

Stairway to Heaven or Highway to Hell: Liquidity, Sweat Equity, and the Uncertain Path to Ownership*

R. VIJAY KRISHNA[†]

CURTIS R. TAYLOR[‡]

MARCH, 2011

Abstract

A principal contracts optimally with an agent to operate a firm 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 this mechanism design problem as a recursive dynamic program in which promised utility to the agent is the relevant state variable. By establishing that output distortions and the stringency of liquidity constraints decrease monotonically in promised utility, we are able to interpret the state variable as the agent's equity in the firm. We establish a bang-bang property of optimal contracts wherein the agent is incentivised only through adjustments to his future utility until achieving a critical level of equity, after which he may be incentivised through cash payments, that is, through instantaneous rents. Thus the incentive scheme resembles what is commonly regarded as a *sweat equity* contract, with all rents, ie, cash payments net of costs, being back loaded. A 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 attains a vested ownership stake in the firm. Finally, properties of the

(*) We thank Rachael Kranton and Philipp Sadowski for helpful comments, Pino Lopomo for conversations that helped the project off the ground, and are grateful for the able research assistance of Joe Mazur and Sergiu Ungureanu.

(†) University of North Carolina, Chapel Hill <rvk@unc.edu>

(‡) Duke University <curtis.taylor@duke.edu>

theoretically optimal contract are shown to be similar to features common in real-world work-to-own franchising agreements and venture capital contracts.

Key Words: liquidity, sweat equity, monotone contracts, dynamic screening, franchising, venture capital, ownership

JEL Classifications: C61, D82, D86, L26

Contents

1	Introduction	2	9	Monotone Contracts: Another Look	29
2	Related Literature	5	10	Conclusion	32
3	The Model	9	Appendices		
4	Contract Design	11	A	Proofs from Section 4	35
5	Monotone Contracts	16	B	Proofs from Section 5	41
6	Dynamics	20	C	Proofs from Section 6	46
7	Discussion and Extensions	22	D	Monotonicity of Maximisers	49
7.1	Social Cost of Liquidity Constraints	22	D.1	Comparative Statics for Constrained Optimisation	49
7.2	The Path to Ownership	23	D.2	The Problem at Hand	51
7.3	Hiring and Firing	24	D.3	Some Equivalent Problems	53
7.4	Path Dependence	25	D.4	Coda	54
8	Applications	26	References		
					55

1. Introduction

Few entities are more representative of the modern economy than the retail franchiser and the venture capital investor. At first glance, these two types of organisations might appear to have little in common. Yet, in many ways, they possess remarkably similar objectives and engage in remarkably similar economic activity. Both the retail franchiser and the venture capitalist have capital, but are

unable, either due to a lack of knowledge of local factors, or because their time and energy are best spent elsewhere, to operate a particular firm or franchise. Both, therefore, contract with an agent who is typically liquidity constrained and who has access to private knowledge about the enterprise either because (as in franchising) he is *on the scene* or because (as in venture financing) he possesses *technical expertise*. The salient features of these contractual situations are that: (i) the agent is liquidity constrained and cannot buy the firm outright, (ii) the relationship is of a long term nature, (iii) the agent acquires private knowledge regarding certain factors influencing profitability, and (iv) the principal incentivises the agent by controlling the scale of operations. In this paper, we provide a normative analysis of the optimal dynamic contract in a general setting possessing these characteristics.

Operationally, we study an infinite-horizon discrete-time model in which the marginal cost of production evolves according to an iid 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. Since 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 enterprise 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 screening model in which the principal incentivises the agent through both instantaneous payments as well as promised future payments. The principal also manages information rents through control of the scale (or level) of operations, that is, the output of the firm, in each period.

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

- ***Sweat equity***: The optimal contract incentivises 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.
- ***Success begets Success***: Future firm size is increasing as a function of

(1) We discuss situations in which the agent possesses initial positive wealth in subsection 7.2.

current firm size. Thus, the firm's scale of operations is positively serially correlated over time.

- ***Easing of liquidity:*** Liquidity constraints ameliorate 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.

Indeed, we survey 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 new doctor joining a medical practice, a new attorney joining a law firm, a new economist joining a consulting firm, etc.

Our analysis leverages the recursive nature of the principal's problem, where the utility promised (in the form of present and future rents) to the agent, v , is the state variable. Building upon techniques recently developed by Quah (2007), we show that the optimal contract is monotone in v . This is our principal finding. Indeed, we show that *all* variables of importance are monotone in promised utility. Besides its technical significance, monotonicity permits us to interpret promised utility v as the agent's ***equity*** in the firm.² In particular,

- All elements of the menu of output choices available to the agent at any point in time are increasing in his equity.
- All elements of the corresponding menu of continuation payoffs to the agent are also increasing in his equity.
- All liquidity constraints confronting the agent attenuate as his equity increases, to the point where if he has enough equity to become a fully vested partner, then liquidity constraints disappear from that point onward.
- For all cost realizations, greater equity implies a greater likelihood of the agent becoming a partner in the next period.

(2) Such an interpretation would be more tenuous if some of the key elements of the contract were not monotone. For instance, if it were the case that output restrictions were more severe (at least for some cost realisations) at higher levels of promised utility, it is not immediately (if at all) clear how one could then regard v as equity, since greater levels of equity suggests not only that the agent is better off, but also that he faces less stringent controls.

Roughly put, these results show that greater equity comports with the agent having greater control of the firm because histories resulting in higher expected payoffs (to the agent) also correspond to greater levels of output, less stringent output controls, greater rents, more liquidity, and a higher likelihood of attaining a permanent ownership stake in the immediate future. 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.

In the next section we briefly survey the relevant literature. We introduce the model formally in section 3, and describe the recursive approach in section 4, where we also establish basic properties of the principal's value function, prove that the optimal contract has the bang-bang property, and derive a simplified version of the principal's contract design problem more amenable to analysis. In section 5 we use the simplified program to prove the monotonicity properties of the optimal contract that facilitate the interpretation of promised utility v as sweat equity. 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 levels of ownership, path dependence of the optimal contract, and the extension where the principal can fire the agent, are analysed in section 7. Section 8 contains the applications of our model mentioned above to work-to-own franchising programs and to venture capital covenants. Since it is somewhat disconnected from the rest of the paper, or at least uses sufficiently different concepts, a discussion of the key ideas underpinning the monotonicity results are deferred to section 9, with some concluding thoughts in section 10. Formal proofs and some purely technical results are relegated to the appendix.

2. *Related Literature*

This paper contributes to a growing literature on optimal dynamic incentive schemes spanning a diverse set of research areas including: social insurance (eg, Fernandes and Phelan, 2000), taxation (eg, Albanesi and Sleet, 2006), and executive compensation (eg, 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 iid over time and the state variable is taken to be the expected present value of the agent's utility under the continuation contract.

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.³ Roughly, DeMarzo and Fishman (2007) explore optimal financial contracting in a general finite-horizon CFD model which 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 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.⁴ To see this plainly, note that in the CFD models each

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

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

value of the state variable gives rise to a distinct level of optimal investment, while 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. Among other things, this means – except in the two-type case – we must employ novel methods to establish monotonicity of the entire menu of output levels in the state variable. As we argued above, this monotonicity is crucial for interpreting the state as the agent’s equity stake in the firm.⁵ While our investigation clearly touches on issues of corporate finance, our focus is rooted in questions of procurement and monopolistic screening more readily identified with industrial organization.⁶

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 incentivised only through adjustments in his future utility up to a threshold, after which he is incentivised with cash payments. The CFD papers naturally interpret this as optimal financial structure; eg, 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 bang-bang property 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.

In addition to the monotonicity of the primal contractual variables, we also prove monotonicity of the Lagrange multipliers for the liquidity constraints (one for each type of the agent). Then, defining dead-weight 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

(5) In fact, monotonicity of investment fails at low levels of the state in some of the CFD models (eg, Clementi and Hopenhayn, 2006) due to an exogenous liquidation value for the firm. The state at which the principal optimally replaces the agent in our model is endogenously determined, and monotonicity of output holds globally.

(6) See, for example, Laffont and Martimort (2002, p 86).

measure of the *price* of the constraints. Namely, dead-weight 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 maximised (where all the multipliers drop to zero and the agent achieves a vested ownership stake in the firm).

More generally, our methods, which build on Quah (2007), allow us to establish monotonicity of the contractual variables in *any* (recursive) dynamic contracting setting where the Principal’s objective is concave and supermodular in all the contractual variables, and all the constraints are linear (in the contractual variables) and increasing in the level of promised utility.⁷ We therefore provide a unified explanation of the source of monotonicity results in many recursive contracting models.

In addition to this study, there are several other recent investigations of screening in dynamic environments. For instance, Bergemann and Välimäki (2010) introduce and analyze a dynamic version of the *VCG* pivot mechanism. (In a similar vein, see Athey and Segal, 2007 and Covallo, 2008.) In two recent working papers, Pavan, Segal and Toikka (2009); Pavan, Segal and Toikka (2010) study dynamic screening in a setting in which the distribution of types may be non-stationary and agents’ payoffs need not be time-separable. They 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 quasi-linear preferences. While their analysis is illuminating, the generality of their model prohibits use of both the recursive and monotone methods that are the lynchpins of our study. Moreover, Pavan, Segal, and Toikka do not address the question of contracting for ownership in the face of liquidity constraints that is the focus of our investigation. Boleslavsky (2009) explores 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’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

(7) Monotonicity in our setting is not a straightforward application of Quah’s result. Although our objective function satisfies his conditions (over a suitably restricted domain), our constraint sets do not. Indeed, the bulk of the proof is in showing that the constraints can be transformed in such a way that the new optimisation problem (with the transformed constrained set) has the same maximisers as the original problem.

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. Although he considers a different setting and does not employ our methods, Battaglini does also find a type of monotonicity in output distortions under an optimal contract. For an initial string of reported low-demand realizations, the consumer is awarded less and less output, but nevertheless makes payments in excess of his valuation of the output (and thus receives negative rents in each period). The first time he reports high demand, however, the contract calls for efficient output for either type from that point forward. In analyzing the process of ownership acquisition, Battaglini (2005) emphasizes the role of initial and persistent private information, while we focus on the importance of transitory private information in the face of liquidity constraints.

3. The Model

Consider a setting in which a principal initially owns a business enterprise and wishes to contract with an agent to operate it. Specifically, the agent will produce output in each period $t = 0, 1, 2, \dots$. Both the principal and agent 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 $R : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is twice continuously differentiable, strictly concave, and $R(0) = 0$.⁸

The principal is *not* a bank who simply lends the agent capital. Instead, we shall suppose that the firm possesses some market power, which leads naturally to the assumption that $R'' < 0$, 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 7.2 below.

The agent's cost of producing q units of output in a given period is θq , where $\theta \in \Theta := \{\theta_1, \dots, \theta_n\}$, and $0 < \theta_1 < \dots < \theta_n < \infty$.⁹ We will frequently

(8) As long as revenue is contractible, it does not matter whether it accrues directly to the principal (who then compensates the agent for costs) or to the agent (who then delivers profits to the principal). We assume the former case in the text.

(9) 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

abuse notation and refer to $i, j \in \Theta$ rather than saying $\theta_i, \theta_j \in \Theta$. The cost parameter θ is drawn independently in each period according to the cumulative probability distribution F where we may assume, without loss of generality, that $\Pr\{\theta = \theta_i\} := f_i > 0$ for all $i \in \Theta$.

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

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

and

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

Then, implicitly define the first-best output levels by $R'(q_i^*) = \theta_i$ for all $i \in \Theta$. For future reference, note that $\infty > q_1^* > q_2^* > \dots > q_n^* > 0$; ie, first-best output is monotone decreasing in type. As always, the agent can leave at any moment in time, to an outside option worth 0 utiles.¹⁰ There are two crucial sources of friction in the model. First, the realization of the cost parameter θ in each period is observed only by the agent. Second, the agent is liquidity constrained and cannot incur a negative cash flow in any period.¹¹ 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,

$$[\text{FB}] \quad v^{\text{FB}} := \frac{1}{1 - \delta} \sum_{i \in \Theta} f_i (R(q_i^*) - \theta_i q_i^*)$$

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

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.

(10) 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.

(11) A third implicit assumption is that the agent cannot borrow sufficient funds from a bank to purchase the firm. There are numerous reasons this might be the case; eg, banks may lack the expertise needed to evaluate the profitability of the business, or they may be unable to provide the requisite brand recognition and/or proprietary methods. As mentioned above, we also suppose that the principal is not a bank.

The timing runs as follows. At the beginning of the game the principal offers the agent an infinite-horizon contract which 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. We now explore the structure of the optimal contract.

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, eg, 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

$$[\text{PK}] \quad \sum_{i=1}^n f_i(u_i + \delta w_i) = v$$

Each w_i summarizes the discounted expected future rents, while v is the expected sum of instantaneous present and discounted future rents.

Incentives: The set of incentive constraints is

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

for all $i, j \in \Theta$.

Liquidity: The agent’s liquidity constraints are simply

$$[L'_i] \quad u_i \geq 0$$

for all $i \in \Theta$. That is, when the agent reports truthfully, the monetary transfer he receives from the principal, m_i , must cover his production costs $\theta_i q_i$. 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 constraints in the future. While 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 in the firm. 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 much more cumbersome.¹²

Participation: The continuation utility w_i is the sum of expected rents, and because instantaneous rents to the agent can never be less than zero, it follows that we must include feasibility constraints that require $w_i \geq 0$ for all i . Thus, 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 $P : \mathbb{R}_+ \rightarrow \mathbb{R}$ that satisfies*

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

subject to: promise keeping (PK), incentive compatibility (C_{ij}), liquidity (L'_i), and feasibility $q_i \geq 0$ and $w_i \geq 0$ for all $i \in \Theta$. 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 (MR₀), namely the assumption that $R'(0) = \infty$, ensures $P'(0) = \infty$. In

(12) See Edmans et al. (2010) for a novel use of ‘incentive accounts’ in the context of executive compensation.

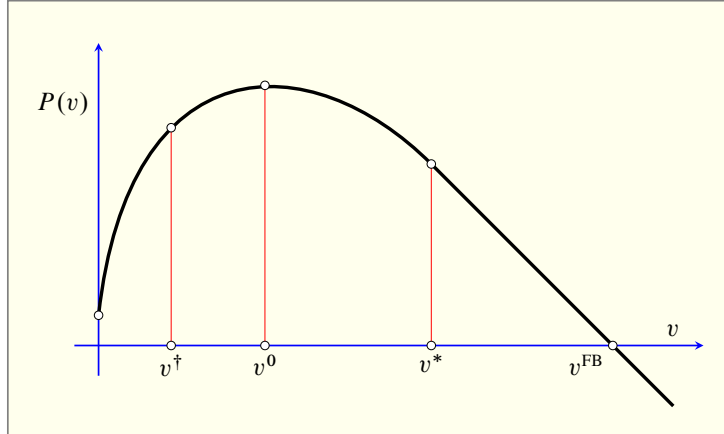


Figure 1: *Principal's Value Function*

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 that the principal initially stakes the agent upon signing the contract.

Note, however, that social surplus (ie, firm value) $P(v) + v$ is maximized at any $v \geq v^*$.¹³ 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.

Proposition 4.1. For any optimal contract (u, q, w) , incentives are provided purely through adjustments in the agent's equity whenever his stake in the franchise 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, for all v , $w_i(v) \leq v^*$ and

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

Proposition 4.1 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

(13) This follows since $P(v) + v$ is continuously differentiable, and has derivative $P'(v) + 1$, which is strictly positive for all $v < v^*$, and is 0 for all $v \geq v^*$.

agent does not reach $v = v^*$ in the next period, it must be that the agent earns no instantaneous rents, but instead is incentivised 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.

In order to obtain a sharper characterization of an optimal contract the following definitions are very useful.

Definition 4.2 (Monotonicity in Type and Equity). Output is said to be *monotonic in type* if for all $v \geq 0$,

$$[M_i] \quad q_i(v) \geq q_{i+1}(v)$$

for all $i = 1, \dots, n-1$. Output is said to be *monotonic in equity* if for all $i \in \Theta$,

$$v' > v \text{ implies } q_i(v') \geq q_i(v)$$

Analogous definitions apply for rent $u_i(v)$ and promised utility $w_i(v)$.

In static mechanism design, inequalities analogous to (M_i) are often referred to as *implementability conditions*. In order to establish our key result that the optimal contract is monotone in equity, it is necessary to reformulate the principal's program in a simpler way (with fewer constraints and choice variables) that is more amenable to analysis. To this end, first consider the binding version of the upward adjacent incentive constraints that say the agent must be indifferent between reporting his true marginal cost and one level higher:

$$[C_i] \quad u_i + \delta w_i = u_{i+1} + \delta w_{i+1} + \Delta_i q_{i+1}$$

for all $i = 1, \dots, n-1$, where $\Delta_i := \theta_{i+1} - \theta_i$.

The following lemma establishes a result familiar from static mechanism design that the large set of incentive constraints (C_{ij}) may be replaced by a much smaller set, namely (M_i) and (C_i) .

Lemma 4.3. If output is monotonic in type (M_i) and the upward adjacent incentive constraints bind (C_i) , then all incentive constraints (C_{ij}) are satisfied. Moreover, there exists a maximal rent optimal contract (u, q, w) in which (M_i) and (C_i) hold, and in any such contract, instantaneous rent and promised utility are also monotonic in type.

Next, the following lemma uses (PK) and (C_{*i*}) to derive a key expression for the agent's current payoff.

Lemma 4.4. In any optimal contract, the agent's payoff satisfies

$$[\text{U}_i] \quad u_i + \delta w_i = v - \sum_{j=1}^{n-1} F_j \Delta_j q_{j+1} + \sum_{j=i}^{n-1} \Delta_j q_{j+1}$$

for all $i = 1, \dots, n$. Moreover, (U_{*i*}) implies (PK) and (C_{*i*}).

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).

The equations (U_{*i*}), which imply (PK) and (C_{*i*}) 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

$$[\text{L}_i] \quad \sum_{j=1}^{n-1} F_j \Delta_j q_{j+1} - \sum_{j=i}^{n-1} \Delta_j q_{j+1} + \delta w_i \leq v$$

for all $i \in \Theta$. Using this version of the liquidity constraints and substituting (PK) directly into the principal's objective yields the following intuitive result.

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

$$[\text{VF}] \quad P(v) = \max_{(q,w)} \sum_i f_i \left[\left(R(q_i) - \theta_i q_i \right) + \delta \left(P(w_i) + w_i \right) \right] - v,$$

*subject to monotonicity in output (M_{*i*}), liquidity (L_{*i*}), and feasibility $q_n \geq 0$ and $w_n \geq 0$. 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 (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, involving n^2 fewer constraints and n fewer choice variables. This version of the program 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), while the term $\sum_i f_i \delta [w_i + P(w_i)]$ is 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 social surplus (ie, the value of the firm) net of any expected information rents.

Most importantly, the version of the principal's problem presented in Theorem 2 is also more amenable to analysis. In particular, in the next section we use this version of the problem to establish our key result that all contractual variables are monotonic in v ; ie, that v may be interpreted as the agent's equity in the firm.

5. Monotone 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, hence first order conditions are both necessary and sufficient. Let λ_i be the Lagrange multiplier associated with the liquidity constraint (L_i) and μ_i the Lagrange multiplier of the implementability constraint $q_i \geq q_{i+1}$ with $q_{n+1} = 0$ for all i . Since $P'(0) = \infty$, we will ignore the constraint $w_n \geq 0$ whenever $v > 0$. Since $P'(v) = -1$ for $v \geq v^*$, we can also ignore the constraint $w_i \leq v^*$. For the moment, let us ignore the constraint (M_1), that is, the constraint (M_i) for $i = 1$. (Lemma 5.2 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 for the *best* type, ie, there is no distortion at the top, which holds here for all $v \geq 0$. The first order condition for q_i , for any $i > 1$, is

$$\begin{aligned} R'(q_i) - \theta_i &= \frac{\Delta_{i-1}}{f_i} \sum_{k=1}^n \lambda_k [F_{i-1} - \mathbb{I}\{k < i\}] - \frac{1}{f_i} (\mu_i - \mu_{i-1}) \\ \text{[FO}q_i] \quad &= \frac{\Delta_{i-1}}{f_i} [F_{i-1} \Lambda_n - \Lambda_{i-1}] - \frac{1}{f_i} (\mu_i - \mu_{i-1}) \end{aligned}$$

where $\Lambda_k = \sum_{j=1}^k \lambda_j$ for all k .

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

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

Finally, the envelope condition is

$$[\text{Env}] \quad P'(v) = -1 + \Lambda_n$$

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

Lemma 5.1. The critical level of equity is

$$[\text{Vest}] \quad v^* = \frac{1}{1-\delta} \sum_{j=1}^{n-1} F_j \Delta_j q_{j+1}^*.$$

Hence, v^* is the present value of receiving expected rents from efficient production (that is, output without distortions) in perpetuity. Moreover, since $P'(v) = -1$ 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.

In order to establish our principal result below, it will be useful to show that the optimal contract does not involve production greater than the socially optimal amount. This is now stated formally.

Lemma 5.2. In any maximal rent optimal contract, the agent never produces more than first-best output, that is $q_i(v) \leq q_i^*$ for $i = 1 \dots, n$ and $v \in [0, v^*]$.

This result indicates that we can, without loss of generality, restrict attention to domains for the choice variables wherein $q \in [0, q_1^*] \times \dots \times [0, q_n^*]$ and $w \in [0, v^*]^n$. Notice that the principal's objective function, $\sum_i f_i \left[\left(R(q_i) - \theta_i q_i \right) + \delta \left(P(w_i) + w_i \right) \right]$, which is simply the expected social surplus, is strictly increasing, supermodular and concave over this restricted domain. This observation enables us to prove that in a maximal rent optimal contract, output and promised utility must be monotonic in the state, allowing us to interpret v as the agent's equity in the firm. Moreover, monotonicity allows us to characterize not only the long-run dynamics of the contractual relationship but to analyze short-run changes as well.

Theorem 3. *The optimal contract is monotone in equity. That is, for all $i \in \Theta$ and all $v, v' \in [0, v^*]$, $v > v'$ implies $(q_i(v), w_i(v)) \geq (q_i(v'), w_i(v'))$.*

Figure 2 illustrates the monotonicity of the quantities and continuation utilities in the state variable, promised utility, when there are three types. (The

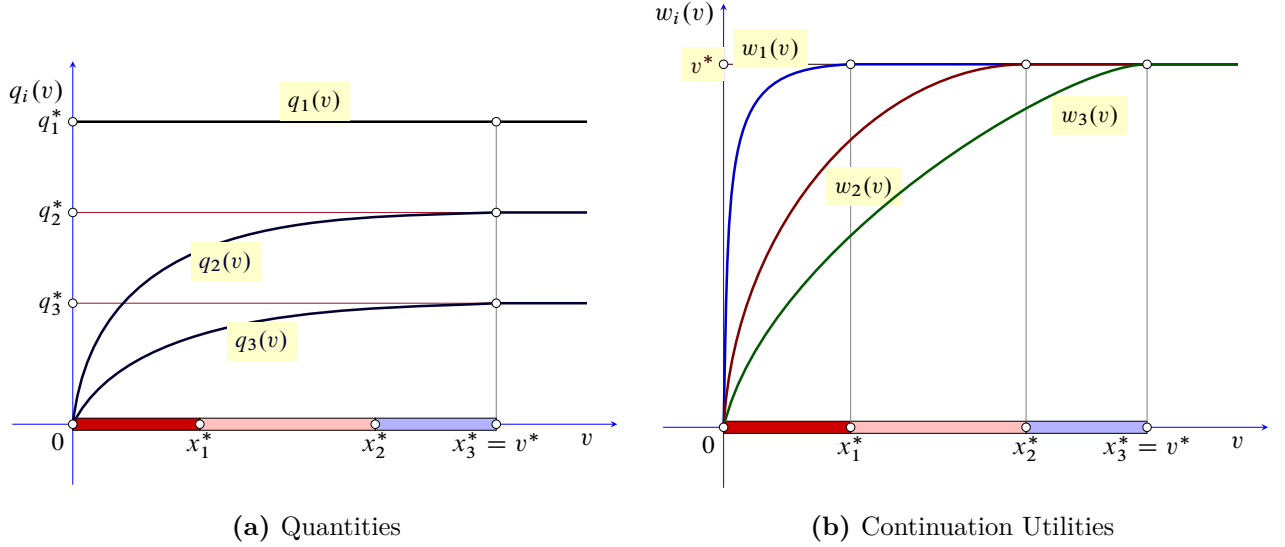


Figure 2: *Monotone Contracts*

cutoff points x_1^* , x_2^* and x_3^* are discussed in proposition 5.3 below.) While this result seems natural, establishing monotonicity is often problematic in dynamic contracting models with more than two types. The proof uses results from Quah (2007), and is in the appendix. Since the main ideas underlying the proof are sufficiently removed from incentive theory, we defer a sketch and discussion of the intuition for the interested reader to section 9 below.

Theorem 3 says that the degree of distortion the contract imposes on the agent's output decreases as his stake, ie, his equity, in the firm grows. Hence, increasing v results in firm growth, while decreasing v results in contraction. At $v = 0$, the contract calls for virtual shutdown (this is lemma A.1 in the appendix): $q_1(0) = q_1^*$ and $q_i(0) = 0$ for $i = 2, \dots, n$.¹⁴ As v increases, output restrictions are relaxed until $v = v^*$, at which point the contract calls for efficient production for all cost realizations: $q_i(v^*) = q_i^*$ for $i = 1 \dots, n$. The agent's promised future utility levels are similarly increasing in sweat equity. At $v = 0$, he *never* receives any rents, implying $w_i(0) = 0$ for $i = 1, \dots, n$. Again, as v increases, promised future utility levels rise monotonically until $v = v^*$, when the agent becomes a **vested partner** with a permanent ownership stake, with $w_i(v^*) = v^*$ for $i = 1, \dots, n$. 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.

(14) To be sure, the assumption $R'(0) = \infty$ implies that the limiting case of $v = 0$ and the concomitant virtual shutdown never occurs on any finite sample path; ie $w_n(v) > 0$ for all $v > 0$.

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 (see Proposition 5.3 immediately below) leading to less strict output controls and still higher levels of promised future utility.

The monotonicity of the optimal contract also reveals information about the Lagrange multipliers. As usual, the multipliers can be thought of as the marginal cost of violating a constraint – in this case, the liquidity constraints. The following proposition collects some useful facts.

Proposition 5.3. The Lagrange multipliers (λ_i) satisfy the following:

- (a) For each v , $\lambda_1(v)/f_1 \leq \dots \leq \lambda_n(v)/f_n$.
- (b) For each i , $\lambda_i(v)$ is continuous and decreasing in v , with $\lambda_i(v^*) = 0$ and $\lim_{v \rightarrow 0} \lambda_i(v) = \infty$.
- (c) There exist $0 < x_1^* \leq \dots \leq x_n^* = v^*$ such that $v < x_i^*$ implies $\lambda_i(v) > 0$, and $v \geq x_i^*$ implies $\lambda_i(v) = 0$. Moreover, $x_1^* < v^*$.

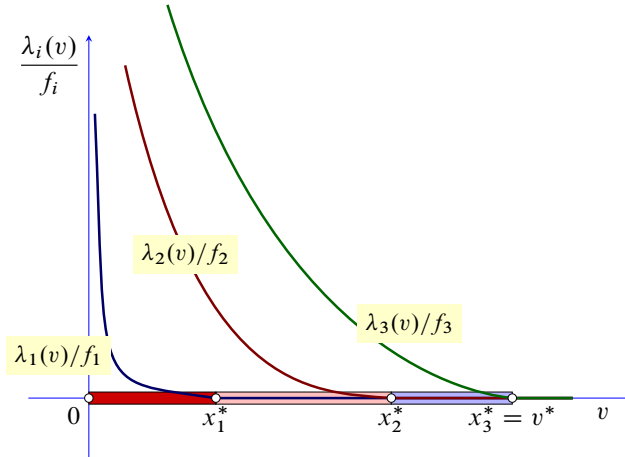


Figure 3: *Cost of Liquidity Constraints*

Figure 3 illustrates the monotonicity of the Lagrange multipliers, as well as the cutoff points, for the case of three types. The envelope condition (Env) and concavity of the value function imply that the sum of the Lagrange multipliers of the liquidity constraints Λ_n must be decreasing. The proposition above is a refinement of that observation. In particular, it says that at each $v < v^*$, there is a subset of the constraints (L_i) that bind, and that this subset is decreasing in

v . The fact that each λ_i is decreasing in v has an important interpretation. As the agent acquires a greater stake in the firm, the cost of violating his liquidity constraints falls. The intuition is that when the agent acquires more equity, then the contract optimally reduces distortions in order to generate more joint surplus. Moreover, equity is particular to the relationship between the principal and the agent, and cannot be traded with anyone outside this interaction. Put differently, equity is a ***relation-specific tradeable asset***, and larger amounts of it alleviate liquidity concerns.

The cutoff points x_i^* for the Lagrange multipliers have another useful consequence. For each v , we may define the probability, $G(v)$, that the agent will become the owner of the firm in the *next period*. For each $v \geq x_1^*$, $G(v) > 0$. If, for example, $v \in [x_k^*, x_{k+1}^*)$, the agent is potentially one step away from obtaining a permanent ownership stake in the firm. Specifically, liquidity constraints 1 through k do not bind at this point, so if the agent reports a cost realization in this range, ie reports θ_j where $j \leq k$, his sweat equity will be v^* in the ensuing period and forever hence. Thus, for a $v \in [x_k^*, x_{k+1}^*)$, $G(v) = F_k = \sum_{i \leq k} f_i$. It follows from proposition 5.3 that $G(v)$ is monotone increasing in v ; indeed, it is a step function. Thus, with a greater stake in the firm, the agent is ever closer, in a precise sense, to becoming a vested partner in the firm.

6. *Dynamics*

We next derive both short- and long-run dynamics of the contractual relationship. We begin with a straightforward, but important, consequence of our definitions, which reveals something about the long-run behaviour of the relationship. The optimal contract induces a process $P'(\cdot)$ that is a martingale. To see this, consider an increase in v by one unit. This can be achieved by increasing all the w_i 's by $1/\delta$. The cost of this to the principal is $\sum_i f_i [1 + P'(w_i)]$. As Thomas and Worrall point out, by the envelope theorem, this is locally optimal, and hence is equal to $P'(v)$. From a slightly different point of view, notice that $P'(v) = -1 + \Lambda_n = \sum_i f_i P'(w_i)$, where the first equality is the envelope condition (Env), and the second equality is obtained by summing the first order conditions for w_i (FO w_i).

An important consequence of the martingale property of P' and the monotonicity of the optimal contract is that a shock of $\theta = \theta_1$ is necessarily *good*, in the sense that the continuation values of sweat equity $w_1 > v$, while a shock of $\theta = \theta_n$

is unambiguously *bad*, wherein $w_n < v$. More generally, we have the following.

Proposition 6.1. In the optimal maximal rent contract, for all $v \in (0, v^*)$, we have $P'(w_n) > P'(v) > P'(w_1)$. Moreover, $w_1(v) > v > w_n(v)$.

This captures the short-run consequences of good and bad shocks. To see the intuition, suppose, for simplicity, that P is strictly concave on $(0, v^*)$. Since P' is a martingale, if the proposition were not true, it would follow that $P'(w_i) = P'(v)$ for all $i \in \Theta$, which implies (if P is strictly concave) that $w_i(v) = v < v^*$ for all $i \in \Theta$. But proposition 4.1 also requires that for such a v , $u_i(v) = 0$, which violates promise keeping (PK), and by incentive compatibility, would require that $q_i = 0$ for all $i > 1$. Therefore, incentive compatibility and promise keeping force the agent to spread out continuation utilities. This is unsurprising, since the role of continuation utilities is precisely to aid in incentive compatibility, by allowing the principal to raise instantaneous surplus, without raising the cost of doing the same. While 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 is a *vested partner* if his equity level reaches v^* .

Theorem 4. *The martingale P' converges almost surely to $P'_\infty = -1$. Thus, the agent becomes a vested partner with probability 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 6.1 above and the continuity of the contract in v .

The economic 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. Recall that total lifetime utility for type i is $u_i + \delta w_i$. Clearly, for any type $i < n$, the total (lifetime) expected rent is $u_i + \delta w_i > 0$, that is, lifetime expected utility is strictly positive. Therefore, the principal faces the choice of either granting rents in the present, via u_i , or relegating them to the future, via w_i . Notice that any instantaneous rent to the agent is spent outside the relationship and therefore does not affect the principal. However, if the principal chooses to provide the necessary

incentives via continuation payoffs w_i , this has the benefit of increasing liquidity in the following period, which is useful for the principal, since it allows her to raise instantaneous surplus in the subsequent period. (Recall that a larger v means a larger feasible set, and output is increasing in v ; see Theorem 3 above.) It is this desire to keep the agent's rents within the relationship for as long as possible that causes the principal to back load payments, and consequently causes v to converge to v^* along almost all sample paths.

7. Discussion and Extensions

7.1. Social Cost of Liquidity Constraints

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 $S(v) = v^{\text{FB}} = \frac{1}{1-\delta} \sum_i f_i [R(q_i^*) - \theta_i q_i^*]$ for all $v \geq v^*$.¹⁵ Moreover, by the envelope condition (Env), we see that $S'(v) = P'(v) + 1 = \Lambda_n(v)$. Therefore, Λ_n measures the marginal social cost of illiquidity (which is decreasing in v). Hence, for any $v < v^*$, the **dead-weight loss** generated by an optimal contract is

$$v^{\text{FB}} - S(v) = \int_v^{v^*} \Lambda_n(x) dx.$$

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 dead-weight loss is consequently nil.

(15) To see this, recall that for all i , $q_i(v^*) = q_i^*$ and $w_i(v^*) = v^*$. Substitution into (VF) then yields $P(v^*) + v^* = \frac{1}{1-\delta} \sum_i f_i [R(q_i^*) - \theta_i q_i^*] = v^{\text{FB}}$, and hence, $S(v^*) = v^{\text{FB}}$. 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, $v^* > v^{\text{FB}}$, the latter being a condition depending on the primitives of the model; recall that (Vest) says $v^* = \frac{1}{1-\delta} \sum_{j=1}^{n-1} F_j \Delta_j q_{j+1}^*$.

7.2. The Path to Ownership

When exploring firm ownership, it is useful to distinguish between two paradigms as discussed by Bolton and Scharfstein (1998). 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, pioneered by Grossman and Hart (1986) and Hart and Moore (1990), 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. Nevertheless, it is possible to include an option for the principal to transfer control of the productive assets to the agent in certain situations, thereby permitting us to regard firm ownership in the Grossman-Hart-Moore sense as well.

It follows from our assumption on the absence of fixed costs, ie, $R(0) = 0$, that first-best profit is nonnegative in every state.¹⁶ Recall that under a maximal rent optimal contract, the agent's equity is capped at v^* and he is incentivised with cash from that point forward. However, once the agent attains equity of v^* , all output restrictions are eliminated, and both the principal and agent are indifferent between providing incentives with cash or further equity adjustments.

Suppose now that $v^* < v^{\text{FB}}$, so that $P(v^*) = v^{\text{FB}} - v^* > 0 = P(v^{\text{FB}})$, and consider a contract under which the agent continues to be incentivised with sweat equity until $v = v^{\text{FB}}$. Indeed, if the agent attains v^* , then he will move monotonically to v^{FB} because (as is easily seen from U_i) $w_n(v) = \frac{v - (1 - \delta)v^*}{\delta} \geq v$ for $v \geq v^*$. Once $v = v^{\text{FB}}$, the principal owes the agent cash flows equal to the first-best value of the enterprise; ie $P(v^{\text{FB}}) = 0$. While this commitment is formally financial, it is easy to imagine the principal simply transferring control of the productive assets to the agent and terminating the contractual relationship at this point.

Our model is somewhat less well suited to analyze transfer of asset ownership in the case when $v^* > v^{\text{FB}}$. In this instance, output distortions are not completely eliminated until the principal owes the agent cash flows in excess of the first-best value of the firm. If $v = v^*$, one can imagine the principal paying the agent a termination fee of $v^* - v^{\text{FB}}$ and transferring control of the firm to him. The trouble is that if relinquishing control is an option formally available to the principal, then she should exercise it before $v = v^*$ because $v^* > v^{\text{FB}}$ implies $0 > P(v^{\text{FB}}) > P(v^*)$. The

(16) If this were not the case, then the agents lack of liquidity would prohibit full ownership of the productive assets.

principal could eliminate the negative part of the value function by relinquishing control to the agent at the point when $P(v) = 0$. Of course, this would impact her incentives to distort output at lower equity levels as well as the value function itself. This, however, would not alter the qualitative nature of the optimal contract.

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

Corollary 7.1. Suppose $v^* \leq v^{\text{FB}}$ and 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 $S(v^0) < v^{\text{FB}}$.
- (b) If $v^0 < y < v^*$, then the agent surrenders y to the principal and receives initial equity y . Initial welfare is $S(y) \in (S(v^0), v^{\text{FB}})$.
- (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. Since $S(y)$ 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; ie, liquidity constraints never bind and the contract is first-best. Finally, while it is common wisdom that incentive problems can be eliminated by *selling the firm to the agent*, note that if $v^* < v^{\text{FB}}$, then it is not necessary to sell the entire firm to the agent because the first-best outcome obtains if his equity position is v^* .

7.3. Hiring and Firing

Suppose there is an infinite pool of identical agents, but that the principal can only contract with one agent at a time. The principal may, however, *fire* the current agent and replace him with a new one. If the principal fires an agent, then she must make a severance payment to him equal to the current level of sweat equity.

Proposition 7.2. There exists a critical level of equity $v^\dagger \in (0, v^0)$ such that it is optimal to fire the agent if sweat equity falls below v^\dagger (also see figure 1).

Lemma C.2 in the appendix 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, Theorem 2.4 of Steele, 2001) provides a bound for this probability, wherein, the probability that $P'(v) \geq C$ is less than $1/(1 + C)$.

To formally incorporate the option to replace an agent it is necessary to introduce a new value function $Q(v)$. For any function $Q : \mathbb{R}_+ \rightarrow \mathbb{R}$ bounded above, let $v_Q^0 \in \arg \max_x Q(x)$. Now let Q be the unique function that satisfies

$$Q(v) = \max \left[Q(v_Q^0) - v, \max_{(q,w)} \mathbf{E} \left[\left(R(q_i) - \theta_i q_i \right) + \delta \left(Q(w_i) + w_i \right) \right] - v \right]$$

s.t. $(M_i), (L_i), q_n \geq 0$ and $w_n \geq 0$

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 since it is exercised with positive probability. Hence, for any $v < v_Q^0$ such that firing is not optimal, $Q(v)$ is increasing. Since $Q(v_Q^0) - v$ is decreasing, there exists a state v^\dagger such that it is optimal to fire the agent if $v < v^\dagger$ and to retain him if $v > v^\dagger$.

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.

7.4. Path Dependence

The maximal rent optimal contract specifies (q, w) as a function of equity, v . Therefore, the evolution of (q, w) depends on the evolution of v . Typically, the evolution of v along any sample path will depend on the order of shocks – and this is true of models of dynamic contracting in general. Nevertheless, there is a very strong form of path dependence that holds in our model. There are two reasons for this: Firstly, 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.¹⁷

More specifically, for any initial $v^{(0)} = v \in (0, v^*)$, there exists an integer $\tau > 0$ such that if the agent repeatedly receives θ_1 shocks over τ periods (which happens with strictly positive probability), he will reach v^* , ie he will have $v^{(\tau)} = v^*$, in τ periods (where $v^{(k)}$ represents the value of v in period k). This relies on two observations. 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, $P'(v^{(\tau)}) < \gamma$ (this is lemma C.2 in the appendix). Of course, the sample path where θ_1 is repeated τ times has strictly positive probability. The second observation is that there exists an $\varepsilon > 0$ such that $\lambda_1(v) = 0$ for all $v \in (v^* - \varepsilon, \infty)$. But this follows from part (c) of proposition 5.3. In sum, we have shown that from any initial level of sweat equity, the agent will reach v^* with positive probability in a finite number of periods.

Therefore, in an arbitrary sample path, the order of the occurrences of shocks matters greatly. In any sample path where θ_1 occurs sufficiently often, the agent strictly prefers to have all the θ_1 shocks in the beginning, since 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), where it is shown that when an agent with a private endowment has 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.

8. Applications

In order to focus on the fundamental economic forces at work, the model analysed 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

(17) This second 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.

discuss the 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 organisational form, especially in retailing. According to Blair and Lafontaine (2005, pages 8–13), 34% of US retail sales in 1986 (almost 13% of GDP) derived from franchised outlets. Estimates on the number of US franchisers vary widely, but listings in directories suggest a figure between 2,500 and 3,000. 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, page 97) 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. While details of sweat equity arrangements vary across franchisers, Quiznos’ *Operating Partner Program* is broadly representative, enabling experienced managers to receive financing from the parent company for all but \$5,000 of the up-front 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.¹⁸

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 . . .’

(18) See Liddle (2010).

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, HBS (2000), a case study by Harvard Business School, 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’. VC contracts are very complex legal instruments, providing investors with numerous control and liquidation rights. VCs typically demand representation on the board of directors and often play an active role in the day-to-day operation of the fledgling company.

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

We find that venture capital financing allow VCs to separately allocate cash flow rights, board rights, voting rights, liquidation rights, and other control rights. These rights are often contingent on observable measures of financial and non-financial performance. 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.

While our stylized model cannot directly address the plethora of sophisticated contingencies and control rights found in typical VC contracts, Kaplan and Stromberg's findings are consistent with the monotonicity of the optimal mechanism in both type and equity. Specifically, v , or sweat equity, is a summary statistic of past performance, and greater sweat equity leads to reductions in output restrictions, less stringent liquidity constraints, and eventually to agent ownership, while lower sweat equity results in more stringent restrictions and liquidity constraints, and ultimately to replacement of the agent. In fact, the founders of poorly performing ventures are often 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. Monotone Contracts: Another Look

In this section, we provide some intuition and sketch the proof to our critical result, Theorem 3, which establishes that the maximal rent contract is monotone in the agent's equity v . In dynamic contracting models, monotonicity in the state variables is often difficult to establish. As noted by Stokey, Lucas and Prescott (1989, p 86), the problem is the same as establishing monotone comparative statics of constrained optimization problems. Seminal work by Milgrom and Shannon (1994) and Topkis (1998) provide the most general conditions under which a set of maximisers is monotone in the parameters of the objective function. These methods, however, often cannot be applied to establish monotonicity in parameters appearing in the constraints because constraint sets are often not sublattices (eg, budget sets) and hence cannot be ranked in the strong set order (see section D.1 in the appendix below for a definition of the strong set order).

However, models in economics typically have more structure (for instance, convexity assumptions) beyond the lattice theoretic structure assumed by Milgrom-Shannon/Topkis. In particular, in a seminal study, Quah (2007) shows, among other things, the following result which is extremely useful to us.

Theorem 5. *Let X be a convex sublattice of \mathbb{R}^n , $U : X \rightarrow \mathbb{R}$ a concave and supermodular function, and $g : X \rightarrow \mathbb{R}$ a continuous, increasing, submodular and convex function. Suppose also that for each value of $k \in \mathbb{R}$, there is a unique solution to $\max_{x \in g^{-1}((-\infty, k])} U(x)$. Then, $k > k'$ implies $\arg \max_{x \in g^{-1}((-\infty, k])} U(x) \geq \arg \max_{x \in g^{-1}((-\infty, k'])} U(x)$.*

The above theorem follows from Corollary 2(ii) of Quah (2007). The chief challenge in applying Quah's techniques in our setting is in verifying that the contract design program belongs to the class of problems amenable to analysis. The first step is provided in lemma 5.2, which says that in any maximal rent optimal contract, the agent never produces more than first-best output, $q_i(v) \leq q_i^*$ for $i = 1 \dots, n$ and $v \geq 0$. (Recall that for $v \geq v^*$, output is always optimal.)

This allows us to restrict attention to a domain where $q \in \times_{i=1}^n [0, q_i^*]$ and $w \in [0, v^*]^n$. It is easily seen that this domain is both convex and a sublattice of \mathbb{R}^{2n} . It is also easy to see that when restricted to this domain, the principal's objective function, $\sum_i f_i \left[(R(q_i) - \theta_i q_i) + \delta (P(w_i) + w_i) \right]$, is strictly increasing, supermodular and concave.

Next, observe that the liquidity constraints (L_i) can be written, in vector form, as

$$Aq + \delta w \leq v \mathbf{1}$$

where $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^n$, and A is the following $n \times n$ matrix

$$\begin{bmatrix} 0 & \Delta_1(F_1 - 1) & \Delta_2(F_2 - 1) & \cdots & \Delta_{n-1}(F_{n-1} - 1) \\ 0 & \Delta_1 F_1 & \Delta_2(F_2 - 1) & \cdots & \Delta_{n-1}(F_{n-1} - 1) \\ 0 & \Delta_1 F_1 & \Delta_2 F_2 & \cdots & \Delta_{n-1}(F_{n-1} - 1) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \Delta_1 F_1 & \Delta_2 F_2 & \cdots & \Delta_{n-1} F_{n-1} \end{bmatrix}$$

Let a_i be the i th row of A and consider the function

$$g(q, w) := \max\{a_1 q + \delta w_1, \dots, a_n q + \delta w_n\}$$

Evidently, the liquidity constraints (L_i) are satisfied if and only if $g(q, w) \leq v$. Hence, if $g(q, w)$ is increasing, convex, and submodular, then we can invoke Quah's

result to establish that output and promised utility are increasing in equity. The problem is that negative elements in the matrix A imply that g is neither increasing nor necessarily submodular. However, in appendix D, we show that we can replace A by a different matrix \hat{A} possessing nonnegative elements without changing the solutions to the original problem. Appealing to results in Topkis (1998), we show that the function

$$\hat{g}(q, w) := \max\{\hat{a}_1 q + \delta w_1, \dots, \hat{a}_n q + \delta w_n\}$$

is increasing, convex, and submodular, which allows us to invoke Quah's result, proving Theorem 3. The results in appendix D are somewhat cumbersome (especially notationally), but we can yet provide some intuition for the result. To do so, we shall consider an abstract programming problem, with the same mathematical structure as the problem considered in Theorem 3 above.

Let $U : \mathbb{R}_+^2 \rightarrow \mathbb{R}$ be a continuous, strictly increasing, concave function, and consider the constrained optimisation problem

$$[\text{G}] \quad \max U(x) \quad \text{subject to} \quad g(x) \leq v$$

where $g(x) = \max[\frac{1}{2}x_1 + \frac{1}{2}x_2, 2x_1 - x_2]$. Notice that the constraint set $D_v := \{x \in \mathbb{R}_+^2 : g(x) \leq v\}$ can be written as $D_v = \{x \in \mathbb{R}_+^2 : Ax \leq v\mathbf{1}\}$, where A is the matrix $\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 2 & -1 \end{bmatrix}$. It is easily seen that the function g is not increasing. It is also not clear if the function g is submodular.¹⁹ On the other hand, the constraint set is easily described as $D_v = \text{conv}\{(0, 2v), (v, v), (\frac{1}{2}v, 0), (0, 0)\}$, is the convex hull of a set of four points.

Consider now, for each v , the expanded constraint set given by $E_v := \text{conv}\{(0, 2v), (v, v), (v, 0), (0, 0)\}$, which clearly contains D_v .²⁰ Evidently, we may write $E_v = \{x \in \mathbb{R}_+^2 : \hat{g}(x) \leq v\}$, where $\hat{g}(x) = \max[\frac{1}{2}x_1 + \frac{1}{2}x_2, x_1]$, or equivalently, as $E_v = \{x \in \mathbb{R}_+^2 : \hat{A}x \leq v\mathbf{1}\}$, where \hat{A} is the matrix $\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{bmatrix}$. This allows us to study the alternate optimisation problem

$$[\text{H}] \quad \max U(x) \quad \text{subject to} \quad \hat{g}(x) \leq v$$

where \hat{g} is defined above.

(19) In particular, it is not clear how we would establish submodularity if g were the minimum of an arbitrary, but finite, set of linear constraints.

(20) We shall explain our construction of E_v momentarily.

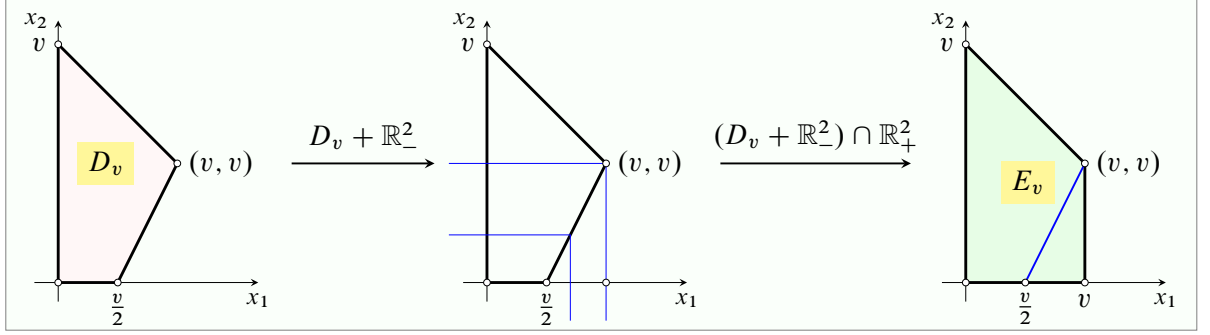


Figure 4: *Transforming the Constraint Set*

Since U is strictly increasing, it is easy to see that any solution to (H) is automatically a solution to (G), and vice versa. Therefore, without loss of generality, we may solve problem (H). But we may also show (as we do in appendix D) that \hat{g} is submodular. It is clear that \hat{g} is continuous, increasing, and convex. Therefore, we have met all the sufficient conditions of corollary 2(ii) of Quah (2007), to establish the monotonicity of the maximisers in v .

We end with a few observations. Notice first that for any $v \geq 0$, $E_v := (D_v + \mathbb{R}_+^2) \cap \mathbb{R}_+^2$, where $D_v + \mathbb{R}_+^2 := \{x + y : x \in D_v, y \in \mathbb{R}_+^2\}$. A generalisation of this construction allows us to expand the constraint set in the contracting problem, without affecting the set of maximisers. A second observation is for each $v \geq 0$, $D_v = vD_1$ and $E_v = vE_1$. This is no coincidence, and we use the positive homogeneity of the constraint set to construct the function \hat{g} for the general problem.

A final comment is in order. In the original contracting problem, there are additional monotonicity (or implementability) constraints on the q 's, which require that for any v and i , $q_i(v) \geq q_{i+1}(v)$. It is easy to see that in the abstract problem considered above, we could impose the constraint $x_1 \geq x_2$, without any change to either the construction of the function \tilde{g} or the conclusion. In appendix D.1, we also provide some intuition for Quah's result.

10. Conclusion

In this paper 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 prove that the optimal contract is monotone in promised utility, facilitating the interpretation of the state variable as the agent's equity in the firm. In particular, we show that in each state, output increases, distortions decrease, liquidity constraints matter less, and the agent's probability of achieving ownership in the immediate future increases in the agent's level of equity.

We establish a bang-bang property of an optimal contract wherein the agent is incentivised only through adjustments to his equity until achieving a critical level, after which he may be incentivised through cash payments. We can, therefore, interpret the incentive scheme as a sweat equity contract, where all rent payments are back loaded. 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, ie, 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* reports leading to higher levels of sweat equity and *bad* 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; ie, 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 incentivised primarily through equity adjustments. Moreover, good outcomes lead to less stringent controls by the franchiser/VC and increased autonomy by the manager, while bad outcomes have the reverse effects.

We believe that the monotone methods employed in this investigation can be fruitfully applied in numerous other settings of dynamic incentives including: regulation, taxation, and procurement. A particularly appealing avenue for future research involves application of the approach to analyse *frequent customer* programs in a setting where liquidity constraints imply a bound on payments rather than on

rents. Also, we conjecture that the methods we use can be generalised in several directions such as continuous types and Markovian shocks. Finally, our principle methodological contribution, adapting the techniques of Quah (2007) to establish the monotonicity of the policy function (in the state variable) of a dynamic program, may be of independent interest beyond the model presented here.

Appendices

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, sweat equity or promised utility v , 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

$$\frac{1}{1-\delta} \sum_{i=1}^n f_i [R(q_i^*) - \theta_i q_i^*] - v$$

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

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

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

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

be endowed with the sup metric, which makes it a complete metric space. Let $\Gamma_0(v)$ be the set of (u, q, w) that satisfy the constraints $q_i \geq 0$, (L_i) , (PK) , (C_{ij}) , and $w_i \geq 0$ for all i, j . Define the operator $\mathbb{T} : \mathcal{F} \rightarrow \mathcal{F}$ as

$$\begin{aligned} (\mathbb{T}Q)(v) &= \max_{(u_i, q_i, w_i)} \sum_i f_i \left[(R(q_i) - \theta_i q_i) - u_i + \delta Q(w_i) \right] \\ &\text{s.t. } (u, q, w) \in \Gamma_0(v) \end{aligned}$$

for each $Q \in \mathcal{F}$. Since $\Gamma_0(v)$ is compact for each v , the maximum is achieved for each v . Moreover, by the bounds established earlier, it is easily seen that $\mathbb{T}Q \in \mathcal{F}$.

Next, define

$$v^b := \frac{1}{1-\delta} \sum_{j=1}^{n-1} F_j(\theta_{j+1} - \theta_j) q_{j+1}^*$$

and let us assume that $Q \in \mathcal{F}$ is such that $Q'(v) = -1$ for all $v \geq v^b$. Consider the relaxed problem

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

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, $w_i = v^b$. By letting

$$u_i(v) := v - \delta v^b - \sum_{j=1}^{n-1} F_j(\theta_{j+1} - \theta_j) q_{j+1}^* + \sum_{j=1}^{n-1} (\theta_{j+1} - \theta_j) q_{j+1}^*$$

we see from (U_i) above that (PK) and $(C_{i,i+1})$ hold with equality, so that all the constraints, including liquidity, are satisfied. Therefore, the contract $(u(v), q_i^*, w_i = v^b) \in \Gamma_0(v)$, and is feasible, and is therefore a solution to the original constrained problem. In particular, for any $Q \in \mathcal{F}$ that is linear, with slope -1 for $v \geq v^b$,

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

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

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

that is, $(\mathbb{T}Q)'(v) = -1$ for all $v \geq v^b$.

By a variation of Theorem 4.6 of Stokey, Lucas and Prescott (1989), we see that the operator \mathbb{T} has a unique fixed point in \mathcal{F} , that we shall call P . Moreover, if $Q \in \mathcal{F}$ is concave and is linear with slope -1 beyond v^b , $\mathbb{T}Q$ has this property too. Therefore, the fixed point of the operator \mathbb{T} must also have this property, that is, the value function P is concave and has the property that $P'(v) = -1$ for all $v \geq v^b$.

We first establish a lower bound on $P'(0)$. By lemma A.1 below, the optimal contract associated with $v = 0$ is $q_1 = q_1^*$, $q_i = 0$ for $i > 1$ and $u_i = w_i = 0$ for all $i \in \Theta$, and we have

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

Since 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 $q_1 = q_1^*$, $q_i = x$, for $i > 1$, $w_i = 0$ for all $i \in \Theta$, and $u_i = (\theta_n - \theta_i)x$ for all $i \in \Theta$. From the second period on, the contract reverts to $v = 0$. Define x by

$$\begin{aligned}\varepsilon &= \sum_{i \in \Theta} f_i u_i \\ &= (\theta_n - \mathbf{E}[\theta]) x.\end{aligned}$$

to satisfy (PK). Note that this contract satisfies all constraints.

The principal's payoff under the proposed contract is

$$\begin{aligned}Q(\varepsilon) &= f_1 (R(q_1^*) - \theta_1 q_1^*) + \sum_{j=2}^n f_j [R(x) - \theta_j x] - \varepsilon + \delta P(0) \\ &= P(0) + \sum_{j=2}^n f_j [R(x) - \theta_j x] - \varepsilon.\end{aligned}$$

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

$$Q(\varepsilon) - P(0) = \frac{\sum_{j=2}^n f_j [R(x) - \theta_j x] - \varepsilon}{\varepsilon}$$

so that

$$\begin{aligned}\lim_{\varepsilon \rightarrow 0} \frac{Q(\varepsilon) - P(0)}{\varepsilon} &= \lim_{\varepsilon \rightarrow 0} \frac{\sum_{j=2}^n f_j [R(x) - \theta_j x] - \varepsilon}{\varepsilon} \\ &= \sum_{j=2}^n f_j \lim_{x \rightarrow 0} \left[\frac{R(x) - \theta_j x}{(\theta_n - \mathbf{E}[\theta])x} \right] - 1 \\ &= \frac{(1 - f_1)R'(0) + f_1 \theta_1 - \mathbf{E}[\theta]}{\theta_n - \mathbf{E}[\theta]} - 1 \\ &= \frac{(1 - f_1)R'(0) + f_1 \theta_1 - \theta_n}{\theta_n - \mathbf{E}[\theta]},\end{aligned}$$

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

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

as required. Notice now that if (MR_0) holds, that is if $R'(0) = \infty$, it follows immediately that $P'(0) = \infty$.

Since 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 Theorem 4.11 on p 85 of Stokey, Lucas and Prescott (1989). Since P is concave and $P'(v) = -1$ for all $v \geq v^b$, 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, $v^* \leq v^b$. (Of course, it is shown in section 5 that in fact $v^* = v^b$.) \square

While 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 A.1. If $v = 0$, any optimal contract entails $u_i = w_i = 0$ for all i , $q_1 = q_1^*$, and $q_i = 0$ for all $i > 1$.

Proof. To see this, recall that feasibility implies $u_i, w_i \geq 0$ for all i . Promise keeping (PK) requires $\sum_i f_i [u_i + \delta w_i] = 0$, which implies $u_i = w_i = 0$ for all i . This observation and $C_{i,i+1}$ in turn imply that $q_i = 0$ for all $i > 1$. The intuition is simply that if there is any output, there must be some rent paid which, due to the liquidity and feasibility constraints, would violate (PK). \square

Proof of proposition 4.1. Notice that the value function can be written as

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

subject to all the constraints. So suppose $w_i(v) < v^*$ for some v , and by way contradiction, $u_i(v) > 0$. Notice that in all the constraints (C_{ij}) and (PK), u_i and w_i appear in the form $u_i + \delta w_i$. Since $u_i > 0$, we can reduce it by an appropriately chosen $\varepsilon > 0$ and increase w_i by ε/δ . This leaves all the (PK) and (C_{ij}) constraints unchanged. Moreover, liquidity constraint (L_i) is also unaffected. Lastly, the q_i 's are left unchanged. Therefore, this new contract is feasible, and is also a strict improvement, since $w_i + P(w_i)$ 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 the contract is a maximal rent contract, where $w_i(v) \leq v^*$ for all v . We have proved above that $u_i(v) \neq 0$ implies $w_i(v) \geq v^*$. But since u_i

is nonnegative, and the contract is maximal rent, this is equivalent to saying that $u_i(v) > 0$ implies $w_i(v) = v^*$, as required. \square

The following lemma breaks down the proof of Lemma 4.3 into easily digestible parts.

- Lemma A.2.** (a) For all v , q is monotone in type, that is $q_i(v) \geq q_{i+1}(v)$.
- (b) The constraints $C_{i,i+1}$ and $C_{i,i-1}$ imply all other constraints C_{ij} .
- (c) If the constraints $C_{i,i+1}$ hold with equality and $q_i(v) \geq q_{i+1}(v)$ for all $i < n$, the constraints $C_{i,i-1}$ holds for all $i > 1$.
- (d) We may assume that the constraints $C_{i,i+1}$ hold with equality; that is, if any of the constraints $C_{i,i+1}$ are slack, there is another contract that gives the principal the same utility, but where $C_{i,i+1}$ holds with equality.
- (e) In any maximal rent contract, $(C_{i,i+1})$ and (L_i) imply that u and w are monotone in type.

Proof. (a) That $q_i \geq q_{i+1}$ follows by adding $C_{i,i+1}$ and $C_{i+1,i}$.

- (b) Suppose all the constraints $C_{i,i+1}$ and $C_{i,i-1}$ hold. Define $z_i = u_i + \delta w_i$, fix some i , and suppose $j > i$. Then,

$$\begin{aligned} z_i &\geq z_{i+1} + (\theta_{i+1} - \theta_i)q_{i+1} \\ &\geq z_{i+2} + (\theta_{i+2} - \theta_i)q_{i+2} \\ &\geq z_j + (\theta_j - \theta_i)q_j \end{aligned}$$

where we have used the facts that $z_i \geq z_{i+1}$ and $q_i \geq q_{i+1}$. The proof that any constraint C_{ij} also holds for $j < i$ is similar, and therefore omitted.

- (c) That monotonicity (in type) of q_i and the equality of $C_{i,i+1}$ implies $C_{i,i-1}$ is standard, and therefore omitted.
- (d) We want to show that $C_{i,i+1}$ holds with equality for all $i < n$. By the results above, we may restrict attention to upward incentive constraints and assume that q is monotone in type. Suppose that some constraint $C_{i,i+1}$ is slack, so that $u_i + \delta w_i > u_{i+1} + \delta w_{i+1} + \Delta_i q_{i+1}$. There are two cases to consider. The first case is when $u_i > u_{i+1}$. We can increase u_{i+1} by ε and reduce u_i by $(f_i/f_{i+1})\varepsilon$, so that (PK) still holds, none of the upward incentive constraints are upset, and the objective is unchanged. We may choose ε so that $C_{i,i+1}$ holds with equality, which proves this case.

The second case is where $u_i \leq u_{i+1}$, which implies $w_i > w_{i+1}$. Replace w_i with $w'_i := w_i - \varepsilon$, and replace w_{i+1} with $w'_{i+1} := w_{i+1} + (f_i/f_{i+1})\varepsilon$, where $\varepsilon > 0$ is chosen so that $C_{i,i+1}$ holds with equality. Notice that since $q_{i+1} \geq 0$, it must be that $w'_i \geq w'_{i+1}$. We want to show this change does not leave the principal any worse off.

To see this, notice that by construction, $f_i w_i + f_{i+1} w_{i+1} = f_i w'_i + f_{i+1} w'_{i+1}$. Therefore, it only remains to show that $f_i P(w'_i) + f_{i+1} P(w'_{i+1}) \geq f_i P(w_i) + f_{i+1} P(w_{i+1})$, which holds if, and only if, $f_{i+1} [P(w'_{i+1}) - P(w_{i+1})] \geq f_i [P(w_i) - P(w'_i)]$. Recall that P is continuously differentiable, so that if $w'_{i+1} \leq w'_i$, the concavity of P implies $P'(w'_{i+1}) \geq P'(w'_i)$. We then observe

$$\begin{aligned} f_{i+1} [P(w'_{i+1}) - P(w_{i+1})] &\geq f_{i+1} P'(w'_{i+1})(w'_{i+1} - w_{i+1}) \\ &= f_i P'(w'_{i+1})\varepsilon \\ &\geq f_i P'(w'_i)(w_i - w'_i) \\ &\geq f_i [P(w_i) - P(w'_i)] \end{aligned}$$

where we have used the fact that $f_i(w_i - w'_i) = f_i \varepsilon = f_{i+1}(w'_{i+1} - w_{i+1})$, and the first and last inequality follow from the definition of the subdifferential, and the second follows from the concavity of P . This proves our claim.

- (e) We shall show that u and w are monotone in type in any maximal rent contract. Suppose first that $u_i < u_{i+1}$ for some i . Then, by the liquidity constraint (L'_i), it must be that $u_{i+1} > 0$. But by proposition 4.1, this implies $w_{i+1} = v^*$. Now, ($C_{i,i+1}$) with implies $\delta w_i \geq (u_{i+1} - u_i) + \delta w_i + \Delta_i q_{i+1} > \delta w_i$, which implies $w_{i+1} > v^*$, which is impossible in a maximal rent contract. Therefore, it must be that u is monotone in type.

Next, let us assume that $w_{i+1} > w_i$ for some i . Once again, ($C_{i,i+1}$) implies $u_i - u_{i+1} \geq \delta(w_{i+1} - w_i) + \delta_i q_{i+1} > 0$, which implies, by (L_{i+1}), that $u_i > 0$. But proposition 4.1 says we must have $w_i = v^*$, which in turn implies $w_{i+1} > v^*$, which is impossible in a maximal rent contract. Therefore, w must also be monotone in type. \square

Proof of Lemma 4.4. Note that (C_i) can be rewritten, for each i from 1 to $n-1$, as

$$u_i + \delta w_i = u_n + \delta w_n + \sum_{j=i}^{n-1} \Delta_j q_{j+1}$$

Taking the expectation of both sides and – as usual – reversing the order of the

double summation on the right yields

$$\sum_{j=1}^n f_j(u_j + \delta w_j) = u_n + \delta w_n + \sum_{j=1}^{n-1} F_j \Delta_j q_{j+1}$$

Notice from (PK), that the left side of this expression is simply v . Hence, we have

$$u_n + \delta w_n = v - \sum_{j=1}^{n-1} F_j \Delta_j q_{j+1}$$

Substituting this back into the original equation above yields (U_i). □

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 maximisers (q, w) . To see this, suppose (q, w) and (q', w') are optimal at some v , but $q \neq q'$. Then, since the feasible set is convex, $(\frac{1}{2}(q + q'), \frac{1}{2}(w + w'))$ is also feasible, and moreover, is a strict improvement over (q, w) and (q', w') , since $R(q)$ is strictly concave. Therefore, it must be that $q = q'$ across all optimal contracts.

By proposition 4.1 and lemma 4.3, we know that for each v , there exists a type i such that for $u_i = 0$ implies $u_j = 0$ for all $j > i$. 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 $w_i = v^*$, since we have a maximal rent contract. In either case, $w_i(v)$ is uniquely determined in a maximal rent contract.

Finally, Theorem 4.6 of Stokey, Lucas and Prescott (1989) shows that the optimal maximal rent contract must be continuous in v . □

B. Proofs from Section 5

First we present the derivation of v^* given in (Vest).

Proof of Lemma 5.1. Since $P'(v^*) = -1$, we have $\Lambda_n(v^*) = 0$, and since $\lambda_i \geq 0$ for all i , it must be that $\lambda_i(v^*) = 0$ for all i .

By the definitions of P and v^* , we also have $\Lambda_n(v) > 0$ for all $v < v^*$. Lemma 4.3, which says that rents are monotone in type, now implies $\lambda_n(v) > 0$ for all $v < v^*$. But by complementary slackness, $u_n(v) = 0$ for all $v < v^*$. Since the optimal contract is continuous in v (see Theorem 2), it follows that $u_n(v^*) = 0$.

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, (Vest) holds by (U_{*i*}) for $i = n$. \square

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 B.1. 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 maximisers, 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 $\langle a_i, q \rangle + \delta w_i \leq v$, where $a_i \in \mathbb{R}^{n-1}$. 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,

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

where the strict inequality follows from the strict concavity of R . This proves the strict concavity of P , as required. \square

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 B.2. The Lagrange multipliers satisfy the following inequalities:

- (a) $\frac{\lambda_i}{f_i} \geq \frac{\lambda_j}{f_j}$ for $i \geq j$
- (b) $\frac{\lambda_i}{f_i} \geq \frac{\Lambda_{i-1}}{F_{i-1}}$ for $i > 1$
- (c) $\frac{\Lambda_{i+1}}{F_{i+1}} \geq \frac{\Lambda_i}{F_i}$ for $i < n$
- (d) $\Lambda_n \geq \frac{\Lambda_i}{F_i}$ for all $i \leq n$

- Proof.* (a) By Lemma 4.3, we know that in a maximal rent contract, for each v , $w_i(v) \geq w_{i+1}(v)$. The concavity of P then implies that $P'(w_i) \leq P'(w_{i+1})$. By the first order condition for w_i , namely (FO w_i), we see that $-1 + \lambda_i/f_i = P'(w_i) \geq P'(w_{i-1}) = -1 + \lambda_{i-1}/f_{i-1}$. This allows us to conclude that $\lambda_i/f_i \geq \lambda_j/f_j$ for all $i \geq j$.
- (b) The previous part tells that for all $i \geq j$, $\lambda_i f_j \geq \lambda_j f_i$. Summing over $j \leq i-1$, we see that $\lambda_i \sum_{j=1}^{i-1} f_j \geq f_i \sum_{j=1}^{i-1} \lambda_j$, which can be rewritten as $\lambda_i F_{i-1} \geq f_i \Lambda_{i-1}$, ie $\lambda_i/f_i \geq \Lambda_{i-1}/F_{i-1}$, as required.
- (c) From the previous part, we know that $F_i \lambda_{i+1} \geq f_{i+1} \Lambda_i$. Adding $F_i \Lambda_i$ to both sides of the inequality, we get $F_i(\Lambda_i + \lambda_{i+1}) \geq (F_i + f_{i+1})\Lambda_i$, which can be rewritten as $F_i \Lambda_{i+1} \geq F_{i+1} \Lambda_i$, as required.
- (d) The previous part tells us that $\Lambda_n \geq \Lambda_{n-1}/F_{n-1} \geq \dots \geq \Lambda_1/F_1$, as required. \square

The following is an easy corollary of the proposition above.

Corollary B.3. If $\lambda_i/f_i = \lambda_j/f_j$ for all i, j , $\Lambda_n = \Lambda_k/F_k$ for all $k < n$.

Proof. Again, $\frac{\lambda_i}{f_i} = \frac{\lambda_j}{f_j}$ can be rewritten as $f_j \left(\frac{\lambda_i}{f_i} \right) = \lambda_j$. Summing over $j = 1, \dots, k$, we get $F_j \left(\frac{\lambda_i}{f_i} \right) = \Lambda_k$. This gives us the equalities $\frac{\lambda_i}{f_i} = \frac{\Lambda_i}{F_i} = \Lambda_n$, as required. \square

An obvious question is whether the optimal contract can ever have greater than optimal production, which was stated as lemma 5.2 in the main text. We are now in a position to establish this lemma.

Proof of lemma 5.2. Recall the first order condition for q_i , (FO q_i), is

$$R'(q_i) - \theta_i = \frac{\Delta_{i-1}}{f_i} [F_{i-1}\Lambda_n - \Lambda_{i-1}] + \frac{1}{f_i} (\mu_{i-1} - \mu_i)$$

where μ_i is the Lagrange multiplier of the constraint $q_i \geq q_{i+1}$, and q_{n+1} is taken to be 0. Also recall that $R(q_i) - \theta_i q_i$ is concave in q_i , achieving its maximum at q_i^* . Therefore, for any $q_i > q_i^*$, $R'(q_i) - \theta_i < 0$. We shall prove the proposition via the following claims.

Claim 1. If, for some $v \geq 0$, there is an i such that the optimal $q_i > q_i^*$, then $\mu_i > 0$.

Proof of Claim 1. This follows from inspection of the first order condition, which requires that $\frac{\Delta_{i-1}}{f_i} [F_{i-1}\Lambda_n - \Lambda_{i-1}] + \frac{1}{f_i} (\mu_{i-1} - \mu_i) < 0$. But proposition B.2(d) tells us that $F_{i-1}\Lambda_n - \Lambda_{i-1} \geq 0$. Moreover, by virtue of being Lagrange multipliers, $\mu_{i-1}, \mu_i \geq 0$. Therefore, it must be that $\mu_i > 0$. \blacktriangle

Claim 2. If, for some $v \geq 0$, there is an i such that the optimal $q_i > q_i^*$, then $q_{i+1} = q_i$.

Proof of Claim 2. In claim 1 above, we established that $q_i > q_i^*$ implies $\mu_i > 0$. But the KKT complementary slackness condition requires that $\mu_i(q_i - q_{i+1}) = 0$, which implies $q_i = q_{i+1}$. \blacktriangle

Returning to the proof at hand, suppose for some $v \geq 0$, there is an i such that $q_i > q_i^*$. Then, by claims 1 and 2, it must be that $q_i = q_{i+1} > q_{i+1}^*$. This, in turn, implies that $\mu_{i+1} > 0$. Therefore, by induction, we see that $q_n > q_n^* > 0$ (where the second inequality is by assumption) and $\mu_n > 0$. But this is impossible, since the KKT complementary slackness condition requires that $\mu_n(q_n - 0) = 0$, which proves the claim. \square

We can now prove the main Theorem of section 5, which tells us that a constrained version of the principal's problem is well behaved, in the sense that the

maximisers are increasing in v . Recall that the principal's problem can be written as the solution to the Bellman equation

$$P(v) = \max_{(q,w)} \sum_i f_i \left[\left(R(q_i) - \theta q_i \right) + \delta \left(w_i + P(w_i) \right) \right] - v$$

subject to monotonicity in output (M_i), liquidity (L_i), and feasibility $q_i \geq 0, w_n \geq 0$. Let the set of all (q, w) that satisfy these constraints be written as $\Gamma(v)$. Theorem 3 says that (q, w) is monotone increasing in v . We can now provide the proof, while deferring mathematical generalities and details to section D.

Proof. By lemma 5.2, $q_i(v) \leq q_i^*$ for all i and v . Therefore, on the domain $[0, q_i^*]$, $R(q_i) - \theta_i$ is strictly increasing. Consider now the domain

$$X = \left\{ (q, w) : q \in \prod_{i=1}^n [0, q_i^*], w \in [0, v^*]^n, q_i \geq q_{i+1} \text{ for } i > 1 \right\}$$

Clearly, X is a sublattice of \mathbb{R}^{2n} , and is also convex. Moreover, the objective, $\sum_i f_i \left[\left(R(q_i) - \theta q_i \right) + \delta \left(w_i + P(w_i) \right) \right] - v$ is strictly increasing, concave and super-modular on the domain X .

In the next section, we shall see that there is a continuous increasing, convex, and submodular function $g : X \rightarrow \mathbb{R}$, that has the properties that (i) for each v , $\Gamma(v)$ is a subset of the lower contour set of g , namely $\{(q, w) : g(q, w) \leq v\}$, and (ii) solving the optimisation problem on this larger domain $\{(q, w) : g(q, w) \leq v\}$ instead of $\Gamma(v)$ leaves the set of maximisers unchanged.

But this means that we have satisfied all the conditions of corollary 2(ii) of Quah (2007). By Theorem 2, we know that there is a unique maximal rent contract, and the maximisers of our program constitute a maximal rent contract. Therefore, by corollary 2(ii) of Quah (2007), the unique maximal rent contract is monotone increasing in v . \square

We end with a proof of proposition 5.3.

Proof of proposition 5.3. To prove part (a), recall that by Lemma 4.3, $w_1 \leq \dots \leq w_n$ which implies $P'(w_n) \geq \dots \geq P'(w_1)$. The claim now follows from the first order condition for w_i ($\text{FO}w_i$). Part (b) is an immediate consequence of the fact that w_i is monotone increasing in v (Theorem 3), and the first order condition ($\text{FO}w_i$). That $0 < x_1^* \leq \dots \leq x_n^* = v^*$ in (c) follows from the fact that for each $i < n$, $\lambda_i(v)$ is decreasing in v and is nonnegative. All that remains is to show that $x_1^* < v^*$.

It suffices to show that there exists an $\varepsilon > 0$ such that $\lambda_1(v) = 0$ for all $v \in (v^* - \varepsilon, \infty)$. To see that this is true, suppose not, that is, suppose $\lambda_1(v) > 0$ for all $v < v^*$. Then, (L_1) holds for all $v < v^*$, which implies $\delta w_1 = v + \sum_i (1 - F_i) \Delta_i q_{i+1}^*$. But the optimal contract (q, w) is continuous in v , and $q_i(v^*) = q_i^*$. Therefore, for some v sufficiently close to v^* , $\delta w_1 > v^*$, which implies $w_1 > v^*$, which is impossible in a maximal rent contract (which requires that $w_1 \leq v^*$). \square

C. Proofs from Section 6

To prove proposition 6.1, we shall need the following lemma.

Lemma C.1. For any $v \geq 0$, $q_i(v) = q_i^*$ for all i implies $\lambda_i(v)/f_i = \lambda_j(v)/f_j = \Lambda_n(v)$ for all i, j .

Proof. From the first order condition for q_i , (FO q_i), it is clear that $F_{i-1}\Lambda_n = \Lambda_{i-1}$ for all $i > 1$, since by assumption, $\mu_i = 0$ for all i . For $i = 2$, this implies $\Lambda_n = \lambda_1/f_1$.

Suppose now that for all $j \leq i - 1$, we have $\lambda_j/f_j = \lambda_{j-1}/f_{j-1} = \Lambda_n$. In particular, we have $\Lambda_n = \Lambda_{i-1}/F_{i-1} = \lambda_{i-1}/f_{i-1}$, which can be rewritten as

$$[C.1] \quad f_{i-1}\Lambda_{i-1} = F_{i-1}\lambda_{i-1}.$$

The condition $F_{k-1}\Lambda_n = \Lambda_{k-1}$ for $k = i + 1$ then gives us $\Lambda_n = \Lambda_i/F_i = \lambda_{i-1}/f_{i-1}$ (where the last equality is the induction hypothesis). This is equivalent to $f_{i-1}\lambda_i + f_{i-1}\Lambda_{i-1} = \lambda_{i-1}f_i + F_{i-1}\lambda_{i-1}$. By equation (C.1), this is equivalent to $\lambda_i/f_i = \lambda_{i-1}/f_{i-1}$. This establishes $\Lambda_n = \lambda_j/f_j = \lambda_{j-1}/f_{j-1}$ for all $j < n$.

For $j = n$, recall that $\Lambda_n = \Lambda_{n-1}/F_{n-1}$. This can be rewritten as $F_{n-1}\lambda_n + F_{n-1}\Lambda_{n-1} = f_n\lambda_{n-1} + F_{n-1}\Lambda_{n-1}$, that is $\lambda_n/f_n = \Lambda_{n-1}/F_{n-1} = \Lambda_n$, as required. \square

We are now ready to prove proposition 6.1, which says that in the unique, optimal maximal rent contract, for all $v \in (0, v^*)$, $P'(w_n) > P'(v) > P'(w_1)$. This is Lemma 5, parts (i) and (ii), of Thomas and Worrall (1990). While their proof is applicable here, we give another proof of the result, that uses the monotonicity of the optimal contract.

Proof of proposition 6.1. Recall that w is monotone in type, that is $w_1 \geq \dots \geq w_n$, which implies $P'(w_n) \geq \dots \geq P'(w_1)$. The claim is that for all $v \in (0, v^*)$, $P'(w_n) > P'(v) > P'(w_1)$. So suppose the claim is not true. Since P' is a martingale, the only possibility then is that $P'(w_1) = \dots = P'(w_n) = P'(v)$. (Notice that this does not imply $w_1 = \dots = w_n = v$, since we haven't established that P is strictly concave.)

Since $P'(0) = \infty$, we know that $w_n > 0$. The first order condition (FO w_i) then implies $\lambda_i/f_i = \lambda_j/f_j$ for all i, j . Corollary B.3 then implies $F_{i-1}\Lambda_n = \Lambda_{i-1}$, and (FO q_i) then implies $q_i = q_i^*$ for all i .

Let $v^{(0)} := v$, and define $K = \sum_{i=1}^{n-1} (1 - F_i)\Delta_i q_{i+1}^*$. Suppose $v^{(k)}$ is such that $q_i(v^{(k)}) = q_i^*$. Lemma C.1 then implies $\lambda_i/f_i = \lambda_j/f_j$ for all i, j which, in turn, implies $P'(w_1) = \dots = P'(w_n) = P'(v^{(k)})$. Define now, $v^{k+1} := w_1(v^{(j)}) = (v^{(j)} + K)/\delta$, and notice that since the optimal $q_i(v)$ is (a) by lemma 5.2, less than $q_i^*(v)$ for all i and v , and (b) by Theorem 3, monotone in v , so it must be that $q_i(v^{k+1}) = q_i^*$ for all i .

In sum, we have a sequence $(v^{(k)})_{k \geq 0}$ with the following properties. (i) $v^{(k+1)} - v^{(k)} > K/\delta$, and (ii) $P'(v^{(k+1)}) = P'(v^{(k)})$ for all k . But by (i), there is some k such that $v^{(k)} \geq v^*$, and by assumption, $P'(v^{(0)}) > -1$. This gives us $P'(v^{(k)}) = P'(v^{(0)}) > -1 = P'(v^*) = P'(v^{(k)})$, which is a contradiction. (In the above, the first equality is (ii) above, the strict inequality is since $v^{(0)} < v^*$, $P'(v^*) = -1$ by definition, and the last equality is by definition of P – see Theorem 1.) \square

We now prove another useful lemma that shows that with positive probability, the martingale P' can take all values in $(-1, \infty)$. This is lemma 5(iii) of Thomas and Worrall (1990) and we follow their proof.

Lemma C.2. For any $v \in (0, v^*)$, and $\gamma > P'(v)$, if state θ_n 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.

Proof. Suppose state θ_n occurs repeatedly. This gives us a sequence $v^0 = v$, $v^1 = w_n(v^0) < v^0$, and $v^\tau = w_n(v^{\tau-1}) < v^{\tau-1}$. Since (v^τ) is a 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, since $P'(0) = \infty$.

Therefore, suppose the claim is not true. This implies there is some $y_n > 0$ such that $\lim_{\tau \rightarrow \infty} v^\tau = y_n$. In other words, $\lim_{\tau \rightarrow \infty} w_n(v^\tau) = y_n$. Since the optimal contract is continuous in v , $w_n(\cdot)$ is continuous in v . Therefore, $w_n(y_n) = y_n$, which

contradicts proposition 6.1 which requires that $w_n(y_n) < y_n$. This gives us the desired contradiction. The second part is similar and therefore omitted. \square

We now move to the proof of proposition 4. Once again, we follow Thomas and Worrall (1990).

Proof of proposition 4. Since P' is a martingale that is bounded below by -1 , it follows that $P' + 1$ is a nonnegative martingale. The (positive) Martingale Convergence Theorem (see, for instance, Theorem 22 of Pollard, 2002), 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.

Consider a sample path with the properties that (i) $\lim_{t \rightarrow \infty} P'(v^t) = C \notin \{-1, \infty\}$, and (ii) state θ_n occurs infinitely often, and define $C =: P'(y)$, so that $\lim_{t \rightarrow \infty} v^t = y$. Consider a subsequence $(\sigma(t))$ such that $\theta^{\sigma(t)} = \theta_n$ for all t , ie this is the subsequence consisting of all the θ_n shocks in the original sequence. Since $(v^{\sigma(t)})$ 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 $\varphi^\sigma(v, \theta_n)$ where $\varphi^\sigma(v^{\sigma(t)}, \theta_n) = v^{\sigma(t+1)}$. Since $\varphi(v, \theta_i)$ is continuous in v , it follows that $\varphi^\sigma(v, \theta_n)$ is also continuous in v . Therefore, the sequence $\varphi^\sigma(v^{\sigma(t)}, \theta_n)$ converges to $\varphi^\sigma(y, \theta_n)$. Moreover, $\varphi^\sigma(y, \theta_n) = \varphi(y, \theta_n) = y$, since $\varphi^\sigma(v^{\sigma(t)}, \theta_n) = v^{\sigma(t+1)}$, and $\lim_{t \rightarrow \infty} v^{\sigma(t)} = \lim_{t \rightarrow \infty} v^t = y$.

But $\lim_{t \rightarrow \infty} P'(v^{\sigma(t)}) = C$ and $\lim_{t \rightarrow \infty} P'(v^{\sigma(t+1)}) = C$, so by the continuity of P' we have $P'(y) = P'(\varphi^\sigma(y, \theta_n)) = P'(\varphi(y, \theta_n)) = C$, contradicting proposition 6.1 which states that $P'(y) < P'(\varphi(y, \theta_n))$. But paths where state θ_n does not occur infinitely often are of probability zero, which proves the proposition. \square

Proof of proposition 7.2. By way of contradiction, suppose it is never optimal to fire an agent, then for all $v \in [0, v^0]$ the principal's payoff is $P(v)$. If she fires the current agent when his sweat equity is v (and never fires another), then her payoff is $P(v^0) - v$. Hence, a contradiction will obtain if there exists $v \in (0, v^0)$ such that $P(v^0) - v > P(v)$ or $P(v) + v < P(v^0)$. Note that $P(0) < P(v^0)$ by definition of v^0 . Since P is continuous, there exists $v > 0$ such that $P(v) + v < P(v^0)$. \square

D. Monotonicity of Maximisers

In this section, we provide some intuition behind Quah's result. We then show how the contracting problem (at any equity level v) may be viewed as an optimisation problem, and describe fully transformed problems that are equivalent in the sense that the contracting and transformed problems have identical solutions. Showing that the transformed problem has maximisers monotone in v allows us to establish the monotonicity of the solutions to the contracting problem.

D.1. Comparative Statics for Constrained Optimisation

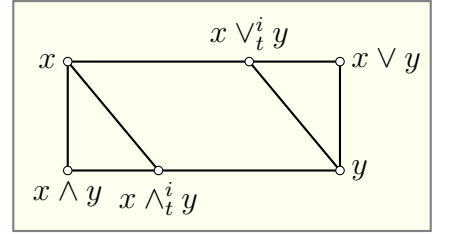
We now provide a simplified version of and some of the intuition behind Quah's results. It is helpful to begin with some classical results from the theory of monotone comparative statics. Let (X, \succeq) be a lattice, with \vee and \wedge respectively denoting the least upper and greatest lower bounds. The **strong set order**, \succeq^p on 2^X induced by \succeq , is defined as follows: $S \succeq^p T$ if, and only if, for any $x \in S$ and $y \in T$, $x \vee y \in S$ and $x \wedge y \in T$. We can now state a result which tells us that larger constraint sets imply the set of maximisers increases. For two subsets S and T of X , and for a supermodular function $U : X \rightarrow \mathbb{R}$, $\arg \max_{x \in S} U(x) \succeq^p \arg \max_{x \in T} U(x)$. Notice now that in a setting where, say, $X = \mathbb{R}_+^n$ and \succeq is the standard Euclidean order, and S and T are budget sets at different wealth levels, S and T cannot be ordered in the strong set order.

But suppose now that there is a family of orders \mathcal{C} on X such that X is a lattice with respect to each order in the family \mathcal{C} , and the function U is supermodular with respect to the order \succeq for each order $\succeq \in \mathcal{C}$. Consider now, the two budget sets S and T . If the family of orders \mathcal{C} is rich enough that $S \succeq^p T$ for some $\succeq \in \mathcal{C}$, then we would still be able to say something about the monotonicity of the set of maximisers. This is precisely the observation of Quah (2007): If X is a convex sublattice of \mathbb{R}^n , endowed with the usual order \succeq , and $U : X \rightarrow \mathbb{R}$ is supermodular *and* concave, then, there exists a family of orders \mathcal{C} on X such that U is supermodular with respect to each order in \mathcal{C} , and \mathcal{C} is rich enough that a large class of constraint sets, for instance budget sets, can be ordered in the strong set order induced by *some* order in \mathcal{C} . Again, the key idea is that given two constraint sets (that are well behaved, for instance two budget sets at different wealth levels), there is an order $\succeq \in \mathcal{C}$ such that the budget sets can be ordered in

the strong set order induced by \succsim . Since U is supermodular with respect to all the orders in \mathcal{C} , the monotonicity result follows.

We now present the construction of the family of orders \cdot . For each $i = 1, \dots, n$, there is a one-parameter family of partial orders $(\succsim_t^i)_{t \in [0,1]}$ on X such that U is supermodular with respect to \succsim_t^i for all $t \in [0, 1]$, that is, U is supermodular with respect to each partial order in the family. In particular, the partial order \succsim_t^i induces least upper and greatest lower bounds \vee_t^i and \wedge_t^i that are defined as: $x \vee_t^i y := y$ if $x_i \leq y_i$ and $x \vee_t^i y := tx + (1 - t)(x \vee y)$ if $x_i > y_i$. Similarly, define $x \wedge_t^i y := x$ if $x_i \leq y_i$ and $x \wedge_t^i y := ty + (1 - t)(x \wedge y)$ if $x_i > y_i$.

Formally, let $\mathcal{C}_i := \{\succsim_t^i: t \in [0, 1]\}$ and say that U is \mathcal{C}_i -supermodular if it is supermodular with respect to the order \succsim_t^i for each $\succsim_t^i \in \mathcal{C}_i$. Then, U is \mathcal{C} -supermodular if it is \mathcal{C}_i -supermodular for all $i = 1, \dots, n$. To gain some intuition for this definition, notice first that if $t = 0$, \mathcal{C} -supermodularity reduces to the standard notion of supermodularity. But for U to be \mathcal{C}_i -supermodular for some i , we require that $U(x) - U(x \wedge_t^i y) \leq U(x \vee_t^i y) - U(y)$, that is, we require supermodularity with respect to the backward bending parallelogram in the figure.



To see the advantage of the family of orders $(\succsim_t^i)_{t \in [0,1]}$, $i = 1, \dots, n$, consider two budget sets $S := \{x \in \mathbb{R}_+^n : \langle q, x \rangle \leq w\}$ and $T := \{x \in \mathbb{R}_+^n : \langle q, x \rangle \leq w'\}$, where $w > w'$ are wealth levels and q is a price vector. Then, it can be shown that there exists a $t^* \in [0, 1]$ such that $S \succsim_{t^*}^{i,p} T$, that is, S dominates T in the strong set order induced by the order $\succsim_{t^*}^i$ for each i . But since U is supermodular with respect to the order $\succsim_{t^*}^i$, we now see that $\arg \max_{x \in S} U(x) \succsim_{t^*}^{i,p} \arg \max_{x \in T} U(x)$.

In sum, by making stronger requirements of the function U , we introduce a family of orders \mathcal{C} on X such that U is \mathcal{C} -supermodular, that is supermodular with respect to each order in \mathcal{C} . For any two budget sets that differ in wealth, there exists an order such that one dominates the other in the induced strong set order. This allows us to show that the set of maximisers is increasing in the induced strong set order. Needless to say, this is but a sampling (and simplistic account) of the results in Quah (2007). He generalises these ideas to allow for ordinal notions (recall that our function U is concave and supermodular, both cardinal notions), and more general constraint sets (our discussion is restricted to budget sets). Notice in particular that budget sets are lower contour sets of functions of the form $\langle q, x \rangle$, which are increasing, convex and submodular functions

of x , and in light of the observation that supermodularity with respect to the standard order on \mathbb{R}^n and concavity imply supermodularity with respect to an entire family of orders, this suggests we can consider constraint sets of the kind described above, namely constraint sets that are the lower contour sets of increasing, convex and submodular functions.

D.2. The Problem at Hand

The first thing to notice is that the optimisation problems we are interested in take the following form:

$$[\text{G}] \quad \max_{x \in X} U(x) \text{ s.t. } Ax \leq v\mathbf{1}$$

where $\mathbf{1} = (1, \dots, 1)$, $X \subset \mathbb{R}_+^n$ is a convex sublattice, and A is an $m \times n$ matrix. Typically, especially in the applications we have in mind, the matrix A will have negative entries. We will make the following standing assumptions about the objective: U is concave, supermodular and strictly increasing.

Suppose instead, that we are solving the problem

$$[\text{N}] \quad \max_{x \in \mathbb{R}_+^n} U(x) \text{ s.t. } \hat{A}x \leq v\mathbf{1}$$

where $\mathbf{1} = (1, \dots, 1)$, and \hat{A} is an $m \times n$ matrix that is nonnegative, that is each entry $\hat{a}_{ij} \geq 0$, and U is concave and supermodular. For each $i = 1, \dots, m$, define the linear functional f_i as $f_i(x) = \langle \hat{a}_i, x \rangle = \sum_{j=1}^n \hat{a}_{ij}x_j$. Define also the function $g : \mathbb{R}_+^n \rightarrow \mathbb{R}$ as

$$g(x) = \max\{f_1(x), \dots, f_m(x)\}$$

Proposition D.1. The function $h(y) = \max\{y_1, \dots, y_n\}$ is submodular.

Proof. Example 2.6.2, Topkis (1998, p 46), shows that $x \mapsto \min\{\alpha_i x_i : i = 1, \dots, n\}$ where $\alpha_i \leq 0$ for $i = 1, \dots, n$, is supermodular on \mathbb{R}^n . Now, $h(x) = \max\{x_1, \dots, x_n\} = -\min\{-x_1, \dots, -x_n\}$. But $x \mapsto \min\{-x_1, \dots, -x_n\}$ is supermodular, therefore h is submodular. \square

Our second proposition says that the constraint set is the lower contour set of an increasing, convex and submodular function.

Proposition D.2. Let f_i be a positive linear functional, representing a row of the matrix \hat{A} as above, so that $f_i \geq 0$, for each $i = 1, \dots, m$. Then, the function

$$g(x) = \max\{f_1(x), \dots, f_m(x)\}$$

is increasing, convex and submodular in x .

Proof. The proof is a simple consequence of a lemma in Topkis (1998). Let X be a lattice, $f_i(x)$ increasing and supermodular, $Z_i \subset \mathbb{R}$, the range of f_i that is convex for $i = 1, \dots, n$, and $h(z_1, \dots, z_n)$ supermodular in (z_1, \dots, z_n) on $\times_{i=1}^n Z_i$, and is decreasing and convex in z_i for $i = 1, \dots, n$, and for all $z_{-i} \in Z_{-i}$. Then, lemma 2.6.4 from Topkis (1998, p 56) says that $h(f_1(x), \dots, f_n(x))$ is supermodular on X .

Suppose now that f_i is a positive linear functional for each $i = 1, \dots, n$, so that the domain of f_i is $Z_i = \mathbb{R}$. Then, $g(x) = \max\{f_1(x), \dots, f_m(x)\}$ is clearly increasing and convex. All that remains to be shown is that g is submodular. Notice that $g(x) = -\min\{-f_1(x), \dots, -f_m(x)\}$. Let $h : \mathbb{R}^n \rightarrow \mathbb{R}$ be defined as $h(z) = \min\{-z_1, \dots, -z_n\}$, so that it satisfies the hypotheses of the lemma stated above. Then, $h(f_1(x), \dots, f_n(x)) = \min\{-f_1(x), \dots, -f_m(x)\} = -g(x)$ is supermodular. Therefore, g is submodular on \mathbb{R}^n , as required. \square

Notice that $x \in \mathbb{R}_+^n$ satisfies $\hat{A}x \leq v\mathbf{1}$ if, and only if, $g(x) \leq v\mathbf{1}$. Recall also that U is concave and supermodular. Then, our optimisation problem can be rewritten as

$$[\text{N}] \quad \max_{x \in \mathbb{R}_+^n} U(x) \text{ s.t. } g(x) \leq v\mathbf{1}$$

where g is as defined above.

Therefore, by Proposition 4 and Theorem 2 of Quah (2007), the solution set is monotone in v . In particular, if U is strictly concave (which implies the solution is unique), the solution is monotone in v .

In the next section, we shall show that in the problems that concern us, we can replace the general matrix A (which could have negative entries) by a matrix \hat{A} that has only nonnegative entries without affecting the set of solutions of the optimisation problem.

D.3. Some Equivalent Problems

We will now show that problems of type (G) can be rewritten as problems of type (N), and are therefore amenable to analysis. First some terminology. A set $E \subset \mathbb{R}^n$ is *polyhedral* if it can be written as the intersection of a finite collection of closed halfspaces. It is a *polytope* if it is a bounded polyhedron.

As in the problem (G), let A be an $m \times n$ matrix, where $a_{1j} > 0$ for $j = 1, \dots, n$. Then, the set $E_1^v := \{x \in \mathbb{R}_+^n : Ax \leq v\mathbf{1}\}$ is a polytope, and hence also polyhedral. Notice that $E_1^v = vE_1^1$, that is, the set of solutions to the linear inequalities is homogeneous in v .

Corollary 19.3.2 of Rockafellar (1970) says that the sum of two convex polyhedral sets is also convex polyhedral. Therefore, $E_2^v = E_1^v + \mathbb{R}_-^n = E_1^v - \mathbb{R}_+^n$ is also convex and polyhedral. Therefore, E_2^v can be written as the intersection of a finite collection of closed half spaces, ie there is a matrix \tilde{A} such that $E_2^v = \{x \in \mathbb{R}^n : \tilde{A}x \leq b\}$, where $b \in \mathbb{R}_{++}^m$. Our first claim is that \tilde{A} is nonnegative, ie $\tilde{a}_{ij} \geq 0$ for each i, j . In what follows, let $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ be the i th unit vector (when \mathbb{R}^n has the standard basis).

Proposition D.3. The matrix \tilde{A} defined above is nonnegative.

Proof. Notice first that $x \in E_2^v$ implies $x - e_i \in E_2^v$, for all $i = 1, \dots, n$. To see this, recall that $x \in E_2^v$ if, and only if, there exist $x_0 \in E_1^v$ and $y \in \mathbb{R}_+^n$ such that $x = x_0 - y$. But $y + e_i \in \mathbb{R}_+^n$, and by definition of E_2^v , $x_0 - (y + e_i) \in E_2^v$. Therefore, $x - e_i = x_0 + (y - e_i) \in E_2^v$, as claimed.

To prove the proposition, let us now assume that $\tilde{a}_{ij} < 0$. By definition of \tilde{A} , there exists $x \in \mathbb{R}^n$ such that $\langle \tilde{a}_i, x \rangle = v$, ie x lies on the hyperplane defined by \tilde{a}_i that has value v . From the assumption that $\tilde{a}_{ij} < 0$, it follows that $\langle \tilde{a}_i, x - e_i \rangle = v - \tilde{a}_{ij} > v$, which implies $x - e_i \notin E_2^v$, which is a contradiction. Therefore, \tilde{A} is nonnegative, as claimed. \square

Since the function U is only defined on \mathbb{R}_+^n , we introduce the domains $D_i^v := E_i^v \cap \mathbb{R}_+^n$, for $i = 1, 2$. Thus, each D_i^v is a polytope. We now show that enlarging the domain does not change the set of solutions to (G).

Proposition D.4. Let D_1^v, D_2^v and U be defined as above, where U is strictly increasing. Then, $x^* \in \arg \max_{x \in D_1^v} U(x)$ if, and only if $x^* \in \arg \max_{x \in D_2^v} U(x)$.

Proof. Let $x_1^* \in \arg \max_{x \in D_1^v} U(x)$ and suppose first that $x_1^* \notin \arg \max_{x \in D_2^v} U(x)$. Then, there exist $y \in D_1^v$, $z \in \mathbb{R}_+^n$ such that $y - z \in D_2^v$ and $U(y - z) > U(x_1^*)$. But U is strictly increasing, so $U(y) > U(y - z) > U(x_1^*)$, which is a contradiction.

For the converse, suppose $x_2^* \in \arg \max_{x \in D_2^v} U(x)$. Clearly, if $x_2^* \notin D_1^v$, there is $z \in \mathbb{R}_+^n$ such that $x_2^* + z \in D_1^v \subset D_2^v$, and moreover, $U(x_2^*) < U(x_2^* + z)$. Therefore, we must have that $x_2^* \in D_1^v$. But this implies $x_2^* \in \arg \max_{x \in D_1^v} U(x)$, as desired. \square

We have therefore shown that E_2^v can be written as $E_2^v := \{x \in \mathbb{R}^n : \tilde{A}x \leq b\}$, where $b \in \mathbb{R}_{++}^m$. It is easy to see that each $b_i > 0$, because if this were not the case, we would have $\{x : \langle \tilde{a}_i, x \rangle \leq b_i\} \cap \mathbb{R}_+^n = \emptyset$, which would imply E_1^v is not a subset of E_2^v , a contradiction.

Define now the new matrix $\hat{A} = [\hat{a}_{ij}]$ by $\hat{a}_{ij} = \tilde{a}_{ij}/(b_i/v)$. It is easy to see that $E_2^v = \{x \in \mathbb{R}^n : \hat{A}x \leq v\mathbf{1}\}$. Moreover, \hat{A} is nonnegative, as in problem (N).

D.4. Coda

Consider, once again, the problem

$$[\text{G}] \quad \max_{x \in \mathbb{R}_+^n} U(x) \text{ s.t. } Ax \leq v\mathbf{1}$$

where $\mathbf{1} = (1, \dots, 1)$, and A is an $m \times n$ matrix, and U is concave, supermodular and strictly increasing. The propositions above establish that x^* solves problem (G) if, and only if, there is a nonnegative matrix \tilde{A} such that x^* solves problem (N). Therefore, the set of solutions $x^*(v)$ is monotone in v , as desired.

References

- Albanesi, Stefania and Christopher Sleet (2006). ‘Dynamic Optimal Taxation with Private Information’. In: *Review of Economic Studies* 73.1, pp. 1–30 (cit. on p. 5).
- Athey, Susan and Ilya Segal (2007). *An Efficient Dynamic Mechanism*. Tech. rep. Stanford University (cit. on p. 8).
- Battaglini, Marco (2005). ‘Long-Term Contracting with Markovian Consumers’. In: *American Economic Review* 95.3, pp. 637–658 (cit. on pp. 8, 9).
- Bergemann, Dirk and Juuso Välimäki (2010). ‘The Dynamic Pivot Mechanism’. In: *Econometrica* 78.2, pp. 771–789 (cit. on p. 8).
- Berle, Adolf A and Gardiner C Means (1968). *The Modern Corporation and Private Property*. revised. Harcourt, Brace and World (cit. on p. 23).
- Biais, Bruno et al. (2007). ‘Dynamic Security Design: Convergence to Continuous Time and Asset Pricing Implications’. In: *Review of Economic Studies* 74.2, pp. 345–390 (cit. on p. 6).
- Blair, Roger D and Francine Lafontaine (2005). *The Economics of Franchising*. New York, NY: Cambridge University Press (cit. on p. 27).
- Boleslavsky, Raphael (2009). *Dynamic Screening in a Long Term Relationship*. Tech. rep. University of Miami (cit. on p. 8).
- Bolton, Patrick and David S Scharfstein (1990). ‘A Theory of Predation Based on Agency Problems in Financial Contracting’. In: *American Economic Review* 80.1, pp. 93–106 (cit. on p. 6).
- (1998). ‘Corporate Finance, the Theory of the Firm, and Organizations’. In: *Journal of Economic Perspectives* 12.4, pp. 95–114 (cit. on p. 23).
- Clementi, Gian Luca and Hugo A Hopenhayn (2006). ‘A Theory of Financing Constraints and Firm Dynamics’. In: *Quarterly Journal of Economics* 121, pp. 229–265 (cit. on pp. 6, 7).
- Covallo, Ruggiero (2008). ‘Efficiency and Redistribution in Dynamic Mechanism Design’. In: *Proceedings of the 9th ACM Conference on Electronic Commerce*. ACM (cit. on p. 8).
- DeMarzo, Peter and Michael J Fishman (2007). ‘Agency and Optimal Investment Dynamics’. In: *Review of Financial Studies* 20.1, pp. 151–188 (cit. on p. 6).
- DeMarzo, Peter and Yuliy Sannikov (2006). ‘Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Agency Model’. In: *Journal of Finance* 61.6, pp. 2681–2724 (cit. on p. 6).
- Edmans, Alex et al. (2010). *Dynamic Incentive Accounts*. Tech. rep. NYU (cit. on p. 12).

- Fernandes, Ana and Christopher Phelan (2000). ‘A Recursive Formulation for Repeated Agency with History Dependence’. In: *Journal of Economic Theory* 91, pp. 223–247 (cit. on p. 5).
- Green, Edward J (1987). ‘Lending and the Smoothing of Uninsurable Income’. In: *Contractual Agreements for Intertemporal Trade*. Ed. by Edward C Prescott and Neil Wallace. University of Minnesota Press (cit. on p. 5).
- Grossman, Sanford J and Oliver D Hart (1986). ‘The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration’. In: *Journal of Political Economy* 94.4, pp. 691–719 (cit. on p. 23).
- HBS Case Study, 9-800-170 (2000). *Venture Capital Negotiations: VC versus Entrepreneur*. Tech. rep. Harvard Business School (cit. on p. 28).
- Hart, Oliver D and John Moore (1990). ‘Property Rights and the Nature of the Firm’. In: *Journal of Political Economy* 98.6, pp. 1119–1158 (cit. on p. 23).
- Kaplan, Steven N and Per Stromberg (2003). ‘Financial Contracting Theory Meets the Real World: An Empirical Analysis of Venture Capital Contracts’. In: *Review of Economic Studies* 70.2, pp. 281–315 (cit. on p. 28).
- Laffont, Jean-Jacques and David Martimort (2002). *The Theory of Incentives*. Princeton, NJ: Princeton University Press (cit. on p. 7).
- Liddle, Alan (2010). ‘Investing in Sweat Equity’. Nation’s Restaurant News. URL: <http://www.nrn.com/article/investing-sweat-equity> (cit. on p. 27).
- Milgrom, Paul (1987). *Adverse Selection without Hidden Information*. Tech. rep. 8742. University of California, Berkeley (cit. on p. 6).
- Milgrom, Paul and Chris Shannon (1994). ‘Monotone Comparative Statics’. In: *Econometrica* 62, pp. 157–180 (cit. on p. 29).
- Mirrlees, James A (1971). ‘An Exploration in the Theory of Optimum Income Taxation’. In: *Review of Economic Studies* 38.114, pp. 175–208 (cit. on p. 8).
- Pavan, Alessandro, Ilya Segal and Juuso Toikka (2009). *Dynamic Mechanism Design: Incentive Compatibility, Profit Maximization and Information Disclosure*. Tech. rep. Stanford University (cit. on p. 8).
- (2010). *Infinite-Horizon Mechanism Design: The Independent-Shock Approach*. Tech. rep. Stanford University (cit. on p. 8).
- Pollard, David (2002). *A User’s Guide to Measure Theoretic Probability*. New York, NY: Cambridge University Press (cit. on p. 48).
- Quadrini, Vincenzo (2004). ‘Investment and liquidation in renegotiation-proof contracts with moral hazard’. In: *Journal of Monetary Economics* 51.4, pp. 713–751 (cit. on p. 6).

- Quah, John K-H (2007). ‘The Comparative Statics of Constrained Optimization Problems’. In: *Econometrica* 57.2, pp. 401–431 (cit. on pp. 4, 8, 18, 30, 32, 34, 45, 49, 50, 52).
- Rockafellar, R Tyrell (1970). *Convex Analysis*. Princeton, NJ: Princeton University Press (cit. on p. 53).
- Sannikov, Yuliy (2008). ‘A Continuous-Time Version of the Principal–Agent Problem’. In: *Review of Economic Studies* 75.3, pp. 957–984 (cit. on p. 5).
- Spear, Stephen E and Sanjay Srivastava (1987). ‘On Repeated Moral Hazard with Discounting’. In: *Review of Economic Studies* 54.4, pp. 599–617 (cit. on p. 5).
- Steele, J Michael (2001). *Stochastic Calculus and Financial Applications*. New York, NY: Springer (cit. on p. 25).
- Stokey, Nancy L, Robert E Lucas Jr and Edward C Prescott (1989). *Recursive Methods in Economic Dynamics*. Cambridge, Ma: Harvard University Press (cit. on pp. 29, 35, 36, 38, 41).
- Thomas, Jonathan and Tim Worrall (1990). ‘Income Fluctuation and Asymmetric Information: An Example of a Repeated Principal-Agent Problem’. In: *Journal of Economic Theory* 51, pp. 367–390 (cit. on pp. 6, 11, 26, 46–48).
- Topkis, Donald M (1998). *Supermodularity and Complementarity*. Princeton, NJ: Princeton University Press (cit. on pp. 29, 31, 51, 52).
- White, Rebecca J, Rodney R D’Souza and John C McIlwraith (2007). ‘Leadership in venture backed companies: going the distance’. In: *Journal of Leadership and Organizational Studies* (cit. on p. 29).