RAND Journal of Economics
Vol. 44, No. 1, Spring 2013
pp. 104–127

Stairway to heaven or highway to hell: Liquidity, sweat equity, and the uncertain path to ownership

R. Vijay Krishna*

Giuseppe Lopomo**

and

Curtis R. Taylor**

We study a setting in which a principal contracts with an agent to operate a firm over an infinite time horizon when the agent is liquidity constrained and privately observes the sequence of cost realizations. We formulate the principal's problem as a dynamic program in which the state variable is the agent's continuation utility, which is naturally interpreted as his equity in the firm. The optimal incentive scheme resembles what is commonly regarded as a sweat equity contract, with all rents back loaded. Payments begin when the agent effectively becomes the owner, and from this point on, all production is efficient. These features are shown to be similar to features common in real-world work-to-own franchising agreements and venture capital contracts.

1. Introduction

■ Consider the common situation in which two parties form a partnership in order to jointly operate a business enterprise. The equity or cash partner (principal) possesses capital but is unable, either due to lack of expertise or because her time and energy are best spent elsewhere, to operate the firm. By contrast, the managing partner (agent) possesses technical know-how but lacks access to the financial resources necessary to launch the enterprise or keep it afloat.

Real-world examples of this type of situation abound: retail franchising, venture capital, real estate development, newly minted professionals joining established firms. The salient features of these contractual settings are that (i) the agent is liquidity constrained and cannot purchase or finance the enterprise himself, (ii) the relationship is of a long-term nature, (iii) the agent has private access to knowledge regarding certain factors influencing profitability, and (iv) the

*University of North Carolina, Chapel Hill; rvk@unc.edu.
**Duke University; giuseppe.lopomo@duke.edu, crtaylor@duke.edu.
We wish to thank the editor (David Martimort) and two anonymous referees as well as Rachel Kranton, Philipp Sadowski, and numerous seminar participants for helpful comments and suggestions. Joe Mazur and Sergiu Ungureanu provided diligent research assistance. Krishna and Taylor also gratefully acknowledge support from the National Science Foundation (grant no. SES-1132187).

principal maintains control rights over some aspects of the operation. In this article, we provide a normative analysis of the optimal dynamic contract for the principal in a general setting possessing these characteristics.

Formally, we study an infinite-horizon discrete-time model in which the marginal cost of production evolves according to an i.i.d. process that the agent privately observes. Both principal and agent have quasilinear time-separable von Neumann-Morgenstern preferences and discount the future at the same rate. Because contracting occurs before the agent learns any private information and because allocation of risk is not germane, full efficiency could be achieved by selling the firm to the agent at its first-best expected present value. This solution, however, is assumed infeasible by supposing that the agent does not possess the requisite capital. In particular, the agent is presumed to be severely liquidity constrained and cannot experience negative cash flow in any period.[1]

These assumptions give rise to a dynamic intratemporal screening model in which the principal incentivizes the agent through both instantaneous payments as well as promised future payments. The principal also manages information rents through control of the scale of operations, that is, the output of the firm.

Our findings relate the evolution of firm dynamics to other features of the contractual relationship. In particular, we show that there is a maximal firm size, that is, a scale of operations, that is achieved if (and only if) the agent becomes a *fully vested partner* in the firm. Moreover, we show the following.

- Backloading of rents: The optimal contract incentivizes the agent exclusively via promised future payments before he becomes a fully vested partner, and exclusively via instantaneous payments if he becomes a fully vested partner.
- Easing of liquidity: Liquidity constraints are ameliorated as the firm grows and vanish completely if the agent becomes a fully vested partner.
- Heaven or hell: In the long run, with probability 1, the firm either grows to the point where the agent becomes a fully vested partner or it shrinks to the point where the principal replaces him.

In fact, our main results are best summarized collectively as a theory of *sweat equity*, wherein the agent works for the principal without receiving rents until the scale of the firm and his equity position grow to the level of ownership or shrink to the point where he is replaced. We summarize evidence below in Section 8 showing that these characteristics of the optimal dynamic contract have close parallels in real-world work-to-own franchise programs and venture capital covenants. They also resonate with features of contracts involving newly hired members of professional partnerships: a doctor joining a medical practice, an attorney joining a law firm, an economist joining a consulting group, and so forth.

In the next section, we briefly survey the relevant literature. We introduce the model formally in Section 3, and describe the recursive approach we employ in Section 4, where we also establish basic properties of the principal's value function, prove that the optimal contract backloads all rents, and derive a simplified version of the principal's contract design problem that is more amenable to analysis. In Section 5, we derive necessary and sufficient first-order conditions characterizing the solution to the principal's problem. We also derive an expression for the critical level of equity at which the agent achieves a vested ownership stake in the firm. In Section 6, we describe the short- and long-run dynamics induced by the optimal contract. The Lagrange multipliers associated with the liquidity constraints or, more precisely, their sum, can be interpreted as the marginal social cost of illiquidity. This, and other issues, related to various

[1] In hidden action models, a restriction that the agent not be paid negative wages following production is typically called a "limited liability constraint." We wish to distinguish this from the "liquidity constraints" in our hidden information setting under which the agent possesses known contractible wealth and must be allocated the requisite operating capital prior to production.

levels of ownership, path dependence of the optimal contract, and versions of the model where the agent has fixed costs of production, or where the principal can fire the agent, are analyzed in Section 7. Section 8 contains the applications of our model mentioned above to work-to-own franchising programs and to venture capital covenants, and some concluding remarks appear in Section 9. Formal proofs and some purely technical results are relegated to the Appendix.

2. Related literature

■ This article belongs to a line of research in dynamic contracting initiated by Baron and Besanko (1984). That study investigates optimal dynamic regulation in a setting where the regulator (principal) possesses full power of commitment and the firm operater (agent) privately observes realizations of marginal cost over time. Because the agent in Baron and Besanko (1984) is not liquidity constrained, the first best is obtained from the second period on in the i.i.d. version of their model, and the only distortion arises in the first period due to the agent's *ex ante* private information about the initial state of marginal cost. This obviates the need for studying more than two periods or using recursive methods. In fact, the agent in our model does not possess any *ex ante* private information, and it therefore would be possible to implement the first best in every period by *selling him the firm* if he were not liquidity constrained. For a discussion of how the optimal contract is altered in our setting when the agent possesses positive initial wealth, see Section 7, especially Corollary 1 and the ensuing remarks.

More generally, our article contributes to a growing literature on optimal dynamic incentive schemes spanning a diverse set of research areas including social insurance (e.g., Fernandes and Phelan, 2000), taxation (e.g., Albanesi and Sleet, 2006), and executive compensation (e.g., Sannikov, 2008). As is common in this body of work, we employ the recursive techniques for analyzing dynamic agency problems pioneered by Green (1987) (who studied social insurance), Spear and Srivastava (1987) (who studied dynamic moral hazard), and especially Thomas and Worrall (1990) (who examined income smoothing under private information), in which shocks are i.i.d. over time and the state variable is taken to be the expected present value of the agent's utility under the continuation contract. A critical difference between Thomas and Worrall (1990) and the setting considered here is the presence of liquidity constraints. In Thomas and Worrall (1990), the agent's utility is unbounded below, so even though his instantaneous consumption must be nonnegative, he can have arbitrarily low consumption utility in any period. The agent's liquidity constraints in our model preclude this possibility.

Of particular relevance is the recent literature on optimal financial contracting in the face of moral hazard. Specifically, Quadrini (2004), Clementi and Hopenhayn (2006), DeMarzo and Sannikov (2006), DeMarzo and Fishman, 2007, and Biais et al. (2007) study various dynamic incarnations of the celebrated *cash flow diversion* (CFD) model.[2] Roughly, DeMarzo and Fishman (2007) explore optimal financial contracting in a general finite-horizon CFD model that DeMarzo and Sannikov (2006) formulate in continuous time with an infinite horizon, and Biais et al. (2007) provide a model bridging the two environments. Clementi and Hopenhayn (2006) study optimal investment and capital structure in a discrete-time infinite-horizon model, and Quadrini (2004) derives the optimal renegotiation-proof contract in a similar environment.

As in our setting, all of these papers assume a risk-neutral but liquidity constrained agent and a risk-neutral wealthy principal. There are, however, several key differences between the environment we study and the one analyzed in the dynamic CFD literature. First and foremost, the underlying problem facing the principal in CFD models involves moral hazard in which the agent must be given incentives either not to expropriate privately observed cash flows for his personal use or to privately exert personally costly effort. (As DeMarzo and Fishman, 2007 demonstrate, these two situations are formally equivalent.) In particular, the information privately observed by the agent in the CFD models is of no operational use to the principal—she always

[2] See Bolton and Scharfstein (1990) for a canonical two-period CFD model.

wants him either to not divert funds or to work hard, depending on the context of the model. Hence, her contemporaneous policy decision of how much to invest is not sensitive to the agent's private information about his action (regarding the amount of cash he expropriated or his effort choice).

Our focus, by contrast, is not on optimal investment dynamics or capital structure, but on the day-to-day operation of the firm. The principal in our model wishes to tailor her contemporaneous policy decision of how much to produce to the agent's private information regarding the marginal cost of operation. Thus, ours is a dynamic model of intratemporal screening that cannot properly be viewed as a setting of moral hazard.[3] To see this plainly, note that in the CFD models each value of the state variable gives rise to a distinct level of optimal investment, whereas in our setting each value of the state variable gives rise to a menu of output levels from which the agent must be given incentives to select the optimal one. Although our investigation clearly touches on issues of corporate finance, our focus is rooted in questions of procurement and monopolistic screening that are more readily identified with industrial organization.[4]

Clearly, some of our results do have parallels in the CFD literature. For instance, we discover a bang-bang property of an optimal contract common among the CFD papers under which the agent is incentivized only through adjustments in his future utility up to a threshold, after which he is incentivized with cash payments. The CFD papers naturally interpret this as optimal financial structure; for example, debt must be retired before dividends can be paid. We, on the other hand, interpret the bang-bang property of the optimal incentive scheme as a sweat equity contract under which the agent works for the principal until he is fired or earns a permanent ownership stake in the firm. However, in both the CFD models as well as in ours, the backloading of rents is a consequence of the twin assumptions that the agent is risk neutral *and* liquidity constrained.

Questions of interpretation and implementation aside, a number of our results have no counterpart in the CFD literature. For instance, we show that there is an endogenously determined positive level of equity that the principal optimally grants the agent at the beginning of the contract. We also characterize the production mandates used to control information rents, including the familiar result from static mechanism design of *no distortion at the top*, which holds in our setting for all values of the state.

Defining deadweight loss to be the difference between the first-best value of the firm and its value (principal's share plus agent's share) at any state allows us to relate the social cost of illiquidity to the analytical measure of the *price* of the constraints. Namely, deadweight loss under the contract is the integral of the sum of the Lagrange multipliers between the current state and the state at which firm value is maximized (where all the multipliers drop to zero and the agent achieves a vested ownership stake in the firm).

In addition to this study, there are several other recent investigations of screening mechanisms in dynamic environments. For instance, Bergemann and Välimäki (2010) introduce and analyze a dynamic version of the Vickrey-Clarke-Groves (VCG) pivot mechanism. (In a similar vein, see Athey and Segal, 2012 and Covallo, 2008.) In a recent paper, Pavan, Segal, and Toikka (2012) study dynamic screening in a setting in which the distribution of types may be nonstationary and agents' payoffs need not be time separable. These authors derive a generalization of the envelope formula of Mirrlees (1971) for incentive-compatible static mechanisms and use this to compute a dynamic representation for virtual surplus in the case of quasilinear preferences. Although their analysis is illuminating, the generality of their model prohibits use of the recursive methods that are the lynchpins of our study. Moreover, Pavan, Segal, and Toikka (2012) do not address directly the question of contracting for ownership in the face of liquidity constraints that is the focus of our investigation.

[3] The conditions under which *ex post* hidden information, as in the CFD models, is analogous to moral hazard are articulated in Milgrom (1987).

[4] See, for example, Laffont and Martimort (2002).

Boleslavsky and Said (forthcoming) explore a dynamic selling mechanism in which a consumer possesses both permanent private information about his propensity to have high or low taste shocks and transitory private information about his current (conditionally independent) shock. The optimal contract in Boleslavsky and Said's model exhibits a type of *immiseration*, in the sense that after a sufficiently long time horizon, the supplier will eventually refuse to serve the consumer.

Battaglini (2005) investigates a dynamic selling procedure in a model where a consumer's taste parameter follows a two-state (high or low) Markov process. The consumer has private information about the initial state of the process as well as subsequent states. For an initial string of reported low-demand realizations, the consumer's allocation is distorted down from the efficient level, but the distortions diminish after each report in the string. Moreover, the first time the consumer reports high demand, the contract calls for efficient output for both types from that point forward. These dynamics contrast sharply with our findings in which each *bad* report leaves the agent in a worse position and efficiency obtains only after a sufficiently long string of *good* reports. In analyzing the process of ownership acquisition, Battaglini (2005) emphasizes the role of initial and persistent private information, whereas we focus on the complementary part played by illiquidity and transitory private information.

3. The model

■ A principal contracts with an agent to produce output in each period $t = 0, 1, 2, \cdots$. Both parties are risk neutral, have time-separable preferences, and have a common discount factor $\delta \in (0, 1)$. If the agent produces q units in a given period, then a contractually verifiable monetary benefit (revenue) $R(q)$ is generated, where $R : \mathbb{R}_+ \to \mathbb{R}_+$ is twice continuously differentiable, strictly concave, and $R(0) = 0$.[5]

The principal is *not* a bank that simply lends the agent capital. Instead, we suppose the firm possesses some market power, which leads naturally to the assumption $R'' < 0$, and which we associate with control of specialized assets such as brand recognition, an exclusive location, a proprietary business formula, or physical capital. The principal generally retains ownership of these assets, although they may be transferred to the agent under certain situations, as we discuss in Section 7 below.

The agent's cost of producing q units of output in a given period is θq, where $\theta \in \Theta := \{\theta_1, \theta_2\}$[6] and $0 < \theta_1 < \theta_2 < \infty$.[7] We will frequently refer to $i, j \in \{1, 2\}$ rather than saying $\theta_i, \theta_j \in \Theta$. The cost parameter θ is drawn independently in each period with $\Pr\{\theta = \theta_i\} := f_i > 0$ for all $i = 1, 2$.

To ensure an interior solution to the contracting problem, we assume

$$R'(0) = \infty \qquad [\text{MR}_0]$$

and

$$\lim_{q \to \infty} R'(q) < \theta_1.$$

Then, implicitly define the first-best output levels by $R'(q_i^*) = \theta_i$ for all $i \in \{1, 2\}$. For future reference, note that $\infty > q_1^* > q_2^* > 0$; that is, first-best output is monotone decreasing in type and is always finite. As always, the agent can leave at any moment in time, to an outside option

[5] As long as revenue is contractible, it does not matter whether it accrues directly to the principal or the agent. We assume the former case in the text.

[6] For ease of exposition, we assume below that marginal cost can take on one of only two values. Our results extend to an environment with an arbitrary (but finite) number of possible marginal cost realizations.

[7] Consider the seemingly more general specification in which output is $x \geq 0$; concave revenue is $B(x)$; and increasing convex cost is $\theta C(x)$. This is equivalent to the specification given in the text under the change of variables $q := C(x)$ and $R(q) := B(C^{-1}(q))$. Moreover, our results also hold under an alternative specification in which revenue is $\theta B(x)$, which is observed only by the agent, and cost is $C(x)$, which is contractually verifiable.

worth 0 utiles.[8] There are two crucial sources of friction in the model. First, the agent is liquidity constrained and cannot incur a negative cash flow in any period. Second, the realization of the cost parameter θ in each period is observed only by the agent.[9] If either of these conditions were relaxed, it would be possible to implement the first-best outcome. For instance, if θ was observed publicly in each period, the principal could simply write a forcing contract that dictated the efficient level of output q_i^* and compensated the agent for his actual costs $\theta_i q_i^*$. If, on the other hand, the agent possessed sufficient liquid resources, he could purchase the franchise from the principal at the outset for its first-best expected present value,

$$v^{\mathrm{FB}} := \frac{1}{1-\delta}\left[f_1\left[R(q_1^*) - \theta_1 q_1^*\right] + f_2\left[R(q_2^*) - \theta_2 q_2^*\right]\right] \qquad \text{[FB]}$$

in which case there would be no residual incentive problem. Hence, it is the combination of illiquidity and private information that links the present with the future, giving rise to a non-trivial dynamic contracting problem.

The timing runs as follows. At the beginning of the game the principal offers the agent an infinite-horizon contract that he may accept or reject. If he rejects, then the game ends and each party receives a reservation payoff of zero. If the agent accepts the principal's offer, the contract is executed.

4. Contract design

■ When designing an optimal contract, the Revelation Principle implies that the principal may restrict attention to incentive-compatible direct mechanisms. Moreover, it is well known (see, e.g., Thomas and Worrall, 1990) that in the setting under study, she also may restrict attention to recursive mechanisms in which the state variable is the agent's lifetime promised expected utility under the contract, denoted by v. For reasons discussed below, we refer to v as the agent's equity (or sweat equity) in the firm. Hence, if the agent's current equity is v and he reports θ_i, then the contract specifies the amount of output he is to produce $q_i(v)$, the amount he is to be compensated by the principal $m_i(v)$, and his level of equity starting next period $w_i(v)$. (To ease notation, we frequently suppress dependence of the contractual terms on v.)

In fact, it is convenient, both notationally and conceptually, to define the agent's instantaneous rent as $u_i := m_i - \theta_i q_i$ and to consider contracts of the form (u, q, w) rather than (m, q, w). We now present the contractual constraints under this formulation.

Promise keeping: The *promise-keeping* constraint that the contract must obey is written

$$f_1(u_1 + \delta w_1) + f_2(u_2 + \delta w_2) = v. \qquad \text{[PK]}$$

The agent's lifetime expected payoff, v, is composed of his expected payoff in the current period, u_i, and his expected continuation payoff, δw_i.

Incentives: The set of incentive constraints is

$$u_i + \delta w_i \geq u_j + \delta w_j + (\theta_j - \theta_i) q_j \qquad [\mathrm{C}_{ij}]$$

for all $i, j = 1, 2$. Simply put, the agent's payoff from truthfully reporting i must be no less than what he could obtain by reporting j, namely the truthful payoff from reporting j plus the applicable cost difference.

Liquidity: The agent's liquidity constraints are written

$$u_i \geq 0 \qquad [\mathrm{L}_i']$$

[8] In fact, the agent's individual rationality constraint never binds (as we discuss below), so the analysis is unaltered whether we assume he has the option to quit in any period or is committed to work for the principal indefinitely.

[9] Implicit in this assumption is that the agent can divert any excess funds to his own consumption without being observed by the principal.

for all $i \in \{1, 2\}$. That is, when the agent reports truthfully, the monetary transfer he receives from the principal, m_i, must cover his production costs $\theta_i q_i$.[10] As written, the liquidity constraints do not permit wealth accumulation by the agent. In other words, he has no method for saving any positive rents $m_i - \theta_i q_i > 0$ to ease liquidity in the future. Although this appears to be a restrictive assumption, it is actually completely innocuous because the principal saves (and dissaves) on the agent's behalf by adjusting his equity v. Of course, the contract could specify that the agent save any positive rents in a verifiable bank account, but this would be functionally equivalent to using equity adjustments and operationally more cumbersome.[11]

Participation: The continuation utility w_i is the sum of expected future rents, and because instantaneous rents to the agent can never be negative, it follows that we must include feasibility constraints that require $w_i \geq 0$ for all $i = 1, 2$. Thus, promise keeping [PK] implies that the agent's lifetime expected utility v is always nonnegative, and the participation constraint that the contract initially offer him nonnegative lifetime utility may be ignored.

The following proposition shows that the principal's problem can be written as a dynamic program and establishes that an optimal contract exists by virtue of being the corresponding policy function.

Theorem 1. The principal's discounted expected utility under an optimal contract, (u, q, w), is represented by a unique, concave, and continuously differentiable function $P : \mathbb{R}_+ \to \mathbb{R}$ that satisfies

$$P(v) = \max_{(u,q,w)} \sum_{i=1,2} f_i[(R(q_i) - \theta_i q_i) - u_i + \delta P(w_i)] \qquad \text{[VF}']$$

subject to promise keeping [PK], incentive compatibility [C_{ij}], liquidity [L_i'], and feasibility $q_i \geq 0$ and $w_i \geq 0$ for $i = 1, 2$. Moreover, there exists $v^* \in (0, \infty)$ such that $P'(v) > -1$ for $0 \leq v < v^*$ and $P'(v) = -1$ for $v \geq v^*$, and $P'(0) = \infty$.

Theorem 1 provides some clues to the structure of an optimal contract, in particular ([MR_0]), namely the assumption that $R'(0) = \infty$ ensures $P'(0) = \infty$. In other words, the principal's payoff is initially increasing in the agent's equity. This, along with the facts that $P'(v) = -1$ for $v \geq v^*$ and that $P(v)$ is concave, implies that there exists a level of equity $v^0 \in (0, v^*)$ satisfying $P'(v^0) = 0$ at which the principal's discounted expected payoff is maximized (see Figure 1). This is the level of equity at which the principal initially stakes the agent upon signing the contract.

Note, however, that the social surplus (i.e., firm value) $P(v) + v$ is maximized at any $v \geq v^*$.[12] In other words, the value of the contractual relationship continues to grow until $v = v^*$. The following result shows that any optimal contract must have a *bang-bang* structure in the sense that all rents are backloaded.

Proposition 1. For any optimal contract (u, q, w), incentives are provided purely through adjustments in the agent's equity whenever his stake in the firm is sufficiently low—in particular,

$$w_i(v) < v^* \text{ implies } u_i(v) = 0.$$

Moreover, there exists a *maximal rent* (optimal) contract in which incentives are provided purely through payment of rents if the agent's stake in the franchise is sufficiently high—specifically, we have $w_i(v) \leq v^*$ for all $v \geq 0$.

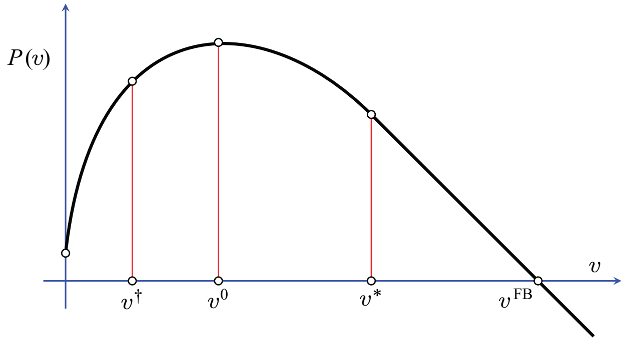
[10] Strictly speaking, the full set of liquidity constraints is $m_j - \theta_i q_j \geq 0$ for all $i, j \in \Theta$. That is, the agent cannot spend more money than the principal gives him in any period, whether or not he reports truthfully. However, because we are imposing incentive compatibility [C_{ij}], there is no loss in generality from restricting attention to the subset of liquidity constraints associated with truthful reporting.

[11] See Edmans et al. (2012) for a novel use of "incentive accounts" in the context of executive compensation.

[12] This follows because $P(v) + v$ is increasing, concave, continuously differentiable, and has the derivative $P'(v) + 1$, which is strictly positive for all $v < v^*$, and is 0 for all $v \geq v^*$.

FIGURE 1

PRINCIPAL'S VALUE FUNCTION



It is easy to see that in any maximal rent contract, if $u_i(v) > 0$, then $w_i(v) = v^*$. Proposition 1 underpins the interpretation of the optimal incentive scheme as a sweat equity contract. For $v < v^*$, if it is the case that $w_i(v) < v^*$, that is, the agent does not reach $v = v^*$ in the next period, it must be that the agent earns no instantaneous rents in state θ_i but instead is incentivized purely through adjustments to his equity position. Once $v = v^*$, however, the agent, as we discuss below, achieves a permanent ownership stake in the firm and earns nonnegative instantaneous rents from that point forward. The proposition also establishes the existence of a useful class of contracts, namely maximal rent contracts, which have the property that they deliver rents to the agent as quickly as possible. That is, instead of making promises of future utility in excess of v^* (where optimal), a maximal rent contract makes future utility promises at the level v^* and instead delivers instantaneous rents. (Note that maximal rent contracts are always optimal.)

The intuition behind this result is that in the dynamic setting, the principal can induce truth telling via two instruments: instantaneous rent u_i and continuation utility w_i, the latter being the sum of expected future rents. The problem with providing incentives through current rent, u_i, is that this must be nonnegative due to the liquidity constraints; thus, the agent can only be rewarded and never penalized. Moreover, any instantaneous rent awarded to the agent is spent outside the contractual relationship and therefore does not benefit the principal. If, however, the principal chooses to provide the necessary incentives through continuation payoffs w_i, then she can reward the agent by adjusting his equity up or penalize him by adjusting it down. Hence, providing incentives through continuation utility has two advantages: it keeps payments inside the relationship and it permits penalties. Once $v = v^*$, liquidity constraints no longer bind (i.e., penalties become irrelevant), and the principal can provide the requisite incentives purely through instantaneous rents.

To aid with the analysis and to obtain a sharper characterization of an optimal contract, it is helpful to reformulate the principal's program in a simpler way (with fewer constraints and choice variables). To this end, first consider the following definition.

Definition 1 (*Monotonicity in type*). Output is said to be *monotonic in type* if for all $v \geq 0$,

$$q_1(v) \geq q_2(v). \qquad [\mathrm{M}_1]$$

Analogous definitions apply for rent $u_1(v)$ and promised utility $w_1(v)$.

The inequality in [M$_1$] requires output to be monotonic in type for each $v \geq 0$, which generalizes the monotonicity requirement encountered in static mechanism design. In the static setting, this inequality is often referred to as an *implementability condition*. Analogous to the static setting, allocations that do not satisfy [M$_1$] for some $v \geq 0$ are not incentive compatible. Thus, the monotonicity (of output) in type is a necessary condition for incentive compatibility.

Next, consider the binding version of the upward adjacent incentive constraints that say the agent of type θ_1 must be indifferent between reporting his true marginal cost and θ_2:

$$u_1 + \delta w_1 = u_2 + \delta w_2 + \Delta q_2, \qquad [\mathrm{C}_1]$$

where $\Delta := \theta_2 - \theta_1$.

The following lemma establishes a result familiar from static mechanism design that the pair of incentive constraints [C_{ij}] may be replaced by [M_1] and [C_1].

Lemma 1. If output is monotonic in type [M_1] and the upward adjacent incentive constraint [C_1] binds, then both incentive constraints [C_{ij}] are satisfied. Moreover, there exists a maximal rent contract (u, q, w) (which is optimal) in which [M_1] and [C_1] hold, and in any such contract, instantaneous rent and promised utility are also monotonic in type.[13]

Next, the following lemma uses [PK] and [C_1] to derive a key expression for the agent's current payoff.

Lemma 2. In any optimal contract (u, q, w), the agent's payoff satisfies

$$u_i + \delta w_i = v - f_1 \Delta q_2 + (2 - i)\Delta q_2 \qquad [\mathrm{U}_i]$$

for $i = 1, 2$. Moreover, if the optimal contract (u, q, w) satisfies [U_i] for $i = 1, 2$, then (u, q, w) also satisfies [PK] and [C_1] for $i = 1, 2$.

Equation [U_i] says that the current payoff to the agent when he is type i is his promised expected level of equity from the prior period (first term on the right) minus his expected information rent (second term) plus his realized information rent (third term). Notice that the realized information rent can be nonzero only if $i = 1$, that is, only the low-cost agent can get any information rent.

The equation [U_i], which implies [PK] and [C_1], can be used to eliminate instantaneous rents, u_i, from the principal's program [VF$'$]. Specifically, the liquidity constraints [L_i'], requiring $u_i \geq 0$, can be recast as

$$f_1 \Delta q_2 - (2 - i)\Delta q_2 + \delta w_i \leq v \qquad [\mathrm{L}_i]$$

for $i = 1, 2$. Using this version of the liquidity constraints and substituting [PK] directly into the principal's objective yields the following intuitive reformulation of the contract design program.

Theorem 2. The principal's value function $P : \mathbb{R}_+ \to \mathbb{R}$ is a solution to the following relaxed program:

$$P(v) = \max_{(q,w)} \sum_{i=1,2} f_i\big[\big(R(q_i) - \theta_i q_i\big) + \delta\big(P(w_i) + w_i\big)\big] - v, \qquad [\mathrm{VF}]$$

subject to monotonicity in output [M_1], liquidity [L_i], and feasibility $q_2 \geq 0$ and $w_2 \geq 0$. Moreover, there is a solution to this program that is a maximal rent contract in which $u_i(v)$ and $w_i(v)$ are monotonic in type. This optimal contract (u, q, w) is unique and continuous in v.

This version of the principal's program is substantially simpler than the one presented in Theorem 1, and it also has an intuitive interpretation. The term $\sum_i f_i[R(q_i) - \theta_i q_i]$ is simply expected instantaneous social surplus (current profit), whereas the term $\sum_i f_i[w_i + P(w_i)]$ is

[13] We note that there exist optimal contracts that are not maximal rent contracts. Consequently, such contracts can have instantaneous rents or promised utilities that are not monotonic in type. To see this, recall from Proposition 1 that there always exists a maximal rent contract. Suppose v is such that $u_i(v) > 0$ and $w_i(v) = v^*$ for some v. Now form a new contract by reducing $u_i(v)$ by ε and increasing $w_i(v)$ by ε/δ. This contract is optimal because it leaves the principal's utility unchanged and also satisfies all the other constraints, but is clearly not monotonic in type in promised utility.

the expected continuation surplus (future profit). Also, v is just the sum of present and future expected rents owed to the agent. Therefore, $P(v)$ is just the dynamic analogue of the objective in the static problem, wherein the principal wants to maximize expected lifetime social surplus (i.e., the value of the firm) net of any expected lifetime information rents.

Note that in the absence of liquidity constraints, the first-order condition for q_i would be $R'(q_i) = \theta_i$, implying $q_i = q_i^*$, and the first-order condition for w_i would be $P'(w_i) + 1 = 0$, implying $w_i = v^*$. Moreover, the principal would set the agent's initial equity at $v = 0$ to ensure his participation. But then [U$_i$] and Lemma 3 in the next section give the agent's first-period rents for high cost realizations as

$$u_2 = -v^*,$$

which is negative. Hence, it is the presence of the binding liquidity constraints that causes the principal to distort output levels away from first best. We investigate these distortions in the next section.

5. Optimal contracts

■ In the previous section, we noted that we can formulate the principal's problem as a dynamic program with only liquidity, implementability, and feasibility constraints. For any value of v, the optimal value of $(q(v), w(v))$ is the solution to a concave programming problem, and hence first-order conditions are both necessary and sufficient. Let λ_i be the Lagrange multiplier associated with the liquidity constraint [L$_i$]. Because $P'(0) = \infty$, we will ignore the constraint $w_2 \geq 0$ whenever $v > 0$. For the moment, let us also ignore the monotonicity constraint [M$_1$]. (Proposition 2 below shows that this is without loss of generality.)

The first-order condition for q_1 is simply $R'(q_1) = \theta_1$, that is, $q_1 = q_1^*$. This is the familiar result from static monopolistic screening that there is *no distortion at the top*, which holds here for all $v \geq 0$. The first-order condition for q_2 is

$$R'(q_2) - \theta_2 = \frac{\Delta}{f_2}\left[f_1\Lambda - \lambda_1\right], \qquad \text{[FO}q_2\text{]}$$

where $\Lambda = \lambda_1 + \lambda_2$.

By Theorem 1, we know that the value function P is continuously differentiable. Therefore, the first-order condition for w_i is

$$P'(w_i) = -1 + \frac{\lambda_i}{f_i}. \qquad \text{[FO}w_i\text{]}$$

Finally, the envelope condition is

$$P'(v) = -1 + \Lambda. \qquad \text{[Env]}$$

The first-order conditions permit calculation of v^* as presented in the following lemma.

Lemma 3. The critical level of equity is

$$v^* = \frac{1}{1-\delta} f_1 \Delta q_2^*. \qquad \text{[Vest]}$$

In words, v^* is the present value of receiving expected information rents from efficient production (that is, output without distortions) in perpetuity. Moreover, because $P'(v) = -1$ for all $v \geq v^*$, it must be that $\lambda_i(v) = 0$ for all i, $v \geq v^*$. That is, v^* is the lowest equity level at which none of the agent's liquidity constraints bind and correspondingly the lowest equity level at which no production levels are distorted.

The following proposition establishes a result familiar from static mechanism design, namely that for $v < v^*$, the principal distorts output levels down (and never up) in order to control information rents.

Proposition 2. In the maximal rent optimal contract (u, q, w), instantaneous rent u, output q, and continuation utility w are all monotone in type for all $v \geq 0$. They satisfy the (necessary and sufficient) first-order conditions [FOq_2], [FOw_i], the usual complementary slackness conditions, and the envelope condition [Env]. The optimal contract is also continuous in v. Moreover, the agent never produces more than first-best output; in fact, $q_2(v) \leq q_2^*$ and $q_1(v) = q_1^*$ for all $v \in [0, v^*]$, and hence [M$_1$] always holds.

The proposition describes the properties of the maximal rent optimal contract and can be viewed as providing a solution to the problem in Theorem 2. In particular, it shows that the contract (u, q, w) is monotone in type for all $v \geq 0$ (this is just Lemma 1) and is continuous in v.

The proposition also provides an upper bound for each q_i, but not a lower bound (other than requiring $q_2 \geq 0$). In particular, at $v = 0$, the contract directs the agent to produce positive output only in the low-cost state (this is Lemma A1 in the Appendix): $q_1(0) = q_1^*$ and $q_2(0) = 0$. At low levels of v, the agent's liquidity constraints are tight and the contract imposes stringent output restrictions along with correspondingly low levels of promised future utility. As v increases, output restrictions are relaxed until $v = v^*$, at which point the contract calls for efficient production for both cost realizations: $q_i(v^*) = q_i^*$ for $i = 1, 2$. The agent's promised future utility levels also rise in sweat equity. At $v = 0$, he *never* receives any rents, implying $w_i(0) = 0$ for $i = 1, 2$. Again, as v increases, promised future utility levels rise until $v = v^*$, when the agent becomes a *vested partner* achieving a permanent ownership stake, with $w_i(v^*) = v^*$ for $i = 1, 2$. As we prove in the next section, if the agent makes a favorable report at this point, he is rewarded with higher equity. This relaxes his liquidity constraints, ultimately leading to less-strict output controls and still-higher levels of promised future utility.[14]

It is worth noting that although the contract is *defined* for the case where $v = 0$, the Inada assumption [MR$_0$] ensures that $P'(0) = \infty$ so that for all $v > 0$, $w_i(v) > 0$ for $i = 1, 2$. Thus, after any finite sample path, there will never be complete shutdown in the high-cost state. To be sure, this observation depends crucially on [MR$_0$]. Indeed, it can be shown that if $R'(0) < \infty$, then $P'(0) < \infty$, so that for $v < v^*$, output in the high-cost state will be shut down permanently after a sufficiently long string of high-cost reports (see Proposition 3 in the next section).

6. Dynamics

■ We next derive both short- and long-run dynamics of the contractual relationship. Our first observation follows directly from summing the first-order conditions for w_i [FOw_i] and substituting from the envelope condition [Env].

Lemma 4. An optimal contract induces a process P' that is a martingale; that is,

$$P'(v) = f_1 P'(w_1) + f_2 P'(w_2).$$

To see this, consider an increase in v by one unit. This can be achieved by increasing all the w_is by $1/\delta$. The cost of this to the principal is $\sum_{i=1,2} f_i\left[1 + P'(w_i)\right] - 1$. By the envelope theorem, this is locally optimal, and hence is equal to $P'(v)$.

An important consequence of the martingale property of P' is that a shock of $\theta = \theta_1$ is necessarily *good*, in the sense that the continuation value of sweat equity $w_1 > v$, whereas a shock of $\theta = \theta_2$ is unambiguously *bad*, $w_2 < v$. Formally, we have the following.

Proposition 3. In an optimal contract, for all $v \in (0, v^*)$, we have $P'(w_2) > P'(v) > P'(w_1)$. Moreover, $w_1(v) > v > w_2(v)$.

This captures the short-run consequences of good and bad shocks. To see the intuition, suppose, for simplicity, that P is strictly concave on $(0, v^*)$. Because P' is a martingale, if the

[14] Strictly speaking, this discussion pertains to monotonicity when moving discretely from $v = 0$ to $v = v^*$. The functions $q_i(v)$, $w_i(v)$, and $\lambda_i(v)$ are continuous for $0 \leq v \leq v^*$, but we have been unable to prove that they are monotone at every point in this range (although we suspect this to be true for suitable specifications of R).

proposition were not true, it would follow that $P'(w_i) = P'(v)$ for $i = 1, 2$, which implies (if P is strictly concave) that $w_i(v) = v < v^*$ for $i = 1, 2$. But Proposition 1 also requires that for such a v, $u_i(v) = 0$, which violates promise keeping [PK], and by incentive compatibility would require that $q_2 = 0$. Therefore, incentive compatibility and promise keeping force the principal to spread out the agent's continuation utilities, rewarding him for favorable (low-) cost reports and penalizing him for unfavorable (high-) cost ones. Although we are unable to establish that P is strictly concave, the proof can be extended to the case where P is merely concave (see the Appendix).

We are now in a position to describe the long-run properties of the optimal contract. Recall that the agent becomes a *vested partner* if his equity level reaches v^*.

Theorem 3. In a maximal rent optimal contract, the agent becomes a vested partner with probability 1. In particular, the martingale P' converges almost surely to $P'_\infty = -1$.

From the martingale convergence theorem, it follows that P' must converge, almost surely, to an integrable random variable P'_∞. The theorem establishes that along almost all sample paths, this limit must be -1. That P' cannot settle down to a finite limit greater than -1 follows from Proposition 3 above and the continuity of the contract in v.

Proposition 3 says that for $v \in (0, v^*)$, the agent is rewarded for reporting a low cost and one in which he is penalized for reporting a high cost. Because $P'(0) = \infty$, an arbitrarily long string of penalties never pushes the agent's continuation utility into the absorbing state at $v = 0$. An arbitrarily long string of rewards, however, will eventually drive his continuation utility into the absorbing state at $v = v^*$. Theorem 3 says that with probability 1, the agent will eventually experience a sufficiently long sequence of rewards to become a vested partner in the firm.

7. Discussion and extensions

□ **The social cost of illiquidity.** Define firm value, or what is the same in this instance, social surplus, under an optimal contract as $S(v) := P(v) + v$. By Theorem 1, $S(v)$ is an increasing, concave, and continuously differentiable function. In particular, we know that $S(v)$ is strictly increasing on $[0, v^*)$, and $S(v) = v^{\mathrm{FB}} = \frac{1}{1-\delta}\sum_{i=1,2} f_i[R(q_i^*) - \theta_i q_i^*]$ for all $v \geq v^*$.[15] Moreover, by the envelope condition [Env], we see that $S'(v) = P'(v) + 1 = \Lambda(v)$. Therefore, Λ measures the marginal social cost of illiquidity (which is decreasing in v). Hence, for any $v < v^*$, the *deadweight loss* generated by an optimal contract that starts with an initial promise of v utiles to the agent is

$$S(v^*) - S(v) = v^{\mathrm{FB}} - S(v) = \int_v^{v^*} \Lambda(x)\,\mathrm{d}x.$$

This cost represents the loss in social surplus arising from the output restrictions the principal imposes to control information rents. As the agent's stake in the enterprise grows, his liquidity constraints become less stringent and output restrictions are *relaxed*. At $v = v^*$, all output levels are first best and deadweight loss is consequently nil.

□ **The path to ownership.** When exploring firm ownership, it is useful to distinguish between two paradigms. One school of thought, due to Berle and Means (1968), defines ownership as residual claims over the cash flows of the firm. A second school identifies ownership of the firm with control rights over productive assets. In our model, the formal contract between the principal and agent is purely financial, identifying firm ownership with the Berle-Means interpretation. It is possible, however, to include an option for the principal to transfer the productive assets to

[15] To see this, recall that for $i = 1, 2$, $q_i(v^*) = q_i^*$ and $w_i(v^*) = v^*$. Substitution into [VF] then yields $P(v^*) + v^* = \frac{1}{1-\delta}\sum_{i=1,2} f_i\big[R(q_i^*) - \theta_i q_i^*\big] = v^{\mathrm{FB}}$ and, hence, $S(v^*) = v^{\mathrm{FB}}$. Moreover, $P(v)$ is continuous and $P'(v) = -1$ for $v \geq v^*$, so $S(v) = S(v^*)$ for $v > v^*$. It also follows from this that $P(v^*) > 0$ if, and only if, $v^* < v^{\mathrm{FB}}$, which always holds (see Lemma 5).

the agent (i.e., to literally sell the firm) under an optimal contract. We begin with the following observation.

Lemma 5. The expected present value of information rents is less than the first-best value of the firm: $v^* < v^{\mathrm{FB}}$.

This lemma says that at equity level v^*, the principal *owes* the agent less than the first-best value of the firm. In particular, at v^* the principal still retains a positive ownership stake, $P(v^*) = v^{\mathrm{FB}} - v^* > 0$.

Recall that under a maximal rent optimal contract, the agent's equity is capped at v^* and he is incentivized with cash from that point forward. However, once the agent attains equity of v^*, all output distortions are eliminated, and both the principal and agent are indifferent between providing incentives with cash or further equity adjustments. Consider a contract under which the agent continues to be incentivized with sweat equity until $v = v^{\mathrm{FB}}$. Indeed, once the agent reaches v^*, he will move monotonically to v^{FB} because (as is easily seen from [U_i]) $w_2(v) = \frac{v-(1-\delta)v^*}{\delta} \geq v$ for $v \geq v^*$. Once $v = v^{\mathrm{FB}}$, the principal owes the agent expected cash flows equal to the first-best value of the firm, namely v^{FB}, at which point $P(v^{\mathrm{FB}}) = 0$. The principal can now simply transfer control of the productive assets to the agent and terminate the contractual relationship.

We conclude the discussion of ownership with a few remarks concerning the situation in which the agent has positive initial wealth. Theorem 1 implies the following result.

Corollary 1. Suppose the agent has initial liquid wealth of $y > 0$.

(a) If $y \leq v^0$, then the agent surrenders y to the principal and receives initial equity v^0. Initial welfare is $S(v^0) < v^{\mathrm{FB}}$.
(b) If $v^0 < y < v^*$, then the agent surrenders y to the principal and receives initial equity y. Initial welfare is $S(y) \in (S(v^0), v^{\mathrm{FB}})$.
(c) If $y \geq v^*$, then the agent surrenders at least v^* and receives a like amount in initial equity. Initial welfare is v^{FB}.

If the agent possesses initial liquid wealth of $y > 0$, then the principal, who has all the bargaining power, can require the agent to buy his way into the contract. If $y < v^0$, then it is optimal for the principal to demand y from the agent and grant him the starting equity level v^0. If $v^0 < y < v^*$, then the principal receives $S(y)$ by requiring the agent to tender all his wealth. Because $S(y)$ is increasing, higher values of y result in a higher initial payoff for the principal. Finally, if $y \geq v^*$, then the agent has enough initial wealth to become a vested partner from the outset; that is, liquidity constraints never bind and the contract is first best. Although it is common wisdom that incentive problems can be eliminated under *ex post* private information by *selling the firm* to the agent, note that $v^* < v^{\mathrm{FB}}$ implies that it is not necessary to sell the *whole* firm because the first-best outcome obtains if the agent's equity position is at least v^*.

□ **Path dependence.** The maximal rent optimal contract specifies (u, q, w) as a function of equity, v. Therefore, the evolution of (u, q, w) depends on the evolution of v. Typically, the evolution of v along any sample path will depend on the order of shocks, which is true of models of dynamic contracting in general. Nevertheless, there is a very strong form of path dependence that holds in our model: in any arbitrary sample path, the order of the occurrences of shocks matters greatly. For instance, in any sample path where θ_1 occurs sufficiently often (the set of such sample paths has full measure), the agent strictly prefers to have all the θ_1 shocks in the beginning, because this will place him at v^* in finitely many periods, giving him a permanent ownership stake in the firm. Notice that this result holds for all revenue functions R that satisfy our assumptions. This is in contrast with a result in Thomas and Worrall (1990), who show that when an agent with a private endowment has Constant absolute risk aversion (CARA) utility, the optimal lending contract with a risk-neutral principal takes a simple form, where it is only the number of times a particular state (private income shock) has occurred that matters, and the order in which the shocks occur is irrelevant.

There are two reasons for this: first, once $v = v^*$, output is always first-best efficient from then on, and in any optimal contract, v never falls below v^* again, and second, from any initial $v > 0$, v^* can be reached in finitely many periods. More specifically, for any initial $v_0 \in (0, v^*)$, there exists an integer $\tau < \infty$ such that if the agent repeatedly receives θ_1 shocks over τ periods (which happens with strictly positive probability), he will reach v^*, that is, he will have $v^{(\tau)} = v^*$, in τ periods. This relies on two observations (see Lemma C1 in the Appendix). The first observation is that for any $v \in (0, v^*)$ and γ such that $P'(v) > \gamma > -1$, there is a $\tau < \infty$ such that if state θ_1 is repeated τ times, $P'(v^{(\tau)}) < \gamma$. The second observation is that for $v^{(\tau)} < v^*$, the sequence $v^{(\tau)} - v^{(\tau-1)}$ is increasing, which implies that $v^{(\tau)}$ reaches v^* in finitely many steps.

The path dependence property described above has another important implication. Fix $v \in (0, v^*)$ and suppose that the agent is promised utility v at time $t = 0$. Let ξ_v denote the time at which the agent becomes fully vested (i.e., his promised utility reaches v^*). Clearly, ξ_v is a random variable in that it depends on the realized shocks; in particular, it is a stopping time. Nevertheless, with probability 1, ξ_v is always finite (Lemma C2 in the Appendix). This property is what distinguishes our strong form of path dependence from the results in, for instance, Thomas and Worrall (1990). In that paper, immiseration occurs (with probability 1), and the agent's lifetime utility goes to $-\infty$ but takes infinitely long to do so, that is, the probability that the agent becomes immiserated in finite time is 0. This is because in Thomas and Worrall (1990) utilities are unbounded below, and although consumption slowly drops to zero, the agent's utility diverges to $-\infty$, which must necessarily take infinitely long along every path.

□ **Fixed costs and liquidation.** Suppose now that the cost of producing q units of output in state θ is $\theta q + c$, where $c > 0$ is a *fixed cost* of production. This clearly renders the cost function concave (and hence, nonconvex). Nevertheless, this does not introduce any significant changes to the value function. As before, if we set $u_i = m_i - \theta_i q_i$, then incentive constraints [C_{ij}] remain unchanged, whereas the promise-keeping constraint [PK] becomes $\sum_{i=1,2} f_i(u_i + \delta w_i) = v + c$, and the liquidity constraint [L_i'] becomes $u_i \geq c$. But using the equation [U_i], the liquidity constraints $u_i \geq c$ are seen to be exactly as in [L_i], that is, the effective liquidity constraints remain unchanged. Therefore, the principal's value function $P_c : \mathbb{R}_+ \to \mathbb{R}$ is a solution to the problem

$$P_c(v) = \max_{(q,w)} \sum_{i=1,2} f_i\big[\big(R(q_i) - \theta_i q_i\big) + \delta\big(P_c(w_i) + w_i\big)\big] - (v + c), \qquad \text{[V-Fix]}$$

subject to monotonicity in output [M_1], liquidity [L_i], and feasibility $q_2 \geq 0$ and $w_2 \geq 0$. There also exists a maximal rent contract as in Theorem 2.

It is easy to see that the value function P_c is concave and differs from P by c, that is, $P(v) = P_c(v) + c/(1 - \delta)$. This is because with a fixed cost, the principal has to compensate the agent, in every state, for the additional fixed cost of c, that is independent of output. Although this does not affect the basic features of the contract, this does allow for $P_c(v) < 0$ for some v even if $P(v) > 0$ for all $v \in [0, v^*]$. Thus, in the presence of a fixed cost $c > 0$, there exists $v_c > 0$ such that if the principal is allowed to liquidate the firm, that is, if she is allowed to make a severance payment to the agent of the amount of the utility owed and then terminate the production technology, she would do so. Of course, this is not an "equilibrium statement," but incorporating the option of liquidating the firm in the value function is nevertheless possible. In the next section, we consider such a value function in the case where the principal can fire the agent and hire another agent to replace him.

□ **Hiring and firing.** Up to now, we have considered the case where the principal can*not* fire the agent, where *firing* the agent entails making a severance payment equal to his promised utility immediately and terminating the relationship. However, it is clear that there are circumstances under which the principal would like to fire the agent, and replace him with a new one, if one were available. To see this, let $v > 0$, and recall that the principal's utility with this level of promised utility is $P(v)$. For firing to be optimal, it must be the case that $P(v) < P(v^0) - v$, that is, the utility from continuing in the relationship is less than the utility from starting anew with another agent after paying the current agent the utility owed him. This condition can be rewritten as $P(v) + v < P(v^0)$. Because P is continuous and because $P(0) + 0 < P(v^0)$, there exists $v^\dagger \in (0, v^0)$ such that $P(v^\dagger) + v^\dagger < P(v^0)$, where $v^\dagger$ is a critical level of equity such that it is optimal to fire the agent if sweat equity falls below $v^\dagger$ (also see Figure 1).

Lemma C1 shows that for any $C > 0$, the process P' is greater than C with strictly positive probability. Hence, there is a strictly positive probability that sweat equity will fall below any positive $v \in (0, v^0)$ and, hence, a positive probability that a given agent will get fired. Moreover, Doob's Maximal Inequality (see, for instance, Chung, 2001) provides a bound for this probability, wherein the probability that $P'(v) \geq C$ is less than $1/(1 + C)$.

Now, suppose there is an infinite pool of identical agents, but that the principal can only contract with one at a time. To formally incorporate the option to replace an agent, it is necessary to introduce a new value function $Q(v)$. For any function $Q : \mathbb{R}_+ \to \mathbb{R}$ bounded above, let $v_Q^0 \in \arg\max_x Q(x)$. Now let Q be the unique function that satisfies

$$Q(v) = \max\left[Q(v_Q^0) - v,\ \max_{(q,w)} \mathbf{E}\left[\left(R(q_i) - \theta_i q_i\right) + \delta\left(Q(w_i) + w_i\right)\right] - v\right]. \quad \text{s.t. } [M_1], [L_i], q_2 \geq 0 \text{ and } w_2 \geq 0.$$

At any level of sweat equity v such that it is not optimal to fire the current agent, $Q(v)$ obviously has the same properties as $P(v)$, although it lies above $P(v)$ for $v < v^*$ because the option to replace the agent has positive value because it is exercised with positive probability. Hence, for any $v < v_Q^0$ such that firing is not optimal, $Q(v)$ is increasing. Because $Q(v_Q^0) - v$ is decreasing, there exists a state $v_Q^\dagger$ such that it is optimal to fire the agent if $v < v_Q^\dagger$ and to retain him if $v > v_Q^\dagger$. An important feature of the value function Q is that it is *not* concave (a feature also present in Clementi and Hopenhayn, 2006).

In essence, the option to reset the process allows the principal to avoid very low levels of sweat equity and the associated large output restrictions. Rather than waiting for the agent to make the long and erratic climb back to v_Q^0, the principal simply pays him off and begins again with a new agent.

8. Applications

■ In order to focus on the underlying fundamental economic forces, the model analyzed above is necessarily stylized. Nevertheless, the environment we investigate, involving a liquidity-constrained entrepreneur who must contract for initial rounds of operating capital, has obvious real-world counterparts. In this section, we briefly discuss two examples mentioned in the Introduction, work-to-own franchise programs and venture capital contracts. In each of these settings, numerous features of the agreements closely parallel aspects of the theoretically optimal contract.

□ **Work-to-own franchising programs.** Franchising is a ubiquitous organizational form, especially in retailing. According to Blair and Lafontaine (2005), 34% of U.S. retail sales in 1986 (almost 13% of GDP) derived from franchised outlets. Estimates on the number of U.S. franchisers vary widely, but listings in directories suggest a figure between 2500 and 3000. The basic reasons for the prevalence of the franchise relationship accord well with our model. The franchiser wishes to expand into a specific market but lacks idiosyncratic knowledge about local factors influencing profitability such as demand and cost fluctuations. The franchisee observes local conditions but lacks brand recognition and an established business formula. Often, the franchisee also lacks sufficient seed capital for getting the business off the ground. For instance, Blair and Lafontaine (2005) suggest that franchisee capital constraints partially explain the wide discrepancy between the franchise fee of $125,000 charged by McDonald's in 1982 and the estimated present value of restaurant profits of between $300,000 and $450,000 over the duration of the contract.

In fact, many franchisers have explicit work-to-own or sweat equity programs designed to allow liquidity-constrained managers to become owners of their own franchises. These arrangements span a wide variety of retail businesses and industries including: *7-11* convenience stores, Big-O-Tires, Charley's Steakery, Fastframe, Fleet Feet Sports, Lawn Doctor, Petland, Outback Steakhouse, and Quiznos sandwiches, to name but a few. Although details of sweat equity arrangements vary across franchisers, Quiznos's *Operating Partner Program* is broadly representative, enabling experienced managers to receive financing from the parent company for all but $5000 of the upfront investment. A recent interview with Quiznos executive John Fitchett highlights the similarities between the restaurant chain's sweat equity program and the theoretically optimal contract discussed above.[16]

Private information and liquidity constraints: "The Operating Partner Program was developed in response to a successful pool of qualified, interested entrepreneurs with restaurant experience who would make great franchise owners, but lack access to the necessary financing."

Sweat equity and ownership: "Operating partners earn a salary and benefits as they work toward full ownership of the restaurant, with 80 percent of profits paying down Quiznos' contribution on a monthly basis. . . .we believe an operating partner that successfully operates the restaurant can reach the point of being able to acquire full ownership in two to five years."

Path dependence and replacement: "For the first year, Quiznos will cover any losses, and the amount will be added to the loan value. After 12 months, if the restaurant has not reached profitability, Quiznos and the operator will determine whether the operator is running his or her restaurant in the most effective way, or if there are other circumstances that may influence the profitability of the restaurant. [We will then] evaluate whether to put a new operator in the restaurant."

□ **Venture capital contracts.** Another contractual setting that accords neatly with our model is the venture capital market. Founders often wish to launch a business based on their personal expertise but do not possess sufficient financial resources. Venture capitalists (VCs) provide liquidity to startups, staging subsequent investments and founder compensation based on various performance criteria. Indeed, a case study by Robinson and Wasserman (2000) reports "A central concept used by VCs in structuring their investments is 'earn in', in which the entrepreneur earns his equity through succeeding at value creation. . .VCs also insist on vesting schedules for options or stock grants, whereby managers earn their stakes over a period of years."

In a pioneering article, Kaplan and Stromberg (2003) investigate 213 VC investments in 119 portfolio companies by 14 VC firms. Their findings also corroborate features of our optimal dynamic mechanism.

> In general, board rights, voting rights, and liquidation rights are allocated such that if the firm performs poorly, the VCs obtain full control. As performance improves, the entrepreneur retains/obtains more control rights. If the firm performs very well, the VCs retain their cash flow rights, but relinquish most of their control and liquidation rights. Ventures in which the VCs have voting and board majorities are also more likely to make the entrepreneur's equity claim and the release of committed funds contingent on performance milestones.

Although our stylized model does not directly address the plethora of contingencies and control rights found in typical VC contracts, Kaplan and Stromberg's (2003) findings are consistent with the main features of the theoretically optimal mechanism. Specifically, v, or sweat equity, is

[16] See Liddle (2010).

a summary statistic of past performance, and greater sweat equity leads to reductions in output distortions, less-stringent liquidity constraints, and eventually to agent ownership, whereas lower sweat equity results in higher distortions, more-stringent liquidity constraints, and ultimately even to replacement of the agent. In fact, the founders of poorly performing ventures are frequently ousted by the VCs, who either take direct control of the company themselves or hire new management. According to White, D'Souza, and McIlwraith (2007), VCs replace the founder with a new CEO in up to 50% of all venture-backed startups.

9. Conclusion

■ In this article, we explore the question of how a principal optimally contracts with an agent to operate a business enterprise over an infinite time horizon when the agent is liquidity constrained and has access to private information about the sequence of cost realizations. We formulate the mechanism design problem as a recursive dynamic program in which promised utility to the agent constitutes the relevant state variable.

We establish a bang-bang property of an optimal contract, wherein the agent is incentivized only through adjustments to his equity until achieving a critical level, after which he may be incentivized through cash payments. We can, therefore, interpret the incentive scheme as a sweat equity contract, where all rent payments are backloaded. The critical level of sweat equity occurs when none of the agent's liquidity constraints bind. At this point, the contract calls for efficient production in all future periods and the agent earns a permanent ownership stake in the enterprise, that is, he becomes a vested partner.

We demonstrate that the derivative of the principal's value function is a martingale, yielding several implications. First, for a given level of sweat equity, the set of cost reports can be partitioned into two subsets, *good* low-cost reports leading to higher levels of sweat equity and *bad* high-cost reports leading to lower levels. Second, if the principal cannot fire the agent, the martingale convergence theorem implies that he will eventually become an owner with probability 1; that is, the contract provides a *stairway to heaven*. On the other hand, if the principal has the option to replace the current agent with a new one, then she will do so after the agent's equity level in the firm becomes sufficiently low, an event that occurs with positive probability. Hence, the contract also embodies a *highway to hell*.

Finally, we show that the properties of the theoretically optimal contract square well with features common in real-world work-to-own franchising agreements and venture capital contracts. In both of these settings, managers are incentivized primarily through equity adjustments. Moreover, good outcomes lead to less-stringent controls by the franchiser/VC and increased autonomy by the manager, whereas bad outcomes have the reverse effects.

In essence, this article can be viewed as addressing the basic question of how an equity partner should optimally contract with a managing partner who possesses no wealth or access to outside sources of capital. The answer we obtain is intuitive. The equity partner should use a sweatequity contract to incentivize the manager, adjusting his ownership stake up when the firm performs well and down when it performs poorly. We show that the (potentially) long and winding road induced by such a contract must ultimately lead to ownership or to dismissal.

Appendix A

□ **Proofs from section 4.** We begin with a proof of Theorem 1.

Proof of Theorem 1. The proof is standard, which allows us to make frequent reference to Stokey, Lucas, and Prescott (1989). Recall that the state variable v, namely sweat equity or promised utility, lies in the set $[0, \infty]$. The principal can always just give the agent v utiles without requiring any production. This would give the agent v utiles and cost the principal $-v$ utiles, thus forming a lower bound for her utility. An upper bound for the principal's value function obtains if we consider the case where there is full information, in which case the principal's utility is

$$\frac{1}{1-\delta}\Big[f_1\big[R(q_1^*) - \theta_1 q_1^*\big] + f_2\big[R(q_2^*) - \theta_2 q_2^*\big]\Big] - v.$$

This entails giving the agent exactly v utiles (net of production costs) but getting efficient output in every state, that is, there are no output distortions. Therefore, the value function $P(v)$ must lie within these bounds, that is, must satisfy

$$0 \leq P(v) + v \leq \frac{1}{1-\delta} \sum_{i=1,2} f_i\left[R(q_i^*) - \theta_i q_i^*\right].$$

Let $C[0, \infty)$ be the space of continuous functions on $[0, \infty)$, and let

$$\mathscr{F} := \left\{ Q \in C[0, \infty) : 0 \leq Q(v) + v \leq \frac{1}{1-\delta} \sum_{i=1,2} f_i\left[R(q_i^*) - \theta_i q_i^*\right] \right\}$$

be endowed with the "sup" metric, which makes it a complete metric space. Let $\mathscr{F}_1$ be the set of all concave functions in $\mathscr{F}$, and let $\mathscr{F}_2$ be the set of all functions $Q \in \mathscr{F}$ such that $Q(v) + v$ is constant for all $v \geq v^\flat$, where $v^\flat := \frac{1}{1-\delta} f_1(\theta_2 - \theta_1)q_2^*$. It is easy to see that both $\mathscr{F}_1$ and $\mathscr{F}_2$ are closed subsets of $\mathscr{F}$.

Let $\Gamma_0(v) := \{(u, q, w) \in \mathbb{R}_+^n \times \mathbb{R}_+^n \times \mathbb{R}_+^n : (u, q, w)$ satisfies [PK], [Cij]$\}$ be the set of feasible (u, q, w). Notice that the liquidity constraints [L_i'] are automatically satisfied because $u_i \geq 0$ holds for all i; also, output and continuation promises are always nonnegative, as discussed in the text. Define the operator $\mathsf{T} : \mathscr{F} \to \mathscr{F}$ as

$$\begin{aligned} (\mathsf{T}Q)(v) = &\max\nolimits_{(u,q,w)} \textstyle\sum_{i=1,2} f_i\left[\left(R(q_i) - \theta_i q_i\right) - u_i + \delta Q(w_i)\right] \\ &\text{s.t. } (u, q, w) \in \Gamma_0(v) \end{aligned}$$

for each $Q \in \mathscr{F}$. It is easy to see that $\mathsf{T}Q(v) + v \geq 0$, and the incentive constraints force $\mathsf{T}Q(v) + v \leq \frac{1}{1-\delta} \sum_{i=1,2} f_i\left[R(q_i^*) - \theta_i q_i^*\right]$. Because $\Gamma_0(v)$ is compact for each v, the maximum is achieved for each v and, by the theorem of the maximum, $\mathsf{T}Q$ is continuous. Therefore, by the bounds established, we have $\mathsf{T}Q \in \mathscr{F}$. We shall now show that if $Q \in \mathscr{F}_1 \cap \mathscr{F}_2$, then $\mathsf{T}Q \in \mathscr{F}_1 \cap \mathscr{F}_2$.

Consider first the case where $Q \in \mathscr{F}_1$, and notice that $\Gamma_0(v)$ is defined by finitely many linear inequalities. This not only implies that $\Gamma_0(v)$ is convex for each $v \geq 0$ but also implies that if $v, v' \geq 0$, $(u, q, w) \in \Gamma_0(v)$, and $(u', q', w') \in \Gamma_0(v')$, then for all $\alpha \in [0, 1]$, $\alpha(u, q, w) + (1-\alpha)(u', q', w') \in \Gamma_0\left(\alpha v + (1-\alpha)v'\right)$. We can now adapt the arguments in Stokey et al. (1989) to conclude that if $Q \in \mathscr{F}_1$, we must also have $\mathsf{T}Q$ is continuous and concave.

Let us now assume that $Q \in \mathscr{F}_2$ so that $Q'(v) = -1$ for all $v \geq v^\flat$. Consider the relaxed problem

$$\begin{aligned} \max_{(u_i, q_i, w_i)} &\sum_i f_i\left[\left(R(q_i) - \theta_i q_i\right) + \delta w_i + \delta Q(w_i)\right] - v \\ &\text{s.t. [PK].} \end{aligned}$$

where $v \geq v^\flat$. It is easy to see that every solution to this problem must have $q_i = q_i^*$. Moreover, *a* solution (but certainly not the unique solution) to this problem has, in addition, $w_i = v^\flat$. By letting

$$u_i(v) := v - \delta v^\flat - f_1 \Delta q_2^* + (2 - i)\Delta q_2^*,$$

we see from [U_i] above that [PK] and [C_{ij}] ($i = 1$, $j = 2$) hold with equality, so that all the constraints, including liquidity, are satisfied. Therefore, the contract $(u(v), q_i^*, w_i = v^\flat) \in \Gamma_0(v)$, and is feasible, and is therefore a solution to the original constrained problem. In particular, for any $Q \in \mathscr{F}_2$,

$$\mathsf{T}Q(v) = \sum_i f_i\left[\left(R(q_i^*) - \theta_i q_i^*\right) + \delta v^\flat(b) + \delta Q\left(v^\flat\right)\right] - v$$

for all $v \geq v^\flat$. Indeed, with the contract $(u(v), q_i^*, w_i = v^\flat) \in \Gamma_0(v)$, for any $v, v' \geq v^\flat$,

$$\mathsf{T}Q(v) - \mathsf{T}Q(v') = -(v - v'),$$

that is, $(\mathsf{T}Q)'(v) = -1$ for all $v \geq v^\flat$, that is, $\mathsf{T}Q \in \mathscr{F}_2$.

It is easy to see that T is monotone ($Q_1 \leq Q_2$ implies $\mathsf{T}Q_1 \leq \mathsf{T}Q_2$) and satisfies discounting ($\mathsf{T}(Q + a) = \mathsf{T}Q + \delta a$, where $0 < \delta < 1$), which implies that T is a contraction mapping on $\mathscr{F}$. We have just established above that if $Q \in \mathscr{F}_1 \cap \mathscr{F}_2$, then $\mathsf{T}Q \in \mathscr{F}_1 \cap \mathscr{F}_2$. But this implies that the unique fixed point of T, which we shall call P, also lies in $\mathscr{F}_1 \cap \mathscr{F}_2$—see Stokey et al. (1989).

We now establish a lower bound on $P'(0)$. By Lemma A1 below, the optimal contract associated with $v = 0$ is $q_1 = q_1^*$, $q_2 = 0$, and $u_i = w_i = 0$ for $i = 1, 2$, and we have

$$P(0) = f_1\left[R(q_1^*) - \theta_1 q_1^*\right] + \delta P(0).$$

Because P is concave, we know $P'(0) \geq [P(\varepsilon) - P(0)]/\varepsilon$ for all $\varepsilon > 0$.

Now, consider a contract such that in the first period $q_1 = q_1^*$, $q_2 = x$, $w_i = 0$, and $u_i = (\theta_2 - \theta_i)x$ for $i = 1, 2$. From the second period on, the contract reverts to $v = 0$. Define x by

$$\varepsilon = f_1 u_1 + f_2 u_2 = (\theta_2 - \mathbf{E}[\theta])\,x.$$

to satisfy [PK]. Note that this contract satisfies all constraints.

The principal's payoff under the proposed contract is

$$\begin{aligned} Q(\varepsilon) &= f_1\left(R(q_1^*) - \theta_1 q_1^*\right) + f_2\big[R(x) - \theta_2 x\big] - \varepsilon + \delta P(0) \\ &= P(0) + f_2\big[R(x) - \theta_2 x\big] - \varepsilon. \end{aligned}$$

Note that $P(\varepsilon) \geq Q(\varepsilon)$ and $\lim_{\varepsilon \to 0} Q(\varepsilon) = P(0)$. Moreover,

$$Q(\varepsilon) - P(0) = \frac{f_2\big[R(x) - \theta_2 x\big] - \varepsilon}{\varepsilon}$$

so that

$$\begin{aligned} \lim_{\varepsilon \to 0} \frac{Q(\varepsilon) - P(0)}{\varepsilon} &= \lim_{\varepsilon \to 0} \frac{f_2\big[R(x) - \theta_2 x\big] - \varepsilon}{\varepsilon} \\ &= f_2 \lim_{x \to 0} \left[\frac{R(x) - \theta_2 x}{\left(\theta_2 - \mathbf{E}[\theta]\right)x}\right] - 1 \\ &= \frac{(1 - f_1)R'(0) + f_1\theta_1 - \mathbf{E}[\theta]}{\theta_2 - \mathbf{E}[\theta]} - 1 \\ &= \frac{(1 - f_1)R'(0) + f_1\theta_1 - \theta_2}{\theta_2 - \mathbf{E}[\theta]}, \end{aligned}$$

where we have used $\varepsilon = x(\theta_2 - \mathbf{E}[\theta])$ in the second equality. This gives us the bound

$$\begin{aligned} P'(0) &= \lim_{\varepsilon \to 0} \frac{P(\varepsilon) - P(0)}{\varepsilon} \\ &\geq \lim_{\varepsilon \to 0} \frac{Q(\varepsilon) - P(0)}{\varepsilon} \\ &= \frac{(1 - f_1)R'(0) + f_1\theta_1 - \theta_2}{\theta_2 - \mathbf{E}[\theta]} \end{aligned}$$

as required. Notice now that if [VF$'$] holds, that is, if $R'(0) = \infty$, it follows immediately that $P'(0) = \infty$.

Because the optimal contract lies in the interior of the feasible set (in an appropriate sense), the continuous differentiability of P follows from standard results as, for instance, in Stokey et al. (1989). Because P is concave and $P'(v) = -1$ for all $v \geq v^\flat$, there is a smallest v such that $P'(v) = -1$; let $v^* := \min\{v : P'(v) = -1\}$. In sum, $P'(v) = -1$ for all $v \geq v^*$ and $P'(v) > -1$ for all $v < v^*$. Moreover, by construction, $v^* \leq v^\flat$. (Of course, it is shown in Section 5 that in fact $v^* = v^\flat$.) *Q.E.D.*

Although we do not (yet) know much about the optimal contract, the following lemma tells us what any optimal contract must look like at $v = 0$.

Lemma A1. If $v = 0$, any optimal contract entails $u_i = w_i = 0$ for all i, $q_1 = q_1^*$, and $q_2 = 0$. *Q.E.D.*

Proof. To see this, recall that feasibility implies $u_i, w_i \geq 0$ for all i. Promise keeping [PK] requires $\sum_i f_i\big[u_i + \delta w_i\big] = 0$, which implies $u_i = w_i = 0$ for all i. This observation and $C_{1,2}$ in turn imply that $q_2 = 0$. The intuition is simply that if $q_2 > 0$, then some rent must paid to the low-cost type θ_1 which, due to the liquidity and feasibility constraints, would violate [PK].

Proof of Proposition 1. Notice that from [PK] the value function can be written as

$$P(v) = \max_{(u,q,w)} \sum_i f_i\big[\big(R(q_i) - \theta_i q_i\big) - u_i + \delta w_i + \delta P(w_i)\big]$$

subject to all the constraints. So suppose $w_i(v) < v^*$ for some v, and by way of contradiction, $u_i(v) > 0$. Notice that in the constraints [C$_{ij}$] and [PK], u_i and w_i appear in the form $u_i + \delta w_i$. Because $u_i > 0$, we can reduce it by an appropriately chosen $\varepsilon > 0$ and increase w_i by ε/δ. This leaves the [PK] and [C$_{ij}$] constraints unchanged. Moreover, liquidity constraint [L$_i$] is also unaffected. Last, the q_is are left unchanged. Therefore, this new contract is feasible, and is also a strict improvement, because $w_i + P(w_i)$ is strictly increasing for $w_i < v^*$ (by Theorem 1), which contradicts the optimality of the original contract. Therefore, it must be that for any optimal contract, $w_i(v) < v^*$ implies $u_i(v) = 0$.

Suppose now that in the optimal contract, we have $w_i(v) > v^*$ for some $\theta_i \in \Theta$ and $v \geq 0$. In analogy with the steps above, we can increase u_i by $\varepsilon = w_i(v) - v^*$, and reduce $w_i(v)$ by ε/δ. Because $P(w) + w$ is linear when $w \geq v^*$, it is easy to see that this new contract is also optimal, but now has $w_i(v) \leq v^*$, that is, the new contract is maximal rent, which completes the proof. *Q.E.D.*

The following lemma breaks down the proof of Lemma 1 into easily digestible parts.

Lemma A2.

(a) For all v, q is monotone in type, that is, $q_1(v) \geq q_2(v)$.
(b) If the constraint $C_{1,2}$ holds with equality and $q_1(v) \geq q_2(v)$, then the constraint $C_{2,1}$ holds.
(c) We may assume that the constraint $C_{1,2}$ holds with equality; that is, if the constraint $C_{1,2}$ is slack, there is another contract that gives the principal at least as much utility but where $C_{1,2}$ holds with equality.
(d) In any maximal rent contract, [C_1] and [L_i] imply that u and w are monotone in type.

Proof.

(a) That $q_1 \geq q_2$ follows by adding $C_{1,2}$ and $C_{2,1}$.
(b) That $q_1(v) \geq q_2(v)$ and the equality of $C_{1,2}$ implies $C_{2,1}$ is standard, and therefore omitted.
(c) We want to show that $C_{1,2}$ holds with equality. By the results above, we may assume that q is monotone in type. Suppose that $C_{1,2}$ is slack, so that $u_1 + \delta w_1 > u_2 + \delta w_2 + \Delta q_2$. There are two cases to consider. The first case occurs when $u_1 > u_2$. We can increase u_2 by ε and reduce u_1 by $(f_1/f_2)\varepsilon$, so that [PK] still holds, the incentive constraints are not upset, and the objective is unchanged. We may choose ε so that $C_{1,2}$ holds with equality, which proves this case.

The second case occurs where $u_1 \leq u_2$, which implies $w_1 > w_2$. Replace w_1 with $w_1' := w_1 - \varepsilon$, and replace w_2 with $w_2' := w_2 + (f_1/f_2)\varepsilon$, where $\varepsilon > 0$ is chosen so that $C_{1,2}$ holds with equality. Notice that because $q_2 \geq 0$, it must be that $w_1' \geq w_2'$. We want to show this change does not leave the principal any worse off.

To see this, notice that by construction, $f_1 w_1 + f_2 w_2 = f_1 w_1' + f_2 w_2'$. Therefore, it only remains to show that $f_1 P(w_1') + f_2 P(w_2') \geq f_1 P(w_1) + f_2 P(w_2)$, which holds if, and only if, $f_2\big[P(w_2') - P(w_2)\big] \geq f_1\big[P(w_1) - P(w_1')\big]$. Recall that P is continuously differentiable, so that if $w_2' \leq w_1'$, the concavity of P implies $P'(w_2') \geq P'(w_1')$. We then observe

$$\begin{aligned} f_2\big[P(w_2') - P(w_2)\big] &\geq f_2 P'(w_2')(w_2' - w_2) \\ &= f_1 P'(w_2')\varepsilon \\ &\geq f_1 P'(w_1')(w_1 - w_1') \\ &\geq f_1\big[P(w_1) - P(w_1')\big], \end{aligned}$$

where we have used the fact that $f_1(w_1 - w_1') = f_1\varepsilon = f_2(w_2' - w_2)$, and the first and last inequalitities follow from the definition of the subdifferential, and the second follows from the concavity of P. This proves our claim.
(d) We shall show that u and w are monotone in type in any maximal rent contract. Suppose first that $u_1 < u_2$. Then, by the liquidity constraint [L_i'], it must be that $u_2 > 0$. But by Proposition 1, this implies $w_2 = v^*$. Now, ($C_{1,2}$) implies $\delta w_1 \geq (u_2 - u_1) + \delta w_1 + \Delta q_2 > \delta w_1$, which implies $w_2 > v^*$, which is impossible in a maximal rent contract. Therefore, it must be that u is monotone in type.

Next, let us assume that $w_2 > w_1$. Once again, ($C_{1,2}$) implies $u_1 - u_2 \geq \delta(w_2 - w_1) + \Delta q_2 > 0$, which implies, by [$L_i'$], that $u_1 > 0$. But Proposition 1 says we must have $w_1 = v^*$, which in turn implies $w_2 > v^*$, which is impossible in a maximal rent contract. Therefore, w must also be monotone in type. *Q.E.D.*

Proof of Lemma 2. Note that [C_1] and [PK] can be rewritten to give

$$\begin{aligned} u_1 + \delta w_1 &= v + f_2\Delta q_2 \\ u_2 + \delta w_2 &= v - f_1\Delta q_2, \end{aligned}$$

which can be rewritten as

$$u_i + \delta w_i = v - f_1\Delta q_2 + (2 - i)\Delta q_2$$

as required by [U_i]. *Q.E.D.*

Proof of Theorem 2. The only part that remains to be proved is the uniqueness of the maximal rent contract. The first claim is that for any $v \geq 0$, there is a unique $q_i(v)$ for each i for all maximizers (q, w). To see this, suppose (q, w) and (q', w') are optimal at some v, but $q \neq q'$. Then, because the feasible set is convex, $\big(\frac{1}{2}(q + q'), \frac{1}{2}(w + w')\big)$ is also feasible and, moreover, is a strict improvement over (q, w) and (q', w'), because $R(q)$ is strictly concave. Therefore, it must be that $q = q'$ across all optimal contracts.

By Proposition 1 and Lemma 1, we know that for each v, $u_2 = 0$ implies $u_1 = 0$. Suppose v is such that $u_i(v) = 0$ for some i. We have already established that $q_j(v)$ is unique for all j. This implies that there is a unique $w_i(v)$ such that [L_i] holds with equality. On the other hand, if $u_i > 0$ for some v, then it must be that $w_i = v^*$, because we have a maximal rent contract. In either case, $w_i(v)$ is uniquely determined in a maximal rent contract.

Finally, Stokey et al. (1989) show that the maximal rent optimal contract must be continuous in v. *Q.E.D.*

Appendix B

□ **Proofs from section 5.** First we present the derivation of v^* given in (Vest).

Proof of Lemma 3. Because $P'(v^*) = -1$, we have $\Lambda(v^*) = 0$, and because $\lambda_i \geq 0$ for all i, it must be that $\lambda_i(v^*) = 0$ for all i.

By the definitions of P and v^*, we also have $\Lambda(v) > 0$ for all $v < v^*$. Lemma 1, which says that rents are monotone in type, now implies $\lambda_2(v) > 0$ for all $v < v^*$. But by complementary slackness, $u_2(v) = 0$ for all $v < v^*$. Because the optimal contract is continuous in v (see Theorem 2), it follows that $u_2(v^*) = 0$.

From the first-order conditions, if $v = v^*$, then $P'(w_i) = -1$, which implies $w_i(v^*) = v^*$ for all i (in a maximal rent contract). Therefore, (12) holds by [U_i] (for $i = 2$). *Q.E.D.*

Next, for ease of exposition, we shall provide some lemmas that are of independent interest and present results in an order somewhat different from the text, which allows this material to be relatively self-contained. We begin with an observation about the implications of local linearity of the value function.

Lemma B1. Let $0 \leq v_\circ < v^\circ$. If P is linear on $[v_\circ, v^\circ]$, any optimal contract must have q constant on $[v_\circ, v^\circ]$, that is, $q(v) = q(v')$ for all $v, v' \in [v_\circ, v^\circ]$.

Proof. It is easy to see that at each v, if (u, q, w) and (u', q', w') are part of optimal contracts (maximal rent or not), it must necessarily be that $q = q'$. This follows from the convexity of the set of maximizers and the strict concavity of R. It is easily seen that we may consider, without loss of generality, maximal rent contracts.

We will prove the contrapositive of the assertion. Let $v, v' \in [v_\circ, v^\circ]$, and suppose q, q' are optimal at v and v', respectively, with $q \neq q'$. For any $\alpha \in (0, 1)$, let $(q^\alpha, w^\alpha) = \alpha(q, w) + (1 - \alpha)(q', w')$. Notice that the constraint [L_i] can be written as $a_i q_2 + \delta w_i \leq v$, where $a_i \in \mathbb{R}$ for $i = 1, 2$. Therefore, (q^α, w^α) is certainly feasible at $v^\alpha := \alpha v + (1 - \alpha)v'$, that is, (q^α, w^α) satisfies [L_i] and (M_i) for all i. Then,

$$
\begin{aligned}
&P\big(\alpha v + (1-\alpha)v'\big) \\
&\geq \sum_i f_i\big[\big(R(q_i^\alpha) - \theta_i q_i^\alpha\big) + \delta w_i^\alpha + \delta P(w_i^\alpha)\big] - v^\alpha \\
&> \alpha \sum_i f_i\big[\big(R(q_i) - \theta_i q_i\big) + \delta w_i + \delta P(w_i)\big] - \alpha v \\
&\quad + (1-\alpha)\sum_i f_i\big[\big(R(q_i') - \theta_i q_i'\big) + \delta w_i' + \delta P(w_i')\big] - (1-\alpha)v' \\
&= \alpha P(v) + (1-\alpha)P(v'),
\end{aligned}
$$

where the strict inequality follows from the strict concavity of R. This proves the strict concavity of P, as required. *Q.E.D.*

The following lemma provides some useful bounds on the Lagrange multipliers. As in the text, we shall assume, unless otherwise mentioned, that the contracts under question are maximal rent contracts.

Proposition B1. The Lagrange multipliers satisfy the following inequalities.

(a) $\dfrac{\lambda_2}{f_2} \geq \dfrac{\lambda_1}{f_1}$.

(b) $\dfrac{\lambda_2}{f_2} \geq \Lambda \geq \dfrac{\lambda_1}{f_1}$.

Proof.

(a) By Lemma 1, we know that in a maximal rent contract, for each v, $w_1(v) \geq w_2(v)$. The concavity of P then implies that $P'(w_1) \leq P'(w_2)$. By the first-order condition for w_i, namely [FOw_i], we see that $-1 + \lambda_1/f_1 = P'(w_1) \leq P'(w_2) = -1 + \lambda_2/f_2$. This allows us to conclude that $\lambda_2/f_2 \geq \lambda_1/f_1$, as claimed.

(b) The previous part tells us that $f_1\lambda_2 \geq f_2\lambda_1$. Adding $f_1\lambda_1$ to both sides gives us $\Lambda \geq \lambda_1/f_1$. Suppose now that $\Lambda > \lambda_2/f_2$. Then, $\Lambda > \lambda_1/f_1$, which implies that $\Lambda > f_1(\lambda_1/f_1) + f_2(\lambda_2/f_2) = \Lambda$, which is impossible. So it must be that $\lambda_2/f_2 \geq \Lambda$, as claimed. *Q.E.D.*

The following is an easy corollary of the proposition above.

Corollary B1. If $\lambda_1/f_1 = \lambda_2/f_2$, then $\lambda_i = f_i\Lambda$ for $i = 1, 2$.

Proof. Follows immediately from part (b) of Proposition B1 above. *Q.E.D.*

An obvious question is whether the optimal contract can ever have greater than optimal production, which was stated as Proposition 2 in the main text. We are now in a position to establish this.

Proof of Proposition 2. All that remains to establish the proposition is to show that $q_2(v) \le q_2^*$ for all $v \ge 0$. Notice that the first-order condition for q_2, [FOq_2], can be written as

$$R'(q_2) - \theta_2 = \frac{\Delta}{f_2}\left[f_1\lambda_2 - f_2\lambda_1\right].$$

Clearly, $q_2 \ge q_2^*$ if, and only if, $\lambda_2/f_2 < \lambda_1/f_1$, which is impossible by part (a) of Proposition B1, thereby completing the proof. *Q.E.D.*

Appendix C

□ **Proofs from sections 6 and 7.**

Proof of Proposition 3. Recall that w is monotone in type, that is, $w_1 \ge w_2$, which implies $P'(w_2) \ge P'(w_1)$. The claim is that for all $v \in (0, v^*)$, $P'(w_2) > P'(v) > P'(w_1)$. So suppose the claim is not true. Because P' is a martingale, the only possibility then is that $P'(w_1) = P'(w_2) = P'(v)$. (Notice that this does not imply $w_1 = w_2 = v$, because we haven't established that P is strictly concave.)

Because $P'(0) = \infty$, we know that $w_2 > 0$. The first-order condition [FOw_i] then implies $\lambda_1(v)/f_1 = \lambda_2(v)/f_2 > 0$. Corollary B1 then implies $\lambda_i = f_i\Lambda$, whereas [FOq_2] implies $q_i(v) = q_i^*$ for all i.

In this case, Proposition 1 and Lemma 2 give $\delta w_1 = v + f_2\Delta q_2^*$ and $\delta w_2 = v - f_1\Delta q_2^*$. The first equation establishes $w_1(v) > v$ directly. From the second equation, $w_2(v) < v$ if and only if $\left[v - f_1\Delta q_2^*\right]/\delta < v$ if, and only if, $v < \frac{1}{1-\delta} f_1\Delta q_2^* = v^*$, which is true by hypothesis. Therefore, we have $w_1(v) > v > w_2(v)$. Moreover, $\delta w_1 - \delta w_2 = \Delta q_2^*$.

We now proceed to show, by contradiction, that $P'(w_1) = P'(w_2) = P'(v)$ is impossible. Let $v_0 := v$, so that $w_2(v_0) < v_0 < w_1(v_0)$, $w_1(v_0) - w_2(v_0) = \Delta q_2^*/\delta$, and $P'(v) > -1$ and $q_i(v) = q_i^*$ for all $v \in [w_2(v_0), w_1(v_0)]$. Consider the sequence $v_k := w_1(v_{k-1})$ and, suppose, as the induction hypothesis, that $P'(\cdot)$ is constant (and strictly greater than -1) on the interval $[w_2(v_{k-1}), v_k]$, with $v_{k-1} \in (w_2(v_{k-1}), v_k)$.

Because $q_i(v) \le q_i^*$ for all v, it follows that $v_k = w_1(v_{k-1}) > w_2(v_k) > w_2(v_{k-1})$, which implies $P'(v_k) = P'(w_2(v_k))$, which in turn implies that $P'(v_k) = P'(v_{k+1})$. Therefore, $P'(\cdot)$ is constant (and strictly greater than -1) on the interval $[w_2(v_0), v_{k+1}]$. Because (v_k) is a strictly increasing sequence that diverges to infinity, we see that $P'(\cdot)$ must then be constant and strictly greater than -1 on the interval $[w_2(v_0), \infty)$, which is impossible because $P'(v) = -1$ for $v \ge v^*$. This completes the proof. *Q.E.D.*

We now prove another useful lemma that shows that with positive probability, the martingale P' can take all values in $(-1, \infty)$.

Lemma C1. For any $v \in (0, v^*)$ and $\gamma > P'(v)$, if state θ_2 is repeated τ times consecutively, then $P'(v^{(\tau)}) > \gamma$ for τ large enough. Similarly, for $-1 < \gamma < P'(v)$, if state θ_1 is repeated τ times consecutively, then $P'(v^{(\tau)}) < \gamma$ for τ large enough. Moreover, there exists $\tau < \infty$ such that if state θ_1 is repeated τ times consecutively, then $P'(v^{(\tau)}) = -1$; that is, $v^{(\tau)} = v^*$.

Proof. Suppose state θ_2 occurs repeatedly. This gives us a sequence $v_0 = v$, $v_1 := w_2(v_0) < v_0$, and $v^{(\tau)} := w_2(v^{(\tau-1)}) < v^{(\tau-1)}$. Because $(v^{(\tau)})$ is a strictly decreasing sequence that is bounded below by 0, it has a limit. The first part is proved if we can show that this limit is 0, because $P'(0) = \infty$.

Therefore, suppose the claim is not true. This implies there is some $y > 0$ such that $\lim_{\tau\to\infty} v^{(\tau)} = y$. In other words, $\lim_{\tau\to\infty} w_2(v^{(\tau)}) = y$. Because the optimal contract is continuous in v, $w_2(\cdot)$ is continuous in v. Therefore, $w_2(y) = y$, which contradicts Proposition 3, which requires that $w_2(y) < y$. This gives us the desired contradiction. The proof of the second part is similar and therefore omitted.

To prove the third part of the claim, consider the sequence $v^{(\tau)} := w_1(v^{(\tau-1)})$. We know that $(v^{(\tau)})$ is a strictly increasing sequence and that $\lim_{\tau\to\infty} v^{(\tau)} = v^*$. By Proposition 1 and Lemma 2, we know that for $v^{(\tau-1)} < v^*$,

$$\delta v_\tau = v_{\tau-1} + f_2\Delta q_2(v^{(\tau-1)}) > 0,$$

and hence

$$v^{(\tau)} - v^{(\tau-1)} \ge \frac{(1-\delta)v^{(\tau-1)}}{\delta} > 0.$$

Thus, $v^{(\tau)} - v^{(\tau-1)}$ is a positive increasing sequence which implies that $v^{(\tau)}$ achieves the limit v^* in a finite number of steps. *Q.E.D.*

We now move to the proof of Theorem 3. Once again, we follow Thomas and Worrall (1990).

Proof of Theorem 3. Because P' is a martingale that is bounded below by -1, it follows that $P' + 1$ is a nonnegative martingale. The martingale convergence theorem (see, for instance, Chung, 2001), says that $P' + 1$ converges almost surely to a nonnegative, integrable limit, $P'_\infty + 1$. Therefore, P' converges almost surely to P'_∞, and the limit is integrable (which implies that $P'_\infty = \infty$ with zero probability). We want to show that $P'_\infty = -1$ almost surely.

Before getting to the details, it is useful to sketch the intuition. Consider a sample path $(\theta^{(t)})$ such that P' converges along this sample path. Suppose that along this sample path, P' converges to some number $C > -1$ and $\hat{v}$ is such that $P'(\hat{v}) = C$. It must be that eventually, all the values that P' takes in this sample path must lie arbitrarily close to C. Therefore, along this path, the step size of the continuation promises $w_1^{(t)} - w_2^{(t)}$ must converge to zero. But this would violate Proposition 3, which says that $w_1(\hat{v}) - w_2(\hat{v})$ is bounded away from zero and the fact that the optimal contract is continuous in v.

Consider a sample path with the properties that (i) $\lim_{t\to\infty} P'(v^t) = C \notin \{-1, \infty\}$ and (ii) state θ_2 occurs infinitely often, and define $C =: P'(y)$, so that $\lim_{t\to\infty} v^t = y$. Consider a subsequence $(\sigma(t))$ such that $\theta^{\sigma(t)} = \theta_2$ for all t, that is, this is the subsequence consisting of all the θ_2 shocks in the original sequence. Because $(v^{\sigma(t)})$ is a subsequence of (v^t), it also converges to y.

Recall that the evolution of promised utility along any sample path can be written as $\varphi(v^t, \theta_i) = v^{t+1}$, where $\varphi(v, \theta_i)$ is continuous in v. This induces the function $\varphi^\sigma(v, \theta_2)$ where $\varphi^\sigma(v^{\sigma(t)}, \theta_2) = v^{\sigma(t+1)}$. Because $\varphi(v, \theta_i)$ is continuous in v, it follows that $\varphi^\sigma(v, \theta_2)$ is also continuous in v. Therefore, the sequence $\varphi^\sigma(v^{\sigma(t)}, \theta_2)$ converges to $\varphi^\sigma(y, \theta_2)$. Moreover, $\varphi^\sigma(y, \theta_2) = \varphi(y, \theta_2) = y$, because $\varphi^\sigma(v^{\sigma(t)}, \theta_2) = v^{\sigma(t+1)}$, and $\lim_{t\to\infty} v^{\sigma(t)} = \lim_{t\to\infty} v^t = y$.

But $\lim_{t\to\infty} P'(v^{\sigma(t)}) = C$ and $\lim_{t\to\infty} P'(v^{\sigma(t+1)}) = C$, so by the continuity of P' we have $P'(y) = P'\big(\varphi^\sigma(y, \theta_2)\big) = P'\big(\varphi(y, \theta_2)\big) = C$, contradicting Proposition 3, which states that $P'(y) < P'\big(\varphi(y, \theta_2)\big)$. But paths where state θ_2 does not occur infinitely often are of probability zero, which proves the proposition. *Q.E.D.*

As in the text in Section 7, let ξ_v denote the time taken for the process to reach v^* when starting from $v \in (0, v^*)$. We can now prove Lemma C2.

Lemma C2. With probability 1, $\xi_v < \infty$.

Proof. Let Θ^∞ be the set of all sample paths. By Theorem 3, along almost every path and from any starting v, the sequence of promised utilities converges to v^*. Fix a path $(\theta^{(t)})$ where θ_1 occurs infinitely often (such paths are of full measure), and let $(v^{(t)})$ denote the induced path of promised utilities. Because $v^{(t)}$ converges to v^*, for any $\varepsilon > 0$ neighborhood of v^* there exists $N > 0$ such that for all $t > N$, $v^{(t)}$ lies in this neighbourhood. But Lemma C1 says that there exists an ε^* neighborhood of -1 such that if $v \in (v^* - \varepsilon^*, v^*)$, then a shock of θ_1 results in a promised utility of v^*. Because the shock θ_1 occurs infinitely often, there is some $t > N$ such that $\theta^{(t)} = \theta_1$, which proves that along such paths, ξ_v is finite. But the paths under consideration have full measure, which implies that with probability 1, $\xi_v < \infty$, as required. *Q.E.D.*

Proof of Lemma 5. Because $R''(q) < 0$, it follows that $R'(q) > \theta_2$ for $q < q_2^*$ and $R'(q) > \theta_1$ for $q < q_1^*$. Recalling that $q_2^* < q_1^*$ then yields

$$\int_0^{q_2^*} \left[R'(q) - \theta_2\right] \mathrm{d}q + \int_{q_2^*}^{q_1^*} \left[R'(q) - \theta_1\right] \mathrm{d}q > 0.$$

Evaluating the integrals (where $R(0) = 0$) and rearranging terms gives

$$R(q_1^*) - \theta_1 q_1^* > (\theta_2 - \theta_1) q_2^*$$

or $f_1\left(R(q_1^*) - \theta_1 q_1^*\right) > f_1(\theta_2 - \theta_1) q_2^*$. Because $f_2\left(R(q_2^*) - \theta_2 q_2^*\right) > 0$, we have

$$f_1\left(R(q_1^*) - \theta_1 q_1^*\right) + f_2\left(R(q_2^*) - \theta_2 q_2^*\right) > f_1(\theta_2 - \theta_1) q_2^*.$$

Dividing both sides by $1 - \delta$ establishes the claim. *Q.E.D.*

References

ALBANESI, S. AND SLEET, C. "Dynamic Optimal Taxation with Private Information." *Review of Economic Studies,* Vol. 73 (2006), pp. 1–30.

ATHEY, S. AND SEGAL, I. "An Efficient Dynamic Mechanism." Working Paper, Stanford University, 2012.

BARON, D.P. AND BESANKO, D. "Regulation and Information in a Continuing Relationship." *Information Economics and Policy,* Vol. 1 (1984), pp. 267–302.

BATTAGLINI, M. "Long-Term Contracting with Markovian Consumers." *American Economic Review,* Vol. 3 (2005), pp. 637–658.

BERGEMANN, D. AND VÄLIMÄKI, J. "The Dynamic Pivot Mechanism." *Econometrica,* Vol. 78 (2010), pp. 771–789.

BERLE, A. AND MEANS, G. *The Modern Corporation and Private Property,* New York: Harcourt, Brace and World, rev. ed, 1968.

Biais, B., Mariotti, T., Plantin, G., and Rochet, J-C. "Dynamic Security Design: Convergence to Continuous Time and Asset Pricing Implications." *Review of Economic Studies,* Vol. 74 (2007), pp. 345–390.

Blair, R. and Lafontaine, F. *The Economics of Franchising,* New York: Cambridge University Press, 2005.

Boleslavsky, R. and Said, M. "Progressive Screening: Long Term Contracting with a Privately Known Stochastic Process." *Review of Economic Studies,* forthcoming.

Bolton, P. and Scharfstein, D. "A Theory of Predation Based on Agency Problems in Financial Contracting." *American Economic Review,* Vol. 80 (1990), pp. 93–106.

Chung, K.L. *A Course in Probability Theory,* San Diego, Calif.: Academic Press, 3d ed., 2001.

Clementi, G.L. and Hopenhayn, H. "A Theory of Financing Constraints and Firm Dynamics." *Quarterly Journal of Economics,* Vol. 121 (2006), pp. 229–265.

Covallo, R. "Efficiency and Redistribution in Dynamic Mechanism Design." In *Proceedings of the 9th ACM Conference on Electronic Commerce*. Chicago: ACM, 2008.

DeMarzo, P. and Fishman, M. "Agency and Optimal Investment Dynamics." *Review of Financial Studies,* Vol. 20 (2007), pp. 151–188.

—— and Sannikov, Y. "Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Agency Model." *Journal of Finance,* Vol. 61 (2006), pp. 2681–2724.

Edmans, A., Gabaix, X., Sadzik, T., and Sannikov, Y. "Dynamic CEO Compensation." *Journal of Finance,* Vol. 67 (2012), pp. 1603–1467.

Fernandes, A. and Phelan, C. "A Recursive Formulation for Repeated Agency with History Dependence." *Journal of Economic Theory,* Vol. 91 (2000), pp. 223–247.

Green, E.J. "Lending and the Smoothing of Uninsurable Income." In E.C. Prescott and N. Wallace, eds, *Contractual Agreements for Intertemporal Trade*. Minneapolis: University of Minnesota Press, 1987.

Harvard Business School. "Venture Capital Negotiations: VC versus Entrepreneur." HBS Case Study no. 9-800-170. Boston: Harvard Business School, 2000.

Kaplan, S. and Stromberg, P. "Financial Contracting Theory Meets the Real World: An Empirical Analysis of Venture Capital Contracts." *Review of Economic Studies,* Vol. 70 (2003), pp. 281–315.

Laffont, J.-J. and Martimort, D. *The Theory of Incentives*. Princetion, NJ: Princeton University Press, 2002.

Liddle, A. "Investing in Sweat Equity." *Nation's Restaurant News*. www.nrn.com/article/investing-sweat-equity, 2010.

Milgrom, P. "Adverse Selection without Hidden Information." Technical Report no. 8742, University of California, Berkeley, 1987.

Mirrlees, J. "An Exploration in the Theory of Optimum Income Taxation." *Review of Economic Studies,* Vol. 38 (1971), pp. 175–208.

Pavan, A., Segal, I., and Toikka, J. "Dynamic Mechanism Design." Working Paper, MIT, 2012.

Quadrini, V. "Investment and Liquidation in Renegotiation-Proof Contracts with Moral Hazard." *Journal of Monetary Economics,* Vol. 51 (2004), pp. 713–751.

Robinson, R.J., and Wassermann, N.T. "Venture Capital Negotiations: VC versus Entrepreneur." Harvard Business School Background Note 800-170, March 2000.

Sannikov, Y. "A Continuous-Time Version of the Principal-Agent Problem." *Review of Economic Studies,* Vol. 75 (2008), pp. 957–984.

Spear, S. and Srivastava, S. "On Repeated Moral Hazard with Discounting." *Review of Economic Studies,* Vol. 54 (1987), pp. 599–617.

Stokey, N.L., Lucas, R.E., Jr, and Prescott, E.C. *Recursive Methods in Economic Dynamics*. Cambridge, Mass.: Harvard University Press, 1989.

Thomas, J.P. and Worrall, T. "Income Fluctuation and Asymmetric Information: An Example of a Repeated Principal-Agent Problem." *Journal of Economic Theory,* Vol. 51 (1990), pp. 367–390.

White, R., D'Souza, R., and McIlwraith, J. "Leadership in Venture Backed Companies: Going the Distance." *Journal of Leadership and Organizational Studies,* Vol. 13 (2007), pp. 121–132.