. For $K = K^* (\sigma)$, the first integral term equals $C' (K^* (\sigma))$ by definition. Thus to have $\hat{K}^G (\sigma) = K^* (\sigma)$, the second integral would have to be zero. But $s_1, e_1^* > 0$ for effort below first best and $s_{12} = y_{12} > 0$ by Assumption 1. So the second integral is negative because $e^* (K, \theta) > e^* (K, \tilde{\theta} (K, \sigma))$ for $\theta > \tilde{\theta} (K, \sigma)$. Since $s(e, K, \theta) - C (K)$ is strictly concave in $K$, this implies $\hat{K}^G (\sigma) < K^* (\sigma)$.

For the case of full pooling of efforts, the maximand is that in (11) but with $\tilde{\theta} (K, \sigma) = \theta$ and $e^* (K, \tilde{\theta} (K, \sigma))$ replaced by $\bar{e} (K, \sigma)$ defined in (5). The definition of general capital in Definition 3 implies that $\bar{e} (K, \sigma)$ defined in (5) is independent of $K$, so the derivative of the maximand in (11) can be written

$$\frac{1}{1 - \delta} \left\{ \int_{\tilde{\theta}} \bar{e}_2 \left( e^* (K, \theta), K, \theta \right) dF (\theta, \sigma) \right.$$ 

$$+ \int_{\tilde{\theta}} \left[ -s_2 (e^* (K, \theta), K, \theta) + s_2 (\bar{e} (K, \sigma), K, \theta) \right] dF (\theta, \sigma) \left\} - C' (K). \quad (38)$$

The same argument can be applied to this as to (37). ■

Proof of Proposition 5. Start with the case of partial pooling of efforts (Case 3 of Proposition 2). For general capital, the derivative of the maximand in (11) is given by
(37). With the definition of $h$ in (8), that can be written

$$
\frac{1}{1 - \delta} \int_{\theta}^{\bar{\theta}} s_2 (e^* (K, \theta), K, \theta) dF (\theta, \sigma) + \frac{1}{1 - \delta} \int_{\hat{\theta}(K, \sigma)}^{\bar{\theta}} h_2 (e^* (K, \hat{\theta} (K, \sigma)), K, \theta) dF (\theta, \sigma) - C' (K).
$$

This must equal zero for $K = \hat{K}^G (\sigma)$. For $K = K^* (\sigma)$, the first integral equals $C' (K^* (\sigma))$ by definition. Thus to have $\hat{K} (\sigma) = K^* (\sigma)$, the second integral term would have to be zero. But by Lemma 1, with $h_2 (e^* (K, \hat{\theta} (K, \sigma)), K, \theta)$ decreasing and strictly concave in $\theta$ for $\theta \in \left[ \hat{\theta} (K, \sigma), \bar{\theta} \right]$,

$$
\int_{\hat{\theta}(K, \sigma)}^{\bar{\theta}} h_2 (e^* (K, \hat{\theta} (K, \sigma)), K, \theta) dF (\theta, \sigma_L) > \int_{\hat{\theta}(K, \sigma_H)}^{\bar{\theta}} h_2 (e^* (K, \hat{\theta} (K, \sigma)), K, \theta) dF (\theta, \sigma_H)
$$

for given $\hat{\theta} (K, \sigma)$. But, from Proposition 3, $\hat{\theta} (K, \sigma_L) \leq \hat{\theta} (K, \sigma_H)$ for given $K$ for both systemic and idiosyncratic risk. Moreover, $s_{12} > 0$ implies that $h_2$ is negative for $\theta > \hat{\theta} (K, \sigma)$. Furthermore, $e^* (K, \theta)$ is increasing in $\theta$. Thus also

$$
\int_{\hat{\theta}(K, \sigma_L)}^{\bar{\theta}} h_2 (e^* (K, \hat{\theta} (K, \sigma_L)), K, \theta) dF (\theta, \sigma_L) > \int_{\hat{\theta}(K, \sigma_H)}^{\bar{\theta}} h_2 (e^* (K, \hat{\theta} (K, \sigma_H)), K, \theta) dF (\theta, \sigma_H).
$$

Since $s(e, K, \theta) - C (K)$ is strictly concave in $K$, this implies $K^* (\sigma_H) - \hat{K}^G (\sigma_H) > K^* (\sigma_L) - \hat{K}^G (\sigma_L)$. For the case of full pooling of efforts, the result follows by replacing $\hat{\theta} (K, \sigma)$ in the above by $\hat{\theta}$. ■

**Proof of Lemma 3.** From the definition of $h$ in (8),

$$
h_2 (e, K, \theta) = s_2 (e, K, \theta) - s_2 (e^* (K, \theta), K, \theta) - s_1 (e^* (K, \theta), K, \theta) e^* (K, \theta),
$$

so

$$
h_{23} (e, K, \theta) = s_{23} (e, K, \theta) - s_{23} (e^* (K, \theta), K, \theta) - s_{12} (e^* (K, \theta), K, \theta) e^* (K, \theta) - s_{11} (e^* (K, \theta), K, \theta) e^* (K, \theta) e^* (K, \theta) e_{12} (K, \theta) - s_{13} (e^* (K, \theta), K, \theta) e^* (K, \theta).
$$

Noting that, by the definition of $e^* (K, \theta)$ in (5),

$$
s_1 (e^* (K, \theta), K, \theta) = 0, \quad \text{for all } \theta \in \left[ \hat{\theta}, \bar{\theta} \right],
$$

29
we can write

\[ h_{23} (e, K, \theta) = s_{23} (e, K, \theta) - s_{23} (e^* (K, \theta), K, \theta) - s_{13} (e^* (K, \theta), K, \theta) e_1^* (K, \theta) \]

\[ - \left[ s_{12} (e^* (K, \theta), K, \theta) + s_{11} (e^* (K, \theta), K, \theta) e_1^* (K, \theta) \right] e_2^* (K, \theta). \]

Moreover, again by the definition of \( e^* (K, \theta) \) in (5),

\[ e_1^* (K, \theta) = \frac{s_{12} (e^* (K, \theta), K, \theta)}{s_{11} (e^* (K, \theta), K, \theta)'}, \]

from which the square bracket in the expression for \( h_{23} \) becomes zero so

\[ h_{23} (e, K, \theta) = s_{23} (e, K, \theta) - s_{23} (e^* (K, \theta), K, \theta) - s_{13} (e^* (K, \theta), K, \theta) e_1^* (K, \theta), \]

which, for \( e < e^* (K, \theta) \), is negative for \( s_{123} \geq 0 \). But \( s_{123} = y_{123} \).

**Proof of Proposition 6.** Consider first the case in which the optimal effort schedule involves partial pooling of \( \theta \) (that is, Part 3 of Proposition 2 applies) for both distribution \( F (\theta, \sigma_{H}) \) and distribution \( F (\theta, \sigma_{L}) \). From Proposition 2, for optimal effort partial pooling, \( \hat{\theta} (K, \sigma) \) is the highest \( \hat{\theta} \) satisfying (6) and thus, for the specification in the proposition with \( \tilde{\theta} (K, \sigma) \) independent of \( K \), satisfies

\[ \frac{\delta}{1 - \delta} \left\{ \int_{\theta} \tilde{\theta} s (e^* (K, \theta), K, \theta) dF (\theta, \sigma) - \tilde{s} (\sigma) \right\} \]

\[ - \int_{\hat{\theta} (K, \sigma)} \left\{ s (e^* (K, \theta), K, \theta) - s (e^* (K, \hat{\theta} (K, \sigma)), K, \theta) \right\} dF (\theta, \sigma) \]

\[ - c (e^* (K, \hat{\theta} (K, \sigma))) = 0. \]

(40)

Differentiation of this with respect to \( K \) gives, as long as optimal effort remains partial pooling,

\[ \frac{\delta}{1 - \delta} \int_{\theta} s_2 (e^* (K, \theta), K, \theta) dF (\theta, \sigma) \]

\[ - \frac{\delta}{1 - \delta} \int_{\hat{\theta} (K, \sigma)} \left\{ s_2 (e^* (K, \theta), K, \theta) - s_2 (e^* (K, \hat{\theta} (K, \sigma)), K, \theta) \right\} dF (\theta, \sigma) \]

\[ - s_1 (e^* (K, \hat{\theta} (K, \sigma)), K, \theta) \left[ e_1^* (K, \hat{\theta} (K, \sigma)) + e_2^* (K, \hat{\theta} (K, \sigma)) \hat{\theta}_1 (K, \sigma) \right] dF (\theta, \sigma) \]

\[ - c' (e^* (K, \hat{\theta} (K, \sigma))) \left[ e_1^* (K, \hat{\theta} (K, \sigma)) + e_2^* (K, \hat{\theta} (K, \sigma)) \hat{\theta}_1 (K, \sigma) \right] = 0. \]

(41)
because $s_1(e^*(K, \theta), K, \theta) = 0$ by definition of $e^*(K, \theta)$ and the integrand of the second integral equals zero for $\theta = \hat{\theta}(K, \sigma)$. It must also be that

$$
\frac{\delta}{1 - \delta} \int_{\hat{\theta}(K, \sigma)}^{\overline{\theta}} s_1 \left( e^* (K, \hat{\theta}(K, \sigma)) , K, \theta (\overline{\theta} (K, \sigma) ) \right) \, dF (\theta, \sigma) \\
- c' \left( e^* (K, \hat{\theta}(K, \sigma)) \right) e_2^* (K, \hat{\theta}(K, \sigma) ) < 0 \quad (42)
$$

because otherwise it would have been possible to increase $\hat{\theta}(K, \sigma)$ while still satisfying (40). Since $e_2^* (K, \theta) > 0$, this implies

$$
\frac{\delta}{1 - \delta} \int_{\hat{\theta}(K, \sigma)}^{\overline{\theta}} s_1 \left( e^* (K, \hat{\theta}(K, \sigma)) , K, \theta \right) \, dF (\theta, \sigma) - c' \left( e^* (K, \hat{\theta}(K, \sigma)) \right) < 0. \quad (43)
$$

From (41),

$$
e_1^* (K, \hat{\theta}(K, \sigma)) + e_2^* (K, \hat{\theta}(K, \sigma)) \hat{\theta}_1 (K, \sigma) \\
= \frac{\delta}{1 - \delta} \left\{ \int_{\theta}^{\overline{\theta}} s_2 (e^* (K, \theta), K, \theta) \, dF (\theta, \sigma) \\
- \int_{\theta(K, \sigma)}^{\overline{\theta}} \left[ s_2 (e^* (K, \theta), K, \theta) - s_2 \left( e^* (K, \hat{\theta}(K, \sigma)), K, \theta \right) \right] \, dF (\theta, \sigma) \right\} \\
\div \left\{ c' \left( e^* (K, \hat{\theta}(K, \sigma)) \right) - \frac{\delta}{1 - \delta} \int_{\hat{\theta}(K, \sigma)}^{\overline{\theta}} s_1 \left( e^* (K, \hat{\theta}(K, \sigma)), K, \theta \right) \, dF (\theta, \sigma) \right\}. \quad (44)
$$

The maximand in (11) can be written

$$
\max_{K \in [0, \overline{K}]} \frac{1}{1 - \delta} \left\{ \int_{\theta}^{\overline{\theta}} s (e^* (K, \theta), K, \theta) \, dF (\theta, \sigma) \\
- \int_{\hat{\theta}(K, \sigma)}^{\overline{\theta}} \left[ s \left( e^* (K, \theta), K, \theta \right) - s \left( e^* (K, \hat{\theta}(K, \sigma)), K, \theta \right) \right] \, dF (\theta, \sigma) \right\} - C (K) \\
\text{subject to } \hat{\theta} (K, \sigma) \text{ the highest value of } \hat{\theta} \text{ satisfying (6).} \quad (45)
$$

The first derivative with respect to $K$ of the corresponding Lagrangean is thus, because
\[ s_1(e^*(K, \theta), K, \theta) = 0 \text{ by definition of } e^*(K, \theta), \]

\[
\frac{1 + \lambda \delta}{1 - \delta} \int_{\theta}^{\bar{\theta}} s_2(e^*(K, \theta), K, \theta) dF(\theta, \sigma) \\
- \left. \frac{1 + \lambda \delta}{1 - \delta} \int_{\theta}^{\bar{\theta}} \{ s_2(e^*(K, \theta), K, \theta) - s_2(e^*(K, \hat{\theta}(K, \sigma)), K, \theta) \\
- s_1(e^*(K, \hat{\theta}(K, \sigma)), K, \theta) [\tilde{e}_1^*(K, \hat{\theta}(K, \sigma)) + \tilde{e}_2^*(K, \hat{\theta}(K, \sigma)) \hat{\theta}_1(K, \sigma)] \} dF(\theta, \sigma) \right.
\]

\[ - \lambda c'(e^*(K, \hat{\theta}(K, \sigma))) [\tilde{e}_1^*(K, \hat{\theta}(K, \sigma)) + \tilde{e}_2^*(K, \hat{\theta}(K, \sigma)) \hat{\theta}_1(K, \sigma)] - C'(K), \quad (46) \]

where \( \lambda \) is a non-negative Lagrange multiplier attached to the constraint (6). In view of (41), the terms in this multiplied by \( \lambda \) are equal to zero, so this derivative can be written

\[
\frac{1}{1 - \delta} \int_{\theta}^{\bar{\theta}} s_2(e^*(K, \theta), K, \theta) dF(\theta, \sigma) \\
- \left. \frac{1}{1 - \delta} \int_{\theta}^{\bar{\theta}} \{ s_2(e^*(K, \theta), K, \theta) - s_2(e^*(K, \hat{\theta}(K, \sigma)), K, \theta) \\
- s_1(e^*(K, \hat{\theta}(K, \sigma)), K, \theta) [\tilde{e}_1^*(K, \hat{\theta}(K, \sigma)) + \tilde{e}_2^*(K, \hat{\theta}(K, \sigma)) \hat{\theta}_1(K, \sigma)] \} dF(\theta, \sigma) \right.
\]

\[ - C'(K). \quad (47) \]

Use of (44) in this gives

\[
\frac{1}{1 - \delta} \left\{ \int_{\theta}^{\bar{\theta}} s_2(e^*(K, \theta), K, \theta) dF(\theta, \sigma) \\
- \int_{\theta}^{\bar{\theta}} [s_2(e^*(K, \theta), K, \theta) - s_2(e^*(K, \hat{\theta}(K, \sigma)), K, \theta)] dF(\theta, \sigma) \right\} \left\{ 1 \\
+ \frac{\delta}{1 - \delta} \int_{\theta}^{\bar{\theta}} s_1(e^*(K, \hat{\theta}(K, \sigma)), K, \theta) dF(\theta, \sigma) \right\} \\
\div \left[ c'(e^*(K, \hat{\theta}(K, \sigma))) - \frac{\delta}{1 - \delta} \int_{\theta}^{\bar{\theta}} s_1(e^*(K, \hat{\theta}(K, \sigma)), K, \theta) dF(\theta, \sigma) \right] - C'(K). \quad (48) \]

We know from above that the square bracket on the bottom line of this is positive. The
whole expression can also be written

\[
\frac{1}{1 - \delta} \left\{ \int_{\theta}^{\theta} \left[ s_{2}(e^{*}(K, \theta), K, \theta) dF(\theta, \sigma) - \int_{\theta(K, \sigma)}^{\theta} \left[ s_{2}(e^{*}(K, \theta), K, \theta) - s_{2}(e^{*}(K, \hat{\theta}(K, \sigma)), K, \theta) \right] dF(\theta, \sigma) \right] \right\} \left\{ 1 + \frac{1 - \delta}{\delta} \frac{c'(e^{*}(K, \hat{\theta}(K, \sigma)))}{\int_{\theta(K, \sigma)}^{\theta} s_{1}(e^{*}(K, \hat{\theta}(K, \sigma)), K, \theta) dF(\theta, \sigma)} - 1 \right\} \right\} \right\} \right\} \right\} \right\} \right\} - C'(K). \quad (49)
\]

For \( \hat{K}^{S}(\sigma) \), (49) must equal zero. For \( \hat{K}^{G}(\sigma) \), (49) the first brace must equal zero. The square bracket on the bottom line of (49) is positive, so the second brace in it is greater than 1. Moreover, because specific capital relaxes the constraint (4), \( \hat{\theta}(K, \sigma) \) is higher for given \( (K, \sigma) \) when capital is specific than when it is general. Furthermore, \( s_{12} > 0 \) implies that the second integral in (49) is positive given \( e^{*}(K, \hat{\theta}(K, \sigma)) < e^{*}(K, \theta) \). Thus, the expression in (49) is larger for given \( K \) when capital is specific than when it is general. Because \( s \) is strictly concave in \( K \) and \( C \) convex, it must therefore be that \( \hat{K}^{S}(\sigma) > \hat{K}^{G}(\sigma) \). Moreover, for \( K^{*}(\sigma) \) for contractible effort, the first integral term in (49) must equal zero. Thus, for \( s_{12} = y_{12} \) sufficiently small, the expression in (49) is larger for given \( K \) when effort is non-contractible and capital is specific than it is when effort is contractible. Thus, under that condition, \( \hat{K}^{S}(\sigma) > K^{*}(\sigma) \).

Now consider the case in which the optimal effort schedule involves full pooling of \( \theta \) (that is, Part 2 of Proposition 2 applies) for both distribution \( F(\theta, \sigma_{U}) \) and distribution \( F(\theta, \sigma_{L}) \). In this case, by Proposition 2, optimal effort is \( e(K, \theta) = \tilde{e}(K, \sigma) \) given by (5) and thus, for the specification in the proposition with \( \tilde{z}(K, \sigma) \) independent of \( K \), satisfies

\[
\frac{\delta}{1 - \delta} \int_{\theta}^{\theta} \left[ s(\tilde{e}(K, \sigma), K, \theta) - \tilde{z}(\sigma) \right] dF(\theta, \sigma) - c(\tilde{e}(K, \sigma)) = 0. \quad (50)
\]

Differentiation of this with respect to \( K \) gives, as long as optimal effort remains full pooling,

\[
\frac{\delta}{1 - \delta} \int_{\theta}^{\theta} \left[ s_{1}(\tilde{e}(K, \sigma), K, \theta) \tilde{e}_{1}(K, \sigma) + s_{2}(\tilde{e}(K, \sigma), K, \theta) \right] dF(\theta, \sigma) - c'(\tilde{e}(K, \sigma)) \tilde{e}_{1}(K, \sigma) = 0. \quad (51)
\]

33
From this,

\[
\tilde{e}_1(K, \sigma) = \frac{\delta}{1 - \delta} \int_0^\theta s_2(\tilde{e}(K, \sigma), K, \theta) \, dF(\theta, \sigma)
\]

\[
= \frac{c'(\tilde{e}(K, \sigma)) - \frac{\delta}{1 - \delta} \int_0^\theta s_1(\tilde{e}(K, \sigma), K, \theta) \, dF(\theta, \sigma)}{c'(\tilde{e}(K, \sigma)) - \delta \int_0^\theta s_1(\tilde{e}(K, \sigma), K, \theta) \, dF(\theta, \sigma)}.
\] (52)

The maximization for this case corresponding to (11) is

\[
\max_{K \in [0, \bar{K}]} \frac{1}{1 - \delta} \int_0^\theta s(\tilde{e}(K, \sigma), K, \theta) \, dF(\theta, \sigma) - C(K) \text{ subject to the constraint (50).} \] (53)

The first derivative with respect to \(K\) of the corresponding Lagrangean is

\[
\frac{1 + \lambda \delta}{1 - \delta} \int_0^\theta \{s_1(\tilde{e}(K, \sigma), K, \theta) \tilde{e}_1(K, \sigma) + s_2(\tilde{e}(K, \sigma), K, \theta)\} \, dF(\theta, \sigma)
\]

\[
- \lambda c'(\tilde{e}(K, \sigma)) \tilde{e}_1(K, \sigma) - C'(K), \] (54)

where \(\lambda\) is a non-negative Lagrange multiplier attached to the constraint (50). In view of (51), the terms in this multiplied by \(\lambda\) are equal to zero, so this derivative can be written

\[
\frac{1}{1 - \delta} \int_0^\theta \{s_1(\tilde{e}(K, \sigma), K, \theta) \tilde{e}_1(K, \sigma) + s_2(\tilde{e}(K, \sigma), K, \theta)\} \, dF(\theta, \sigma) - C'(K). \] (55)

Use of (52) in this gives

\[
\frac{1}{1 - \delta} \left\{ \int_0^\theta s_2(\tilde{e}(K, \sigma), K, \theta) \, dF(\theta, \sigma) \right\}
\]

\[
\times \left\{ 1 + \frac{\frac{\delta}{1 - \delta} \int_0^\theta s_1(\tilde{e}(K, \sigma), K, \theta) \, dF(\theta, \sigma)}{c'(\tilde{e}(K, \sigma)) - \frac{\delta}{1 - \delta} \int_0^\theta s_1(\tilde{e}(K, \sigma), K, \theta) \, dF(\theta, \sigma)} \right\} - C'(K). \] (56)
The whole expression can also be written

\[
\frac{1}{1 - \delta} \left\{ \int_{\hat{\theta}}^{\bar{\theta}} s_2 (e^* (K, \theta), K, \theta) \, dF (\theta, \sigma) \right. \\
- \left. \frac{1}{1 - \delta} \int_{\hat{\theta}}^{\bar{\theta}} [s_2 (e^* (K, \theta), K, \theta) - s_2 (\bar{e} (K, \sigma), K, \theta)] \, dF (\theta, \sigma) \right\} \\
\times \left\{ 1 + \frac{\delta}{1 - \delta} \int_{\hat{\theta}}^{\bar{\theta}} \frac{s_1 (\bar{e} (K, \sigma), K, \theta) \, dF (\theta, \sigma)}{c' (\bar{e} (K, \sigma))} - \frac{\delta}{1 - \delta} \int_{\hat{\theta}}^{\bar{\theta}} s_1 (\bar{e} (K, \sigma), K, \theta) \, dF (\theta, \sigma) \right\} - C' (K) \tag{57}
\]

or

\[
\frac{1}{1 - \delta} \left\{ \int_{\hat{\theta}}^{\bar{\theta}} s_2 (e^* (K, \theta), K, \theta) \, dF (\theta, \sigma) \right. \\
- \left. \frac{1}{1 - \delta} \int_{\hat{\theta}}^{\bar{\theta}} [s_2 (e^* (K, \theta), K, \theta) - s_2 (\bar{e} (K, \sigma), K, \theta)] \, dF (\theta, \sigma) \right\} \\
\times \left\{ 1 + 1 \div \left[ \frac{1 - \delta}{\delta} \int_{\hat{\theta}}^{\bar{\theta}} \frac{s_1 (\bar{e} (K, \sigma), K, \theta) \, dF (\theta, \sigma)}{c' (\bar{e} (K, \sigma))} - 1 \right] \right\} - C' (K). \tag{58}
\]

This has exactly the same form as (49) except that the limit of integration \( \hat{\theta} (K, \sigma) \) is replaced by \( \theta \). But that difference in the limit of integration makes no difference to the argument that follows, so the same conclusion applies.

**Proof of Proposition 7.** Consider first the case in which the optimal effort schedule exhibits partial pooling (that is, Part 3 of Proposition 2 applies) for both distribution \( F (\theta, \sigma_H) \) and distribution \( F (\theta, \sigma_L) \). Then, under the conditions of the proposition, (49) applies with either systemic or idiosyncratic risk. Moreover, \( s_1 (e^* (K, \hat{\theta} (K, \sigma)), K, \theta) = y_1 (e^* (K, \hat{\theta} (K, \sigma)), K, \theta) - c' (e^* (K, \hat{\theta} (K, \sigma))) \), so

\[
\frac{c' (e^* (K, \hat{\theta} (K, \sigma)))}{\int_{\hat{\theta} (K, \sigma)}^{\bar{\theta} (K, \sigma)} s_1 (e^* (K, \hat{\theta} (K, \sigma)), K, \theta) \, dF (\theta, \sigma)} = \frac{1}{\int_{\hat{\theta} (K, \sigma)}^{\bar{\theta} (K, \sigma)} \left[ \frac{y_1 (e^* (K, \hat{\theta} (K, \sigma)), K, \theta)}{c' (e^* (K, \hat{\theta} (K, \sigma)))} - 1 \right] \, dF (\theta, \sigma)}
\]
and

\[
\frac{y_1(e^*(K, \hat{\theta}(K, \sigma)), K, \theta)}{c'(e^*(K, \hat{\theta}(K, \sigma)))} - 1 = 0, \quad \text{for } \theta = \hat{\theta}(K, \sigma),
\]

\[
\frac{d}{d\theta} \left[ \frac{y_1(e^*(K, \hat{\theta}(K, \sigma)), K, \theta)}{c'(e^*(K, \hat{\theta}(K, \sigma)))} - 1 \right] = \frac{y_{13}(e^*(K, \hat{\theta}(K, \sigma)), K, \theta)}{c'(e^*(K, \hat{\theta}(K, \sigma)))} > 0, \quad \text{for } y_{13} > 0,
\]

\[
\frac{d^2}{d\theta^2} \left[ \frac{y_1(e^*(K, \hat{\theta}(K, \sigma)), K, \theta)}{c'(e^*(K, \hat{\theta}(K, \sigma)))} - 1 \right] = \frac{y_{133}(e^*(K, \hat{\theta}(K, \sigma)), K, \theta)}{c'(e^*(K, \hat{\theta}(K, \sigma)))} \geq 0, \quad \text{for } y_{133} \geq 0.
\]

Thus, the negative of the fraction is concave and strictly decreasing in \(\theta\) so, by Lemma 1,

\[
\int_{\hat{\theta}(K, \sigma_H)}^{\bar{\theta}} \left[ \frac{y_1(e^*(K, \hat{\theta}(K, \sigma_H)), K, \theta)}{c'(e^*(K, \hat{\theta}(K, \sigma_H)))} - 1 \right] dF(\theta, \sigma_H) > \int_{\hat{\theta}(K, \sigma_L)}^{\bar{\theta}} \left[ \frac{y_1(e^*(K, \hat{\theta}(K, \sigma_L)), K, \theta)}{c'(e^*(K, \hat{\theta}(K, \sigma_L)))} - 1 \right] dF(\theta, \sigma_L). \quad (59)
\]

So, the square bracket on the bottom line of (49) is smaller for \(\sigma_H\) than for \(\sigma_L\) and the second brace in (49) is larger for \(\sigma_H\) than for \(\sigma_L\). Thus it must be that, for \(s_{12} = y_{12}\) sufficiently small, \(R^S(\sigma_H) - K^*(\sigma_H) > R^S(\sigma_L) - K^*(\sigma_L)\).

Now consider the case in which the optimal effort schedule involves full pooling (that is, Part 2 of Proposition 2 applies) for both distribution \(F(\theta, \sigma_H)\) and distribution \(F(\theta, \sigma_L)\). Under the conditions of the proposition, (58) applies to this case with either systemic or idiosyncratic risk. But (58) has exactly the same form as (49) except that the limit of integration \(\hat{\theta}(K, \sigma)\) is replaced by \(\bar{\theta}\). The only difference that makes to the argument that follows is that, to get a strict inequality in Lemma 1 when the distributions have the same mean requires \(y_{133} > 0\). With that change in assumption, the same conclusion applies.

**Appendix B  Derivations of equations in Section 7**

**Derivation of Equation (15).** For the specification in (12) and (13)

\[
s(e, K, \theta) = \theta^\gamma K^\alpha e^\delta - ce^n. \quad (60)
\]

Thus, the first-order condition for optimal choice of effort conditional on \(K\) and \(\theta\), \(e^*(K, \theta)\), is

\[
\beta \theta^\gamma K^\alpha e^*(K, \theta)^{\beta - 1} - nce^*(K, \theta)^{n - 1} = 0,
\]

36
which can be written

\[
\frac{\beta}{e^* (K, \theta)} y (e^* (K, \theta), K, \theta) = nce^* (K, \theta)^{n-1}
\]

or

\[
e^* (K, \theta)^n = \frac{\beta}{nc} y (e^* (K, \theta), K, \theta)
\]  

(61)

or

\[
e^* (K, \theta) = \left(\frac{\beta}{nc}\right)^{\frac{1}{n}} y (e^* (K, \theta), K, \theta)^{\frac{1}{n}}.
\]

Used in the production function (12), this gives

\[
y (e^* (K, \theta), K, \theta) = \theta^\gamma K^\alpha \left(\frac{\beta}{nc}\right)^{\frac{\delta}{n}} y (e^* (K, \theta), K, \theta)^{\frac{\delta}{n}},
\]

which can be solved for \(y (e^* (K, \theta), K, \theta)\) to give

\[
y (e^* (K, \theta), K, \theta)^{1-\frac{\delta}{n}} = \theta^\gamma K^\alpha \left(\frac{\beta}{nc}\right)^{\frac{\delta}{n}}
\]

or

\[
y (e^* (K, \theta), K, \theta) = \left[\theta^\gamma K^\alpha \left(\frac{\beta}{nc}\right)^{\frac{\delta}{n}}\right]^{\frac{1}{1-\frac{\delta}{n}}}.
\]

Thus, from (61),

\[
e^* (K, \theta)^n = \frac{\beta}{nc} \left[\theta^\gamma K^\alpha \left(\frac{\beta}{nc}\right)^{\frac{\delta}{n}}\right]^{\frac{1}{1-\frac{\delta}{n}}} = \left(\frac{\beta}{nc} \theta^\gamma K^\alpha\right)^{\frac{1}{1-\frac{\delta}{n}}}.
\]  

(62)

Equation (15) follows directly.

**Derivation of Equations (16) and (17).** Use of (62) in (60) gives

\[
s (e^* (K, \theta), K, \theta) = \theta^\gamma K^\alpha \left(\frac{\beta}{nc} \theta^\gamma K^\alpha\right)^{\frac{1}{1-\frac{\delta}{n}}} - c \left(\frac{\beta}{nc} \theta^\gamma K^\alpha\right)^{\frac{1}{1-\frac{\delta}{n}}}
\]

\[
= \left(\frac{\beta}{nc}\right)^{\frac{1}{1-\frac{\delta}{n}}} \left(\theta^\gamma K^\alpha\right)^{\frac{1}{1-\frac{\delta}{n}}} - c \left(\frac{\beta}{nc}\right)^{\frac{1}{1-\frac{\delta}{n}}} \left(\theta^\gamma K^\alpha\right)^{\frac{1}{1-\frac{\delta}{n}}}.
\]
from which (16) follows. Moreover, from (60) and (62) for given \( \tilde{\theta} \),
\[
\begin{align*}
    s \left( e^* \left( K, \tilde{\theta} \right), K, \theta \right) &= \theta^\gamma K^\alpha e^* \left( K, \tilde{\theta} \right) - ce^* \left( K, \tilde{\theta} \right) \\
    &= \theta^\gamma K^\alpha \left( \frac{\beta}{nc} \tilde{\theta}^\gamma K^\alpha \right)^{\frac{\beta}{n} \frac{1}{1-\beta/n}} - c \left( \frac{\beta}{nc} \tilde{\theta}^\gamma K^\alpha \right)^{\frac{1}{1-\beta/n}} \\
    &= \left( \frac{\beta}{nc} \tilde{\theta}^\gamma K^\alpha \right)^{\frac{1}{1-\beta/n}} \left[ \theta^\gamma K^\alpha \left( \frac{\beta}{nc} \tilde{\theta}^\gamma K^\alpha \right)^{\frac{\beta}{n} - 1} - c \right] \\
    &= \left( \frac{\beta}{nc} \tilde{\theta}^\gamma K^\alpha \right)^{\frac{1}{1-\beta/n}} \left[ \theta^\gamma K^\alpha \left( \frac{\beta}{nc} \tilde{\theta}^\gamma K^\alpha \right)^{-1} - c \right],
\end{align*}
\]
from which (17) follows.

**Derivation of Equations (20) and (21).** With the definition of \( \hat{S} (\sigma) \) in (19) and the specification of the effort cost function in (13), (6) becomes
\[
\hat{S} (\sigma) = ce^* \left( K, \tilde{\theta} \right)^n, \text{ for } \tilde{\theta} = \hat{\theta} \left( K, \sigma \right).
\]
which is the form taken by the constraint on the choice of optimal capital in (11) when capital is general. Use of (64) in the top line of (63) for \( \tilde{\theta} = \hat{\theta} \left( K, \sigma \right) \) gives
\[
\begin{align*}
    s \left( e^* \left( K, \hat{\theta} \left( K, \sigma \right) \right), K, \theta \right) &= \theta^\gamma K^\alpha \left( \frac{\hat{S} (\sigma)}{c} \right)^{\frac{\beta}{n}} - c \frac{\hat{S} (\sigma)}{c} \\
    &= \theta^\gamma K^\alpha \left( \frac{\hat{S} (\sigma)}{c} \right)^{\frac{\beta}{n}} - \hat{S} (\sigma).
\end{align*}
\]
The derivative of this with respect to \( K \) is
\[
\frac{d}{dK} s \left( e^* \left( K, \hat{\theta} \left( K, \sigma \right) \right), K, \theta \right) = \theta^\gamma K^\alpha c^{-1} \left( \frac{\hat{S} (\sigma)}{c} \right)^{\frac{\beta}{n}}.
\]
Moreover, from (16),
\[
\begin{align*}
    \frac{d}{dK} s \left( e^* \left( K, \theta \right), K, \theta \right) &= \left( 1 - \frac{\beta}{n} \right) \left( \frac{\beta}{nc} \right)^{\frac{\beta}{n} \frac{1}{1-\beta/n}} \theta^{\frac{\gamma}{1-\beta/n}} K^{-\frac{\alpha}{1-\beta/n} - 1} \theta^{-\frac{\gamma}{1-\beta/n}} K^{-\frac{\alpha}{1-\beta/n} - 1} \\
    &= \alpha \left( \frac{\beta}{nc} \right)^{\frac{\beta}{n} \frac{1}{1-\beta/n}} \theta^{\frac{\gamma}{1-\beta/n}} K^{-\frac{\alpha}{1-\beta/n} - 1}.
\end{align*}
\]
So the derivative with respect to \( K \) of the maximand in (11) with constraint (64) used to substitute for \( \hat{\theta} \left( K, \sigma \right) \) is, because the two integrands take the same value for \( \theta =
The first-order condition for optimal \( K \) is obtained by setting this equal to zero which, divided through by \( K^{k-1} \) gives the condition for \( \hat{K}^G (\sigma) \)

\[
\hat{K}^G (\sigma) = \left( \frac{\hat{S}(\sigma)}{c} \right)^{\frac{1-\beta/n}{\beta}} \left( \frac{\beta}{nc} \right)^{\frac{1}{2}} \hat{\theta}^G (\sigma)^{-\frac{\gamma}{2}}.
\]
Use of this in (65) and multiplication through by \((1 - \delta) / a\) gives the first-order condition for optimal general capital in terms of \(\hat{\theta}^G(\sigma)\) as

\[
\left( \frac{\beta}{nc} \right)^{\frac{\beta n}{1-\beta/n}} \left[ \left( \frac{\hat{S}(\sigma)}{c} \right)^{1-\frac{\beta n}{a}} \left( \frac{\beta}{nc} \right)^{-\frac{1}{2}} \hat{\theta}^G(\sigma)^{-\frac{1}{x}} \right] \int_{\hat{\theta}}^{\hat{\theta}^G(\sigma)} \theta^{1-\frac{\gamma}{\beta/n}} dF(\theta, \sigma)
\]

or

\[
\left( \frac{\beta}{nc} \right)^{\frac{\beta n}{1-\beta/n} - \frac{1}{a}} \left( \frac{\hat{S}(\sigma)}{c} \right)^{1-\frac{\beta n}{a}} \hat{\theta}^G(\sigma)^{-\frac{1}{x}} \int_{\hat{\theta}}^{\hat{\theta}^G(\sigma)} \theta^{1-\frac{\gamma}{\beta/n}} dF(\theta, \sigma)
\]

or

\[
\left( \frac{\beta}{nc} \right)^{\frac{\beta n}{1-\beta/n} + \frac{1}{a}} \left( \frac{\hat{S}(\sigma)}{c} \right)^{1-\frac{\beta n}{a}} \hat{\theta}^G(\sigma)^{1-\frac{\beta n}{a}} \int_{\hat{\theta}}^{\hat{\theta}^G(\sigma)} \theta^{1-\frac{\gamma}{\beta/n}} dF(\theta, \sigma)
\]

or

\[
\left( \frac{\beta}{nc} \right)^{\frac{1-\beta n}{a}} \left( \frac{\hat{S}(\sigma)}{c} \right)^{1-\frac{\beta n}{a}} \hat{\theta}^G(\sigma)^{1-\frac{\beta n}{a}} \int_{\hat{\theta}}^{\hat{\theta}^G(\sigma)} \theta^{1-\frac{\gamma}{\beta/n}} dF(\theta, \sigma)
\]

or

\[
\hat{S}(\sigma)^{1-\frac{\beta n}{a}} \hat{\theta}^G(\sigma)^{1-\frac{\beta n}{a}} \int_{\hat{\theta}}^{\hat{\theta}^G(\sigma)} \theta^{1-\frac{\gamma}{\beta/n}} dF(\theta, \sigma) + \hat{\theta}^G(\sigma)^{1-\frac{1}{\beta/n}} \int_{\hat{\theta}^G(\sigma)}^{\hat{\theta}} \theta^{1-\frac{\gamma}{\beta/n}} dF(\theta, \sigma) = (1 - \delta) \frac{kC}{a} \left( \frac{\beta}{nc} \right)^{1-\frac{k}{a} \gamma^{1-\frac{\beta n}{a}}} \cdot \frac{1}{c} \cdot \hat{\theta}^G(\sigma)^{1-\frac{\beta n}{a}} \int_{\hat{\theta}}^{\hat{\theta}^G(\sigma)} \theta^{1-\frac{\gamma}{\beta/n}} dF(\theta, \sigma)
\]

40
With the calibration \( \gamma = 1 - \beta/n \), this becomes

\[
S(\sigma)^{1 - (1 - \frac{\beta}{n}) \frac{k}{\alpha}} \hat{\theta}^G(\sigma) \left(1 - \frac{\beta}{n}\right)^{\frac{k}{\alpha} - 1} \left\{ \int_{\theta}^{\hat{\theta}^G(\sigma)} \theta dF(\theta, \sigma) \right. \\
+ \hat{\theta}^G(\sigma)^{\frac{\beta}{n}} \int_{\theta}^{\hat{\theta}} \theta^{\frac{1}{\beta} - \frac{\beta}{n}} dF(\theta, \sigma) \right\} = (1 - \delta) \frac{kC}{\alpha} \left(\frac{\beta}{n}\right)^{1 - \frac{k}{\alpha}} e^{\frac{\beta}{n} k}. 
\]

Note that the right-hand side of this is independent of \( \sigma \) and that the left-hand side can be written as the expression in (20). (21) follows directly from taking the ratio of the expression in (66) for \( \sigma_L \) to that for \( \sigma_H \) along with the calibration \( \gamma = 1 - \beta/n \).

**Derivation of Equations (22) and (23).** It is convenient for calculations to use constraint (6) to solve for \( K \) in terms of \( \tilde{\theta} \) and then use this to optimise over \( \tilde{\theta} \) rather than over \( K \) in the choice of optimal capital in (11). From (16) and (17),

\[
s*e^*(K, \theta), K, \theta) - s*e^*(K, \tilde{\theta}), K, \theta) = \left(1 - \frac{\beta}{n}\right) \left(\frac{\beta}{nc}\right)^{\frac{\beta}{n} - \frac{\beta}{n}} K^{\frac{-\beta}{n} \frac{\beta}{n}} - \left(\frac{\beta}{nc}\right)^{\frac{\beta}{n} - \frac{\beta}{n}} \tilde{\theta}^{\frac{-\beta}{n} \frac{\beta}{n}} \left[\left(\frac{\theta}{\tilde{\theta}}\right)^{\gamma} - \frac{\beta}{n}\right] \\
= \left(\frac{\beta}{nc}\right)^{\frac{\beta}{n} - \frac{\beta}{n}} K^{\frac{-\beta}{n} \frac{\beta}{n}} \left\{ \left(1 - \frac{\beta}{n}\right) \theta^{\frac{\beta}{n} \frac{\beta}{n}} - \tilde{\theta}^{\frac{\beta}{n} \frac{\beta}{n}} \left[\left(\frac{\theta}{\tilde{\theta}}\right)^{\gamma} - \frac{\beta}{n}\right] \right\}. 
\]

Then, with the use of (15), (6) can be written

\[
\frac{\delta}{1 - \delta} \left\{ \int_{\theta}^{\tilde{\theta}} \left(1 - \frac{\beta}{n}\right) \left(\frac{\beta}{nc}\right)^{\frac{\beta}{n} - \frac{\beta}{n}} \theta^{\frac{\beta}{n} \frac{\beta}{n}} K^{\frac{-\beta}{n} \frac{\beta}{n}} dF(\theta, \sigma) - s(K, \sigma) \\
- \int_{\theta}^{\tilde{\theta}} \left(\frac{\beta}{nc}\right)^{\frac{\beta}{n} - \frac{\beta}{n}} K^{\frac{-\beta}{n} \frac{\beta}{n}} \left[\left(1 - \frac{\beta}{n}\right) \theta^{\frac{\beta}{n} \frac{\beta}{n}} - \tilde{\theta}^{\frac{\beta}{n} \frac{\beta}{n}} \left[\left(\frac{\theta}{\tilde{\theta}}\right)^{\gamma} - \frac{\beta}{n}\right] \right] dF(\theta, \sigma) \right\} \\
- c \left(\frac{\beta}{nc}\right)^{\frac{\beta}{n}} K^{\frac{-\beta}{n} \frac{\beta}{n}} = 0 \tag{67}
\]
or

\[
\frac{\delta}{1 - \delta} \left\{ \left( 1 - \frac{\beta}{n} \right) \left( \frac{\beta}{nc} \right)^{\frac{\beta}{n}} K^{\frac{n}{\beta - 1}} \int_{\tilde{\theta}}^{\tilde{\theta}} \left[ \left( 1 - \frac{\beta}{n} \right) \tilde{\theta}^{\frac{\gamma}{n}} - \tilde{\theta}^{\frac{\gamma}{n}} \left( \left( \frac{\theta}{\tilde{\theta}} \right)^{\gamma} - \frac{\beta}{n} \right) \right] dF (\theta, \sigma) \right\} 
- \left( \frac{\beta}{nc} \right)^{\frac{\beta}{n}} K^{\frac{n}{\beta - 1}} \int_{\tilde{\theta}}^{\tilde{\theta}} \left[ \left( 1 - \frac{\beta}{n} \right) \tilde{\theta}^{\frac{\gamma}{n}} - \tilde{\theta}^{\frac{\gamma}{n}} \left( \left( \frac{\theta}{\tilde{\theta}} \right)^{\gamma} - \frac{\beta}{n} \right) \right] dF (\theta, \sigma)
\right\} 
- e \left( \frac{\beta}{nc} \right)^{\frac{\beta}{n}} K^{\frac{n}{\beta - 1}} \left( \frac{\beta}{nc} \right)^{\frac{\beta}{n}} \tilde{\theta}^{\frac{\gamma}{n}} = \frac{\delta}{1 - \delta} s (K, \sigma)
\] (68)

or

\[
\left( \frac{\beta}{nc} \right)^{\frac{\beta}{n}} K^{\frac{n}{\beta - 1}} \left\{ \frac{\delta}{1 - \delta} \left( 1 - \frac{\beta}{n} \right) E \left( \tilde{\theta}^{\frac{\gamma}{n}} | \sigma \right) 
- \frac{\delta}{1 - \delta} \int_{\tilde{\theta}}^{\tilde{\theta}} \left[ \left( 1 - \frac{\beta}{n} \right) \tilde{\theta}^{\frac{\gamma}{n}} - \tilde{\theta}^{\frac{\gamma}{n}} \left( \left( \frac{\theta}{\tilde{\theta}} \right)^{\gamma} - \frac{\beta}{n} \right) \right] dF (\theta, \sigma)
\right\} 
= \frac{\delta}{1 - \delta} s (K, \sigma)
\] (69)

or

\[
K^{\frac{n}{\beta - 1}} = \frac{\delta}{1 - \delta} s (K, \sigma) \left( \frac{\beta}{nc} \right)^{\frac{\beta}{n}} K^{\frac{n}{\beta - 1}} \left\{ \frac{\delta}{1 - \delta} \left( 1 - \frac{\beta}{n} \right) E \left( \tilde{\theta}^{\frac{\gamma}{n}} | \sigma \right) 
- \frac{\delta}{1 - \delta} \int_{\tilde{\theta}}^{\tilde{\theta}} \left[ \left( 1 - \frac{\beta}{n} \right) \tilde{\theta}^{\frac{\gamma}{n}} - \tilde{\theta}^{\frac{\gamma}{n}} \left( \left( \frac{\theta}{\tilde{\theta}} \right)^{\gamma} - \frac{\beta}{n} \right) \right] dF (\theta, \sigma)
\right\}^{-1}
\] (70)

or

\[
K = \left( \frac{\delta}{1 - \delta} s (K, \sigma) \right)^{\frac{1 - \beta}{\beta - 1}} \left( \frac{\beta}{nc} \right)^{\frac{\beta}{n}} K^{\frac{n}{\beta - 1}} \left\{ \frac{\delta}{1 - \delta} \left( 1 - \frac{\beta}{n} \right) E \left( \tilde{\theta}^{\frac{\gamma}{n}} | \sigma \right) 
- \frac{\delta}{1 - \delta} \int_{\tilde{\theta}}^{\tilde{\theta}} \left[ \left( 1 - \frac{\beta}{n} \right) \tilde{\theta}^{\frac{\gamma}{n}} - \tilde{\theta}^{\frac{\gamma}{n}} \left( \left( \frac{\theta}{\tilde{\theta}} \right)^{\gamma} - \frac{\beta}{n} \right) \right] dF (\theta, \sigma)
\right\}^{-\frac{1 - \beta}{\beta - 1}}
\] (71)
or

\[
K = \left( \frac{\delta}{1 - \delta} s(K, \sigma) \right)^{1 - \beta/n} \left( \frac{\beta}{n \bar{c}} \right)^{-\delta/n} \left\{ \frac{\delta}{1 - \delta} \left( 1 - \frac{\beta}{n} \right) E \left( \theta^{1 - \beta/n} \mid \sigma \right) \right. \\
\left. - \frac{\delta}{1 - \delta} \int_{\tilde{\theta}}^\theta \left[ \left( 1 - \frac{\beta}{n} \right) \theta^{1 - \beta/n} - \tilde{\theta}^{1 - \beta/n} \theta^\gamma + \frac{\beta}{n} \tilde{\theta}^{1 - \beta/n} \right] dF(\theta, \sigma) \right\}^{-1} - \frac{\beta}{n} \tilde{\theta}^{1 - \beta/n} \right\}^{1 - \beta/n}. (72)
\]

For specific capital, \( s(K, \sigma) \) is independent of \( K \) and so is written here as \( s(\sigma) \), so the right-hand side is independent of \( K \) for given \( \tilde{\theta} \). We can thus, with some abuse of notation for this derivation, replace \( K \) by \( K(\tilde{\theta}, \sigma) \) to give

\[
K(\tilde{\theta}, \sigma) = \left( \frac{\delta}{1 - \delta} s(\sigma) \right)^{1 - \beta/n} \left( \frac{\beta}{n \bar{c}} \right)^{-\delta/n} \left\{ \frac{\delta}{1 - \delta} \left( 1 - \frac{\beta}{n} \right) E \left( \theta^{1 - \beta/n} \mid \sigma \right) \right. \\
\left. - \frac{\delta}{1 - \delta} \int_{\tilde{\theta}}^\theta \left[ \left( 1 - \frac{\beta}{n} \right) \theta^{1 - \beta/n} - \tilde{\theta}^{1 - \beta/n} \theta^\gamma + \frac{\beta}{n} \tilde{\theta}^{1 - \beta/n} \right] dF(\theta, \sigma) \right\}^{-1} - \frac{\beta}{n} \tilde{\theta}^{1 - \beta/n} \right\}^{1 - \beta/n}. (73)
\]

The derivative of this with respect to \( \tilde{\theta} \) is, because the integrand equals zero for \( \theta = \tilde{\theta} \),

\[
K_1(\tilde{\theta}, \sigma) = \left( \frac{\delta}{1 - \delta} s(\sigma) \right)^{1 - \beta/n} \left( \frac{\beta}{n \bar{c}} \right)^{-\delta/n} \left\{ \frac{\delta}{1 - \delta} \left( 1 - \frac{\beta}{n} \right) E \left( \theta^{1 - \beta/n} \mid \sigma \right) \right. \\
\left. - \frac{\delta}{1 - \delta} \int_{\tilde{\theta}}^\theta \left[ \left( 1 - \frac{\beta}{n} \right) \theta^{1 - \beta/n} - \tilde{\theta}^{1 - \beta/n} \theta^\gamma + \frac{\beta}{n} \tilde{\theta}^{1 - \beta/n} \right] dF(\theta, \sigma) \right\}^{-1} \left\{ \frac{\gamma}{1 - \beta/n} \tilde{\theta}^{1 - \beta/n} - \frac{\beta}{n} \tilde{\theta}^{1 - \beta/n} \right\} dF(\theta, \sigma) \\
- \left. \frac{\beta}{n} \tilde{\theta}^{1 - \beta/n} \right\} \right\}^{1 - \beta/n} \left\{ \frac{\delta}{1 - \delta} \left( 1 - \frac{\beta}{n} \right) \left[ \frac{\gamma}{1 - \beta/n} \tilde{\theta}^{1 - \beta/n} - \frac{\beta}{n} \tilde{\theta}^{1 - \beta/n} \right] dF(\theta, \sigma) \right. \\
\left. - \frac{\beta}{n} \tilde{\theta}^{1 - \beta/n} \right\} \right\} \right\} \right\} (74)
\]

43
or

\[ K_1 \left( \tilde{\theta}, \sigma \right) = - \left( \frac{\delta}{1 - \delta} \right)^{1 - \beta/n} \left( \frac{\beta}{n \alpha} \right)^{\beta/n} \left( \frac{1 - \beta/n}{\alpha} \right) \left\{ \frac{\delta}{1 - \delta} \left( 1 - \frac{\beta}{n} \right) E \left( \theta^{1 - \gamma_{p/n}} | \sigma \right) \right. \]

\[ - \frac{\delta}{1 - \delta} \int^{\tilde{\sigma}} \left[ \left( 1 - \frac{\beta}{n} \right) \theta^{1 - \gamma_{p/n}} - \tilde{\theta}^{\gamma_{p/n}} \theta^\gamma + \frac{\beta}{n} \tilde{\theta}^{1 - \gamma_{p/n}} \right] dF \left( \theta, \sigma \right) \]

\[ - \frac{\beta}{n} \tilde{\theta}^{1 - \gamma_{p/n}} \right\}^{1 - \beta/n - 1} \left\{ \frac{\delta}{1 - \delta} \int^{\tilde{\sigma}} \left[ \tilde{\theta}^{\gamma_{p/n} - 1} \theta^\gamma - \tilde{\theta}^{1 - \gamma_{p/n} - 1} \right] dF \left( \theta, \sigma \right) \]

\[ - \tilde{\theta}^{1 - \gamma_{p/n} - 1} \} \right. \] (75)

or

\[ K_1 \left( \tilde{\theta}, \sigma \right) = - \left( \frac{\delta}{1 - \delta} \right)^{1 - \beta/n} \left( \frac{\beta}{n \alpha} \right)^{\beta/n} \left( \frac{1 - \beta/n}{\alpha} \right) \left\{ \frac{\delta}{1 - \delta} \left( 1 - \frac{\beta}{n} \right) E \left( \theta^{1 - \gamma_{p/n}} | \sigma \right) \right. \]

\[ - \frac{\delta}{1 - \delta} \int^{\tilde{\sigma}} \left[ \left( 1 - \frac{\beta}{n} \right) \theta^{1 - \gamma_{p/n}} - \tilde{\theta}^{\gamma_{p/n}} \theta^\gamma + \frac{\beta}{n} \tilde{\theta}^{1 - \gamma_{p/n}} \right] dF \left( \theta, \sigma \right) \]

\[ - \frac{\beta}{n} \tilde{\theta}^{1 - \gamma_{p/n}} \right\}^{1 - \beta/n - 1} \left\{ \frac{\delta}{1 - \delta} \int^{\tilde{\sigma}} \left[ \tilde{\theta}^{\gamma_{p/n} - 1} \theta^\gamma - \tilde{\theta}^{1 - \gamma_{p/n} - 1} \right] dF \left( \theta, \sigma \right) \]

\[ - \tilde{\theta}^{1 - \gamma_{p/n} - 1} \} \right. \} \]. (76)

So, with the use of (73),

\[ \frac{K_1 \left( \tilde{\theta}, \sigma \right)}{K \left( \tilde{\theta}, \sigma \right)} = - \gamma \frac{\beta/n}{\alpha} \left\{ \frac{\delta}{1 - \delta} \left( 1 - \frac{\beta}{n} \right) E \left( \theta^{1 - \gamma_{p/n}} | \sigma \right) \right. \]

\[ - \frac{\delta}{1 - \delta} \int^{\tilde{\sigma}} \left[ \left( 1 - \frac{\beta}{n} \right) \theta^{1 - \gamma_{p/n}} - \tilde{\theta}^{\gamma_{p/n}} \theta^\gamma + \frac{\beta}{n} \tilde{\theta}^{1 - \gamma_{p/n}} \right] dF \left( \theta, \sigma \right) \]

\[ - \frac{\beta}{n} \tilde{\theta}^{1 - \gamma_{p/n}} \right\} \left\{ \frac{\delta}{1 - \delta} \int^{\tilde{\sigma}} \left[ \tilde{\theta}^{\gamma_{p/n} - 1} \theta^\gamma - \tilde{\theta}^{1 - \gamma_{p/n} - 1} \right] dF \left( \theta, \sigma \right) \]

\[ - \tilde{\theta}^{1 - \gamma_{p/n} - 1} \} \right. \} \}. (77)
For the Cobb-Douglas formulation, the maximand in (11) in the paper is

\[ \frac{\delta}{1 - \delta} \left\{ \int_{\theta}^{\bar{\theta}} \left( 1 - \frac{\beta}{n} \right) \left( \frac{\beta}{nc} \right)^{\frac{\beta}{n}} \left( \theta \right)^{\frac{\gamma}{n}} K \theta^{1 - \frac{\beta}{n}} dF(\theta, \sigma) \right\} \]

\[ - \int_{\theta(K, \sigma)}^{\bar{\theta}} \left( 1 - \frac{\beta}{n} \right) \left( \frac{\beta}{nc} \right)^{\frac{\beta}{n}} K^{\frac{\beta}{n}} \left[ \left( 1 - \frac{\beta}{n} \right) \theta^{1 - \frac{\beta}{n}} - \theta(K, \sigma)^{1 - \frac{\beta}{n}} \left( \frac{\theta}{\theta(K, \sigma)} \right)^{\frac{\gamma}{n}} - \frac{\beta}{n} \right] dF(\theta, \sigma) \]

\[- CK^k, \quad (78)\]

which can be written as

\[ \left( \frac{\beta}{nc} \right)^{\frac{\beta}{n}} K^{\frac{n}{1 - \beta/n}} \left( 1 - \frac{\beta}{n} \right) E \left( \theta^{1 - \frac{\beta}{n}} \right) \left\{ \left( 1 - \frac{\beta}{n} \right) E \left( \theta^{1 - \frac{\beta}{n}} \right) \right\} \]

\[- CK^k. \quad (79)\]

With the constraint used to substitute \( K(\hat{\theta}^S(\sigma), \sigma) \) for \( K \), the derivative of this with respect to \( \hat{\theta}(K, \sigma) \) gives the first-order condition for the optimal value \( \hat{\theta}^S(\sigma) \)

\[ \left( \frac{\beta}{nc} \right)^{\frac{\beta}{n}} \frac{\alpha}{1 - \beta/n} K \left( \hat{\theta}^S(\sigma), \sigma \right)^{\frac{\beta}{n} - 1} K_1 \left( \hat{\theta}^S(\sigma), \sigma \right) \left( 1 - \frac{\beta}{n} \right) E \left( \theta^{1 - \frac{\beta}{n}} \right) \left\{ \left( 1 - \frac{\beta}{n} \right) E \left( \theta^{1 - \frac{\beta}{n}} \right) \right\} \]

\[- \int_{\theta^S(\sigma)}^{\bar{\theta}} \left( 1 - \frac{\beta}{n} \right) \theta^{1 - \frac{\beta}{n}} - \theta^S(\sigma)^{1 - \frac{\beta}{n}} \theta^\gamma + \frac{\beta}{n} \theta^S(\sigma)^{1 - \frac{\beta}{n}} dF(\theta, \sigma) \]

\[ + \left( \frac{\beta}{nc} \right)^{\frac{\beta}{n}} K \left( \hat{\theta}^S(\sigma), \sigma \right)^{\frac{n}{1 - \beta/n}} \left( 1 - \frac{\beta}{n} \right) \]

\[ \times \left\{ - \int_{\theta^S(\sigma)}^{\bar{\theta}} \left[ - \frac{\gamma \beta/n}{1 - \beta/n} \hat{\theta}^S(\sigma)^{\frac{\beta}{n} - 1} \theta^\gamma + \frac{\beta}{n} \frac{\gamma}{1 - \beta/n} \hat{\theta}^S(\sigma)^{\frac{\beta}{n} - 1} \right] dF(\theta, \sigma) \right\} \]

\[- CK \left( \hat{\theta}^S(\sigma), \sigma \right)^{k-1} K_1 \left( \hat{\theta}^S(\sigma), \sigma \right) = 0. \quad (80)\]
For the calibration $\gamma = 1 - \beta / n$, this becomes

\[
\left( \frac{\beta}{n} \right)^{\frac{\beta}{1-\beta/n}} K \left( \hat{\theta}^S (\sigma), \sigma \right)^{\frac{\beta}{1-\beta/n} - 1} K \left( \hat{\theta}^S (\sigma), \sigma \right)^{\frac{\delta}{1-\delta}} \left\{ \left( 1 - \frac{\beta}{n} \right) E (\theta | \sigma) \right. \\
- \int_{\hat{\theta}^S(\sigma)}^\theta \left[ \left( 1 - \frac{\beta}{n} \right) \theta - \hat{\theta}^S (\sigma)^{\frac{\beta}{n} \theta^{1-\frac{\beta}{n}} + \frac{\beta}{n} \hat{\theta}^S (\sigma) \right] dF(\theta, \sigma) \right. \\
+ \left. \left( \frac{\beta}{n} \right)^{1-\frac{\beta}{1-\beta/n}} K \left( \hat{\theta}^S (\sigma), \sigma \right)^{\frac{\beta}{1-\beta/n} - \frac{\delta}{1-\delta}} \right. \\
\times \left\{ - \frac{\beta}{n} \int_{\hat{\theta}^S(\sigma)}^\theta \left[ 1 - \hat{\theta}^S (\sigma)^{\frac{\beta}{n} - 1} \theta^{1-\frac{\beta}{n}} \right] dF(\theta, \sigma) \right. \\
- CkK \left( \hat{\theta}^S (\sigma), \sigma \right)^{k-1} K_1 \left( \hat{\theta}^S (\sigma), \sigma \right) = 0. \quad (81)
\]

Division by $K \left( \hat{\theta}^S (\sigma), \sigma \right)^{k-1} K_1 \left( \hat{\theta}^S (\sigma), \sigma \right)$, by $\left( \frac{\beta}{n} \right)^{1-\frac{\beta}{1-\beta/n}}$ and by $\frac{\delta}{1-\delta}$ gives

\[
\frac{\alpha}{1 - \beta / n} K \left( \hat{\theta}^S (\sigma), \sigma \right)^{\frac{n}{1-\beta/n} - k} \left\{ \left( 1 - \frac{\beta}{n} \right) E (\theta | \sigma) \right. \\
- \int_{\hat{\theta}^S(\sigma)}^\theta \left[ \left( 1 - \frac{\beta}{n} \right) \theta - \hat{\theta}^S (\sigma)^{\frac{\beta}{n} \theta^{1-\frac{\beta}{n}} + \frac{\beta}{n} \hat{\theta}^S (\sigma) \right] dF(\theta, \sigma) \right. \\
+ \left. \left. K \left( \hat{\theta}^S (\sigma), \sigma \right)^{\frac{n}{1-\beta/n} + 1-k} K_1 \left( \hat{\theta}^S (\sigma), \sigma \right)^{-1} \right. \\
\times \left. \left\{ - \frac{\beta}{n} \int_{\hat{\theta}^S(\sigma)}^\theta \left[ 1 - \hat{\theta}^S (\sigma)^{\frac{\beta}{n} - 1} \theta^{1-\frac{\beta}{n}} \right] dF(\theta, \sigma) \right. \right. \\
= \frac{1 - \delta}{\delta} Ck \left( \frac{\beta}{n} \right)^{-\frac{\beta}{1-\beta/n}} \frac{\beta}{n} \int_{\hat{\theta}^S(\sigma)}^\theta \left[ 1 - \hat{\theta}^S (\sigma)^{\frac{\beta}{n} - 1} \theta^{1-\frac{\beta}{n}} \right] dF(\theta, \sigma) \right. \right. \\
- \left. Ck \left( \frac{\beta}{n} \right)^{-\frac{\beta}{1-\beta/n}} \frac{\beta}{n} \int_{\hat{\theta}^S(\sigma)}^\theta \left[ 1 - \hat{\theta}^S (\sigma)^{\frac{\beta}{n} - 1} \theta^{1-\frac{\beta}{n}} \right] dF(\theta, \sigma) \right. \right. \right. \\
= \frac{1 - \delta}{\delta} Ck \left( \frac{\beta}{n} \right)^{-\frac{\beta}{1-\beta/n}}. \quad (82)
\]

or

\[
K \left( \hat{\theta}^S (\sigma), \sigma \right)^{\frac{n}{1-\beta/n} - k} \left\{ \alpha E (\theta | \sigma) \right. \\
- \frac{\alpha}{1 - \beta / n} \int_{\hat{\theta}^S(\sigma)}^\theta \left[ \left( 1 - \frac{\beta}{n} \right) \theta - \hat{\theta}^S (\sigma)^{\frac{\beta}{n} \theta^{1-\frac{\beta}{n}} + \frac{\beta}{n} \hat{\theta}^S (\sigma) \right] dF(\theta, \sigma) \right. \\
- \frac{\alpha}{1 - \beta / n} \int_{\hat{\theta}^S(\sigma)}^\theta \left[ \left( 1 - \frac{\beta}{n} \right) \theta - \hat{\theta}^S (\sigma)^{\frac{\beta}{n} \theta^{1-\frac{\beta}{n}} + \frac{\beta}{n} \hat{\theta}^S (\sigma) \right] dF(\theta, \sigma) \right. \\
- \frac{\alpha}{1 - \beta / n} \int_{\hat{\theta}^S(\sigma)}^\theta \left[ \left( 1 - \frac{\beta}{n} \right) \theta - \hat{\theta}^S (\sigma)^{\frac{\beta}{n} \theta^{1-\frac{\beta}{n}} + \frac{\beta}{n} \hat{\theta}^S (\sigma) \right] dF(\theta, \sigma) \right. \\
= \frac{1 - \delta}{\delta} Ck \left( \frac{\beta}{n} \right)^{-\frac{\beta}{1-\beta/n}}. \quad (83)
\]

Substitution for $K(\hat{\theta}^S (\sigma), \sigma)$ from (73) and for $K(\hat{\theta}^S (\sigma), \sigma) / K_1 (\hat{\theta}^S (\sigma), \sigma)$ from (77)
with the calibration \( \gamma = 1 - \beta / n \) gives

\[
\left( \frac{\delta}{1 - \delta} \hat{\sigma} (\sigma) \right)^{1 - k \frac{1 - \beta / n}{n}} \left( \frac{\beta}{n c} \right)^{k \delta / n - \frac{\beta / n}{1 - \delta / n}} \left\{ \frac{\delta}{1 - \delta} \left( 1 - \frac{\beta}{n} \right) E (\theta | \sigma) - \frac{\beta}{n} \hat{\theta}^S (\sigma) \right\}^{k \frac{1 - \beta / n}{n} - 1} \}

\[- \frac{\delta}{1 - \delta} \int_{\hat{\theta}^S (\sigma)} \left( \left( 1 - \frac{\beta}{n} \right) \theta - \hat{\theta}^S (\sigma) \hat{\theta} \theta^1 - \frac{\beta}{n} + \frac{\beta}{n} \hat{\theta}^S (\sigma) \right) dF (\theta, \sigma) \}

\[- \frac{\alpha}{1 - \beta / n} \int_{\hat{\theta}^S (\sigma)} \left[ \left( 1 - \frac{\beta}{n} \right) \theta - \hat{\theta}^S (\sigma) \hat{\theta} \theta^1 - \frac{\beta}{n} + \frac{\beta}{n} \hat{\theta}^S (\sigma) \right] dF (\theta, \sigma) \}

\[- \frac{\delta}{1 - \delta} \int_{\hat{\theta}^S (\sigma)} \left( \left( 1 - \frac{\beta}{n} \right) \theta - \hat{\theta}^S (\sigma) \hat{\theta} \theta^1 - \frac{\beta}{n} + \frac{\beta}{n} \hat{\theta}^S (\sigma) \right) dF (\theta, \sigma) \}

\[- \frac{\beta}{n} \left( 1 - \hat{\theta}^S (\sigma) \hat{\theta} \theta^1 - \frac{\beta}{n} \right) dF (\theta, \sigma) \}

\[- \frac{1}{\delta} \left[ \left( 1 - \hat{\theta}^S (\sigma) \hat{\theta} \theta^1 - \frac{\beta}{n} \right) dF (\theta, \sigma) \right] = \frac{1 - \delta}{\beta} Ck \left( \frac{\beta}{n c} \right)^{\frac{\beta / n}{1 - \beta / n}} \cdot (84) \]

When divided through by \( \alpha \left( \frac{\beta}{n c} \right)^{k \delta / n - \frac{\beta / n}{1 - \beta / n}} \), the right-hand side of this is independent of \( \sigma \) and the left-hand side is just the expression in (22). (23) follows from (73) with \( \gamma = 1 - \beta / n \) for the optimal values \( \hat{\theta} = \hat{\theta}^S (\sigma) \) and \( \hat{R}^S (\sigma) = K \left( \hat{\theta}^S (\sigma) , \sigma \right) \).

**Effect of change in idiosyncratic risk on general capital.** The following argument shows that the effect of a change in idiosyncratic risk on general capital is greater in magnitude than that of a change in systemic risk for \( \left( 1 - \frac{\beta}{n} \right) \frac{k}{\alpha} > 1 \) as with the parameter values in Table 1. It follows from (20) that, for two different levels of risk \( \sigma_H \) and \( \sigma_L \),

\[
\hat{\theta}^G (\sigma_H) \left( 1 - \frac{\beta}{n} \right) \frac{k}{\alpha} \left\{ E (\theta | \sigma_H) - \int_{\hat{\theta}^G (\sigma_H)} \left[ \theta - \hat{\theta}^G (\sigma_H) \hat{\theta} \theta^1 - \frac{\beta}{n} \right] dF (\theta, \sigma_H) \right\} \]

\[
= \left( \frac{\hat{S} (\sigma_H)}{\hat{S} (\sigma_L)} \right)^{1 - \frac{\beta}{n}} \hat{\theta}^G (\sigma_L) \left( 1 - \frac{\beta}{n} \right) \frac{k}{\alpha} \left\{ E (\theta | \sigma_L) - \int_{\hat{\theta}^G (\sigma_L)} \left[ \theta - \hat{\theta}^G (\sigma_L) \hat{\theta} \theta^1 - \frac{\beta}{n} \right] dF (\theta, \sigma_L) \right\} \cdot (85) \]

For \( \left( 1 - \frac{\beta}{n} \right) \frac{k}{\alpha} > 1 \), The total derivative of this with respect to \( \hat{S} (\sigma_H) / \hat{S} (\sigma_L) \) for given \( \hat{\theta}^G (\sigma_L) \) has the same sign as the total derivative with respect to \( x \equiv [\hat{S} (\sigma_H) / \hat{S} (\sigma_L)] \left( 1 - \frac{\beta}{n} \right) \frac{k}{\alpha} \).
for given $\kappa > 0$ of

$$
\hat{\vartheta}^G (\sigma_H) \left(1 - \frac{\beta}{\alpha} \right) \frac{1}{\alpha} \int_{\hat{\vartheta}^G(\sigma_H)}^\theta \left[ \theta - \hat{\vartheta}^G (\sigma_H) \theta^{1-\frac{\beta}{\alpha}} \right] dF (\theta, \sigma_H) = \kappa.
$$

Because the integrand in this equals zero for $\theta = \hat{\vartheta}^G (\sigma_H)$, this differentiation gives

$$
\left[ \left(1 - \frac{\beta}{\alpha} \right) \frac{k}{\alpha} - 1 \right] \hat{\vartheta}^G (\sigma_H) \left(1 - \frac{\beta}{\alpha} \right) \frac{1}{\alpha} \frac{d\hat{\vartheta}^G (\sigma_H)}{dx} \left\{ E (\theta, \sigma_H) - \int_{\hat{\vartheta}^G(\sigma_H)}^\theta \left[ \theta - \hat{\vartheta}^G (\sigma_H) \theta^{1-\frac{\beta}{\alpha}} \right] dF (\theta, \sigma_H) \right\} = \kappa
$$
or

$$
\frac{d\hat{\vartheta}^G (\sigma_H)}{dx} \left[ \left(1 - \frac{\beta}{\alpha} \right) \frac{k}{\alpha} - 1 \right] \hat{\vartheta}^G (\sigma_H) \left(1 - \frac{\beta}{\alpha} \right) \frac{1}{\alpha} \frac{d\hat{\vartheta}^G (\sigma_H)}{dx} \left\{ E (\theta, \sigma_H) - \int_{\hat{\vartheta}^G(\sigma_H)}^\theta \left[ \theta - \hat{\vartheta}^G (\sigma_H) \theta^{1-\frac{\beta}{\alpha}} \right] dF (\theta, \sigma_H) \right\} = \kappa
$$
or

$$
\frac{d\hat{\vartheta}^G (\sigma_H)}{dx} \hat{\vartheta}^G (\sigma_H) \left(1 - \frac{\beta}{\alpha} \right) \frac{1}{\alpha} \frac{d\hat{\vartheta}^G (\sigma_H)}{dx} \left\{ E (\theta, \sigma_H) - \int_{\hat{\vartheta}^G(\sigma_H)}^\theta \left[ \theta - \hat{\vartheta}^G (\sigma_H) \theta^{1-\frac{\beta}{\alpha}} \right] dF (\theta, \sigma_H) \right\} = \kappa.
$$

For $\left(1 - \frac{\beta}{\alpha} \right) \frac{k}{\alpha} > 1$, the terms in square brackets in this are all positive and thus so is the term in braces. Since $\kappa > 0$, this implies $d\hat{\vartheta}^G (\sigma_H) / dx > 0$. Since for idiosyncratic risk $\hat{\vartheta} (\sigma_H) / \hat{\vartheta} (\sigma_L)$ is smaller than for systemic risk when $\sigma_H$ is more risky than $\sigma_L$, this implies that $\hat{\vartheta}^G (\sigma_H)$ is lower for idiosyncratic than for systemic risk for given $\hat{\vartheta}^G (\sigma_L)$. 
From (21) and (85),

\[
\left[ \hat{K}_G (\sigma_L) \right]^{-\frac{\alpha - \beta}{\pi n}} \left\{ E (\theta, \sigma_H) - \int_{\hat{\theta}_G (\sigma_H)}^{\overline{\theta}} \theta \left[ 1 - \left( \frac{\hat{\theta}_G (\sigma_H)}{\theta} \right)^{\beta \pi} \right] dF (\theta, \sigma_H) \right\} = E (\theta, \sigma_L) - \int_{\hat{\theta}_G (\sigma_L)}^{\overline{\theta}} \theta \left[ 1 - \left( \frac{\hat{\theta}_G (\sigma_L)}{\theta} \right)^{\beta \pi} \right] dF (\theta, \sigma_L).
\]

The right-hand side of this is independent of \( \hat{S} (\sigma_H) / \hat{S} (\sigma_L) \) for given \( \hat{\theta}_G (\sigma_L) \), so the left-hand side must be too. Thus, for \((1 - \frac{\beta}{\pi}) \frac{k}{\alpha} > 1\), the effect on \( \hat{K}_G (\sigma_L) / \hat{K}_G (\sigma_H) \) of an increase in \( \hat{S} (\sigma_H) / \hat{S} (\sigma_L) \) must be the opposite of the effect of the increase on

\[
E (\theta, \sigma_H) - \int_{\hat{\theta}_G (\sigma_H)}^{\overline{\theta}} \theta \left[ 1 - \left( \frac{\hat{\theta}_G (\sigma_H)}{\theta} \right)^{\beta \pi} \right] dF (\theta, \sigma_H),
\]

which, given \( d\hat{\theta}_G (\sigma_H) / dx > 0 \), is the same as the effect of an increase in \( \hat{\theta}_G (\sigma_H) \) for given \( E (\theta, \sigma_H) \). Because the integrand equals zero for \( \theta = \hat{\theta}_G (\sigma_H) \), the derivative of this expression with respect to \( \hat{\theta}_G (\sigma_H) \) is

\[
\int_{\hat{\theta}_G (\sigma_H)}^{\overline{\theta}} \frac{\beta}{n} \left( \frac{\hat{\theta}_G (\sigma_H)}{\theta} \right)^{\frac{\beta \pi - 1}{\pi}} dF (\theta, \sigma_H) > 0.
\]

So an increase in \( \hat{S} (\sigma_H) / \hat{S} (\sigma_L) \) reduces \( \hat{K}_G (\sigma_L) / \hat{K}_G (\sigma_H) \). Thus, since idiosyncratic risk has a lower value of \( \hat{S} (\sigma_H) / \hat{S} (\sigma_L) \) than systemic risk, it has a higher value of \( \hat{K}_G (\sigma_L) / \hat{K}_G (\sigma_H) \) and, hence a lower value of \( \hat{K}_G (\sigma_H) \) for given \( \hat{K}_G (\sigma_L) \).

**Concavity of first derivative term.** To derive conditions under which \( h_2 (e, K, \theta) \) is concave in \( \theta \), use the simplified notation without arguments for the function \( s \) but \(*\) to indicate that the first argument is \( e^* (K, \theta) \). So, for example,

\[
s_{ij} = s_{ij} (e, K, \theta); \quad s^*_{ij} = s_{ij} (e^* (K, \theta), K, \theta).
\]
Then
\[
h_{233}(e, K, \theta) = s_{233} - s_{223} - s_{123}e_2^* - s_{12}e_2^2 - s_{112}e_2^2 - s_{113}e_2^* - s_{11}e_2^* - s_{113}e_2^* - s_{112}e_2^2 - s_{113}e_2^* - s_{113}e_2^* - s_{113}e_2^* - s_{113}e_2^* - s_{113}e_2^* + 2s_{13}e_2^* - s_{113}e_2^* - s_{113}e_2^* - s_{113}e_2^* - s_{113}e_2^* - s_{113}e_2^* - (s_{12} + s_{11}e_2^*) e_2^2 - s_{113}e_2^* - (s_{12} + 2s_{11}e_2^*) e_2^2 - s_{113}e_2^* - s_{113}e_2^*,
\]
the second equality following because \( s_1^* = 0 \). Moreover,
\[
e_1^*(K, \theta) = \frac{s_{12}(e^*(K, \theta), K, \theta)}{s_{11}(e^*(K, \theta), K, \theta)}
\]
\[
e_2^*(K, \theta) = \frac{s_{13}(e^*(K, \theta), K, \theta)}{s_{11}(e^*(K, \theta), K, \theta)}
\]
\[
e_{12}^*(K, \theta) = -\frac{1}{s_{11}(e^*(K, \theta), K, \theta)^2} \{ s_{11}(e^*(K, \theta), K, \theta) [s_{113}(e^*(K, \theta), K, \theta) e_2^*(K, \theta) + s_{133}(e^*(K, \theta), K, \theta) - s_{13}(e^*(K, \theta), K, \theta) [s_{111}(e^*(K, \theta), K, \theta) e_2^*(K, \theta) + s_{113}(e^*(K, \theta), K, \theta)]] \}.
\]
From these.
\[
e_{22}^* = -\frac{1}{s_{11}^2} \{ s_{11}^*[s_{133} + s_{113}e_2^*] - s_{13}^*[s_{113} + s_{111}e_2^*] \}.
\]
With the use of these
\[
h_{233}(e, K, \theta) = s_{233} - s_{223} + 2s_{123} \frac{s_{13}^*}{s_{11}^*} - s_{112}^* \left( \frac{s_{13}^*}{s_{11}^*} \right)^2 - s_{113}^* \frac{s_{12}^*}{s_{11}^*} + s_{113}^* \frac{s_1^*}{s_{11}^*} - \left( s_{12}^* - s_{11}^* \frac{s_{12}^*}{s_{11}^*} \right) \frac{1}{s_{11}^2} \left\{ s_{11}^* \left[ s_{133}^* - s_{113}^* \frac{s_{13}^*}{s_{11}^*} \right] - s_{13}^* \left[ s_{133}^* - s_{111}^* \frac{s_{13}^*}{s_{11}^*} \right] \right\}
\]
\[
- \left( -s_{11}^* \frac{s_{13}^*}{s_{11}^*} + 2s_{11}^* \frac{s_{12}^*}{s_{11}^*} \right) \frac{1}{s_{11}^2} \left[ s_{11}^* \left( -s_{113}^* \frac{s_{13}^*}{s_{11}^*} + s_{113}^* \right) - s_{13}^* \left( -s_{111}^* \frac{s_{13}^*}{s_{11}^*} + s_{113}^* \right) \right]
\]
\[
+ s_{133}^* \frac{s_{12}^*}{s_{11}^*} - s_{113}^* \left( \frac{s_{12}^*}{s_{11}^*} \right)^2
\]
or
\[
h_{233}(e, K, \theta) = s_{233} - s_{223} + 2s_{123} \frac{s_{13}^*}{s_{11}^*} - s_{112}^* \left( \frac{s_{13}^*}{s_{11}^*} \right)^2 - s_{113}^* \frac{s_{12}^*}{s_{11}^*} + s_{113}^* \frac{s_1^*}{s_{11}^*} - \left( s_{12}^* - s_{11}^* \frac{s_{12}^*}{s_{11}^*} \right) \frac{1}{s_{11}^2} \left\{ s_{11}^* \left[ s_{133}^* - s_{113}^* \frac{s_{13}^*}{s_{11}^*} \right] - s_{13}^* \left[ s_{111}^* \frac{s_{13}^*}{s_{11}^*} + s_{113}^* \right] \right\}
\]
\[
- \left( -s_{11}^* \frac{s_{13}^*}{s_{11}^*} + 2s_{11}^* \frac{s_{12}^*}{s_{11}^*} \right) \frac{1}{s_{11}^2} \left[ s_{11}^* \left( -s_{113}^* \frac{s_{13}^*}{s_{11}^*} + s_{113}^* \right) - s_{13}^* \left( -s_{111}^* \frac{s_{13}^*}{s_{11}^*} + s_{113}^* \right) \right]
\]
\[
+ s_{133}^* \frac{s_{12}^*}{s_{11}^*} - s_{113}^* \left( \frac{s_{12}^*}{s_{11}^*} \right)^2
\]
or

\[ h_{233}(e, K, \theta) = s_{233} - s^*_2 + 2s^*_{123} \left( \frac{s^*_{13}}{s^*_{11}} \right) - s^*_{112} \left( \frac{s^*_{13}}{s^*_{11}} \right)^2 - s^*_{113} \left( \frac{s^*_{13}}{s^*_{11}} \right)^2 + s^*_{113} \left( \frac{s^*_{12}}{s^*_{11}} \right)^2 \]

or

\[ h_{233}(e, K, \theta) = s_{233} - s^*_2 + 2s^*_{123} \left( \frac{s^*_{13}}{s^*_{11}} \right) - s^*_{112} \left( \frac{s^*_{13}}{s^*_{11}} \right)^2 - 3s^*_{113} \left( \frac{s^*_{12}}{s^*_{11}} \right)^2 + s^*_{113} \left( \frac{s^*_{12}}{s^*_{11}} \right)^2 \]

or

\[ h_{233}(e, K, \theta) = s_{233} - s^*_2 + 2s^*_{123} \left( \frac{s^*_{13}}{s^*_{11}} \right) - s^*_{112} \left( \frac{s^*_{13}}{s^*_{11}} \right)^2 + s^*_{111} \left[ s^*_{12} - s^*_{13} \right] \]

\[ + \left(s^*_{113} \left[ -3s^*_{12} s^*_{13} + s^*_{13} \left( \frac{s^*_{13}}{s^*_{11}} \right) + 2s^*_{13} - s^*_{12} \right] \right) + s^*_{113} \left[ 2s^*_{12} - s^*_{13} \right]. \]

**Appendix C Calculations for calibration**

All numerical calculations were done with Scientific WorkPlace 6.0.27.

**C.1 Bloom et al. (2018) calibrations**

In Bloom et al. (2018), total factor productivity (TFP) is the product of two components, an aggregate component that affects all firms and an idiosyncratic component that affects an individual firm. The log of each component of TFP is assumed to follow a first-order autoregressive process with the variance of innovations constant within each of the two regimes with Markov switching between regimes.

In the formulation used here with stationary TFP, the autoregressive formulation with innovations normally distributed corresponds to \( \theta \) being lognormal. Table 5 in Bloom et al. (2018) gives estimates of \( \sigma_L^A = 0.0067 \) for the aggregate component (\( L \) refers to the low-risk regime, \( H \) to the high-risk regime) and \( \sigma_L^Z = 0.051 \) for the idiosyncratic component for the log autoregressive process, with \( \sigma_L^A / \sigma_L^Z = 1.6 \) and \( \sigma_H^Z / \sigma_L^Z = 4.1 \). But these apply to TFP with exponent 1 in the production function, whereas in the model used here TFP is given by \( \theta^γ \). The standard errors for \( \ln \theta \) in the calibration used here need to be adjusted accordingly. For \( X \sim N(\mu, \sigma^2) \), then \( aX \sim N(a\mu, a^2\sigma^2) \) for \( a > 0 \). Also, for \( X_H \sim N(\mu_1, \sigma_H^2) \) and \( X_L \sim N(\mu_2, \sigma_L^2) \) independent, \( X_H + X_L \sim \)
\( N (\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2) \). With these formulae, \( \gamma \ln \theta \sim N \left( \mu_1 + \mu_2, \frac{1}{\gamma^2} \left( \sigma_1^2 + \sigma_2^2 \right) \right) \) for \( j = L, H, \) so \( \ln \theta \sim N \left( \mu_1 + \mu_2, \frac{1}{\gamma^2} \right) \). The values for \( \sigma_L \) and \( \sigma_H \) in Table 1 are calculated using this formula with \( \gamma = 1 - \beta / n \) as used in the calibration.

A calibration for the other parameters corresponding to Bloom et al. (2018, Table 4) is
\[
k = 1, \alpha = \frac{1}{4}, \beta = \frac{1}{2}, n = 1 \implies \left( 1 - \frac{\beta}{n} \right) k = \frac{1}{2} \times 4 = 2. \tag{87}
\]
The values of \( \alpha \) and \( \beta \) are derived from a labour share of 2/3, a capital share of 1/3 and an isoelastic demand with 33% markup.

### C.2 Lognormal distribution

When \( x \) is log-normally distributed, \( \ln x \sim N (\mu, \sigma^2) \) and the probability density function for \( x \) is given by
\[
\frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}. \tag{88}
\]
The mean and variance are given by
\[
E(x) = e^{(\mu + \sigma^2/2)}; \text{var}(x) = (e^{\sigma^2} - 1)e^{2\mu + \sigma^2}. \tag{89}
\]
Note that changing \( \sigma \) changes the mean unless \( \mu \) is changed correspondingly. The normalisation used in the calibration that \( E(x) = 1 \) requires \( \mu = -\sigma^2/2 \), in which case the variance is just \( e^{\sigma^2} - 1 \). With that normalisation, the probability density function becomes
\[
\frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x + \sigma^2/2)^2}{2\sigma^2}}. \tag{90}
\]

### C.3 General capital

For \( \theta \) lognormal with mean 1, the expression in (20) for general capital becomes
\[
\hat{S} (\sigma) \hat{\theta}^G (\sigma)^{\left(1-\frac{\beta}{n}\right)} \frac{k}{\alpha} \left\{ 1 - \int_{\hat{\theta}^G (\sigma)}^{\infty} \left[ \theta - \hat{\theta}^G (\sigma)^{\frac{\beta}{n}} \theta^{\frac{\beta}{n}} \right] \frac{1}{\theta \sigma \sqrt{2\pi}} e^{-\frac{\left(\ln \theta + \frac{\sigma^2}{2}\right)^2}{2\sigma^2}} d\theta \right\}
\]
\[
= \hat{S} (\sigma) \hat{\theta}^G (\sigma)^{\left(1-\frac{\beta}{n}\right)} \frac{k}{\alpha} \left\{ 1 - \int_{\hat{\theta}^G (\sigma)}^{\infty} \left( 1 - \hat{\theta}^G (\sigma)^{\frac{\beta}{n}} \theta^{-\frac{\beta}{n}} \right) \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{\left(\ln \theta + \frac{\sigma^2}{2}\right)^2}{2\sigma^2}} d\theta \right\}.
\]
The results in Table 2 for systemic risk with \( \hat{S} (\sigma) \) independent of \( \sigma \) and the parameter values in Table 1, are calculated in the following way. \( \hat{\theta}^G (\sigma_H) \) for given \( \hat{\theta}^G (\sigma_L) \) can be
calculated by solving

\[
\hat{\theta}^G (\sigma_H) - \hat{\theta}^G (\sigma_H) \int_{\hat{\theta}^G (\sigma_H)}^{\infty} \left( 1 - \hat{\theta}^G (\sigma_H)^{1/2} \theta^{-1/2} \right) \frac{1}{4.07 \sigma \sqrt{2\pi}} e^{-\frac{\left( \ln \theta + \frac{4.07 \sigma^2}{2} \right)^2}{2(4.07 \sigma^2)^2}} \, d\theta
\]

\[
\hat{\theta}^G (\sigma_L) - \hat{\theta}^G (\sigma_L) \int_{\hat{\theta}^G (\sigma_L)}^{\infty} \left( 1 - \hat{\theta}^G (\sigma_L)^{1/2} \theta^{-1/2} \right) \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{\left( \ln \theta + \frac{4.07 \sigma^2}{2} \right)^2}{2\sigma^2}} \, d\theta
\]

for \( \sigma \equiv \sigma_L = 0.10 \). The change in capital can be calculated directly from (21) as

\[
\frac{\hat{K}^G (\sigma_L)}{\hat{K}^G (\sigma_H)} = \left[ \frac{\hat{\theta}^G (\sigma_H)}{\hat{\theta}^G (\sigma_L)} \right]^2.
\]

The ratio of productivity with \( \sigma_H \) to that with \( \sigma_L \) can then be calculated directly from (18) as

\[
\frac{\hat{\theta}^G (\sigma_L)}{\hat{\theta}^G (\sigma_H)} \left[ 1 - \int_{\hat{\theta}^G (\sigma_H)}^{\infty} \left( 1 - \hat{\theta}^G (\sigma_H)^{\frac{\theta}{\pi}} \theta^{-\frac{\theta}{\pi}} \right) \frac{1}{4.07 \sigma \sqrt{2\pi}} e^{-\frac{\left( \ln \theta + \frac{4.07 \sigma^2}{2} \right)^2}{2(4.07 \sigma^2)^2}} \, d\theta \right]^{-1}.
\]

\[
1 - \int_{\hat{\theta}^G (\sigma_L)}^{\infty} \left( 1 - \hat{\theta}^G (\sigma_L)^{\frac{\theta}{\pi}} \theta^{-\frac{\theta}{\pi}} \right) \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{\left( \ln \theta + \frac{4.07 \sigma^2}{2} \right)^2}{2\sigma^2}} \, d\theta.
\]

C.4 Specific capital

The results in Table 3 for idiosyncratic risk with \( \varepsilon (\sigma) \) independent of \( \sigma \) and the parameter values in Table 1 are calculated using the same procedure as for general capital but with (22) replacing (20) and (23) replacing (21).

References


54

