A Theory of Dynamic Contracting under Financial Constraints

Ilia Krasikov† Rohit Lamba‡

December 2016

Abstract

We study a dynamic principal-agent model where the agent has access to a persistent private technology but is strapped for cash. Financial constraints are generated by the periodic interaction between incentives (private information) and feasibility (being strapped for cash). This interaction produces dynamic distortions that are an additive sum of two effects: backloading of incentives and illiquidity. A sequence of bad technology shocks increase distortions and monotonically push the agent further away from efficiency. An endogenous number of good shocks are required for the agent to become liquid, and eventually for the contract to become efficient. Efficiency is an absorbing state. The optimal allocation can be implemented through a mechanism which is precisely pinned down by a dynamic information operator. The shares of principal and agent in the net present value of economic surplus endogenous to the evolution of technology shocks. Surplus itself is increasing in the share of the agent, and in its type contingent utility spread. By comparing the agent’s utility with and without financial constraints, the model provides a foundation for the usefulness of limited liability in dynamic contracts. Persistence of agent’s private information lends an empirical punch to financial contracting.

0 Introduction

Financial constraints are ubiquitous, and economists have converged to a view that their incorporation in our simplified models is inevitable, even desirable. The emphasis has been pronounced post the Great Recession, so much so that a close cousin—financial frictions—is now colloquially understood to be an important causal element of inefficient allocation of capital at a micro level, and a key driver of crises episodes at the macro.

In an elegant short note, Nobuhiro Kiyotaki advocates "a mechanism design approach to illustrate how different environments of private information and limited commitment generate different financial frictions."1 In the spirit of the said agenda, this paper posits financial

---

*Thanks to the participants at the Pennsylvania Economic Theory Conference, Vanderbilt Mechanism Design Conference, and Northwestern University for their comments.
†Pennsylvania State University, izk113@psu.edu
‡Pennsylvania State University, rlamba@psu.edu.
1Kiyotaki [2011].
constraints as a product of the interaction between (i) "persistent agency frictions", and (ii) "limitations on the ability of agents to generate timed cash flows".

Agency frictions can be modeled as moral hazard or adverse selection. All dynamic financial contracting papers thus far have explored moral hazard (or cash flow diversion models), and almost all have focussed on iid technologies. In contrast, this paper looks at a dynamic adverse selection (or screening) model with persistence in technology. It provides a tractable model of endogenously generated financial constraints, and delivers clear insights on the short-run and long-run outcomes of contracting in such an environment.

Persistence in technology shocks lends an important empirical relevance to financial contracting. It seems like a natural assumption for durable technologies owned and managed by the same firm in a dynamic setting. For example, using Compustat data on firms from 1962 to 2009, Imrohoroglu and Tuzel [2014] find the average persistence in total factor productivity to be 0.7. Most dynamic financial contracting models thus far have restricted themselves to iid models of cash flow diversion, primarily for tractability.\(^2\)

Specifically, we study a dynamic contracting model with persistent asymmetric information and cash or limited liability constraints. A big firm (principal) repeatedly producing a complex finished good contracts with a smaller firm (agent) that supplies an important input. The marginal cost of production for the smaller firm is its private information, and it is imperfectly correlated across time. Moreover, the small firm lacks capital- it cannot post a bond, collateralize its assets or forgo payments. In simple terms it is strapped for cash. Mathematically, this restriction demands positivity of per-period (or stage) utility of the agent.\(^3\)

The big firm is tasked with designing a contract which sets the supply of inputs by the small firm, and regular payments for its production. As motivated, financial constraints are produced by the simultaneous interaction of persistent private information and the cash-strapped constraint. Relaxing either of these leads to greater or complete efficiency in the optimal contract, arguably limiting the empirical appeal of the model.

Ideally, the big firm would like the small firm to produce the efficient quantity in return for payments determined by its bargaining power or outside option. However, agency frictions prevent the efficient allocation from being implemented. The optimal quantity (or allocation) is determined by the trade-off between efficiency and rents:

\[
\max_q \left[ \text{surplus} - \text{information rent} \right] \Rightarrow \frac{d_q(\text{surplus})}{d_q(\text{information rent})} = \frac{\text{marginal benefit}}{\text{marginal cost}} \tag{\star}
\]

Equation (\(\star\))- pointwise optimization of quantities for each type of the agent- forms the backbone of our analysis. The left-hand side reports the derivative of surplus produced by a

\(^2\)Two exceptions known to us are Sannikov [2014], and Fu and Krishna [2016]. They employ moral hazard and cash flow diversion models, respectively, as opposed to our model of dynamic screening.

\(^3\)There is a large literature explaining the inability of start-ups or small businesses to raise requisite financial capital, which forms the core empirical motivation of our model. For example, surveying 1050 CFOs across the US, Europe and Asia, Campello, Graham and Harvey [2010] show the considerable impact of credit constraints on real firm behavior in the aftermath of the Great Recession. Exploiting a change in policy by the Indian government, Banerjee and Duflo [2014] show that most firms in their study were credit constrained, and a relaxation of the same led to a spurt in growth.
certain type— for example, is it a low cost ("good shock") or high cost ("bad shock") type? Economically speaking it captures the marginal benefit associated with the optimal supply contract. The right hand side reports the derivative of information rent that has to be paid to the small firm to incentivize it to follow the recommendation of the big firm. It captures the marginal cost associated with incentive provision.\footnote{This insight has been employed extensively in the literature on contract theory, see for example Stole [2001], and auctions (where it is called virtual valuation), see for example Myerson [1981].} Our paper pins down how optimal distortions determined by the equation evolve in a dynamic contract with financial constraints.

**Structure of the Optimal Contract.** The contract offered by the big firm is described by two vectors: allocation ($q$) and expected utility ($U$) of the small firm. The latter refers to the discounted sum of lifetime utility of the small firm at any point of the contract tree. It can be broken down into two components: current (or stage) and promised (or continuation) utility. Let $h^t$ be a typical history of cost realizations. For example, $h^3 = (\theta_L, \theta_H, \theta_H)$ is a three period history where a low cost was followed by two consecutive high costs.\footnote{A low cost realization is always better for economic surplus than a high cost realization.} A plausible contract identifies quantities and expected utilities that satisfy incentive compatibility and the cash-strapped constraint along every such history. An optimal contract maximizes the big firm’s profit amongst the set of plausible ones.

The evolution of $q$ and $U$ at the optimum is explained through two pictures. Figure 1a depicts a sequence of technology shocks. For $i = L, H$, let $q(\theta_1|h^t)$ and $U(\theta_1|h^t)$ be the allocation and expected utility for cost realizations ($h^t, \theta_1$). At this point, if the right hand side of equation (\star) is zero, then $q(\theta_1|h^t) = q^e(\theta_1)$, that is the efficient quantity is supplied. If it is positive, then $q(\theta_1|h^t) = q^e(\theta_1) - d(\theta_1|h^t)$ where $d$ measures the history dependent optimal distortion.

As is standard, the low cost type always supplies the efficient quantity: $q(\theta_L|h^t) = q^e(\theta_L)$. 

![Diagram](https://example.com/diagram.png)

Figure 1: The optimal contract

(a) $q(\theta_H|h^t, \theta_L) > q(\theta_H|h^t) > q(\theta_H|h^t, \theta_H)$  
(b) Evolution of the optimal contract
On the other hand, each "bad shock" increases optimal distortions: \( q(\theta_H|h^i, \theta_H) < q(\theta_H|h^i) < q^*(\theta_H) \). This is in striking contrast to models of dynamic mechanisms without financial constraints that emphasize progressively decreasing distortions along all histories (see Besanko [1985] and Battaglini [2005]).

Allowing for a long-term contract helps mitigate the problem of agency frictions- this mitigation is achieved through backloading of incentives. Financial constraints restrict the extent of backloading. Dynamic distortions in our framework are an additive sum of two effects: backloading of incentives and illiquidity due to financial constraints; the latter increases with each "bad shock", overturning the standard result.

Moreover, the realization of a "good shock" decreases the optimal distortion: \( q(\theta_H|h^i) < q(\theta_H|h^i, \theta_L) \). An endogenous number of consecutive "good shocks", say \( n(h^i) \), are required for the optimal distortion to reach zero. For every additional "bad shock", as distortions increase, the number increases: \( n(h^i, \theta_H) \geq n(h^i) \). It can rise very quickly as consecutive bad shocks push the contract away from efficiency. Once the optimal distortion reaches zero it stays zero, that is, efficiency is an absorbing state. In the long run, the optimal contract almost surely converges to the efficient allocation.

As a function of the optimal quantities, each period the big firm provides the small firm with transfers and makes a long-term promise of expected utility: \( U(\theta_L|h^i) \). With reference to Figure 1a, the expected utilities of both the low cost and high type go up after a "good shock" and go down after a "bad shock". That is, \((U(\theta_L|h^i, \theta_H), U(\theta_H|h^i, \theta_H)) \ll (U(\theta_L|h^i), U(\theta_H|h^i)) \ll (U(\theta_L|h, \theta_L), U(\theta_H|h^i, \theta_L)) \).

Moreover, two thresholds on the vector of expected utilities divide the evolution of the optimal contract into three regions- illiquidity, liquidity and efficiency; see Figure 1b. A contract typically starts in the illiquid region- both the incentive and cash-strapped constraint bind to produce financial constraints that bite. A low cost type keeps the contract in illiquidity or can transition it to liquidity. A high cost type decreases the expected utility of the small firm which keeps it illiquid. After an endogenous number of low cost realizations, the expected utility of the small firm reaches a critical threshold at which the big firm agrees to lax the cash-strapped constraint and provide the small firm with some credit. This is called the liquid region.

Liquidity is not an absorbing state, a high cost realization can push the small firm back into illiquidity. The liquid region forms a one-step boundary towards efficiency. Once liquid, the realization of one more low cost pushes the expected utility of the small firm beyond the second threshold that propels the optimal contract into the absorbing state of efficiency.\(^6\)

At a technical level, we use a mixture of sequential and recursive approaches to characterize the optimal contract. A novelty we bring to the table is the existence of a "shell", a subset of the recursive domain which houses the optimal constrained contract. To the best of our knowledge, this is a new feature of dynamic contracts. We show that as long as the optimal contract is inefficient, the expected utility of the agent must always lie in this shell. It allows us to show all the aforementioned monotonicity properties of the evolution of the optimal contract. It

\(^6\)In a cash flow diversion model, Fu and Krishna [2016] also establish liquidity as a one-step boundary towards efficiency. However, all the short-run properties of monotonicity in \( q, U \) and \( n \) are unique to our paper.
also clearly elucidates the optimal contract in the two limiting cases: perfectly persistent and iid types.

**An Economic Implementation.** In addition to providing a complete characterization of the optimal contract, we propose an economically meaningful implementation. At the end of every period, the sum of the continuation utilities of the two firms constitutes the economic surplus. It is equal to the net present value of the "meta firm" borne out of their partnership. In keeping with literature, we call the continuation utility of the big firm (principal) its profit and that of the small firm (agent) its promised utility. The two dimensional vector of profit and promised utility constitutes the *capital structure* of the meta firm, which is endogenous to the realization of technology shocks.$^7$

At any level of promised utility, the expected utility of the small firm is marked up on the realization of a "good shock" and marked down for a "bad shock". The difference between the two values is termed the utility spread. After a history $h^{t}$, the utility spread is simply $U(\theta_t|h^{t}) - U(\theta_H|h^{t})$. It is a measure of the dynamic information rent paid to the low cost type over and above the high cost one. Using techniques from mechanism design, we can exactly pin down the utility spread from the so called dynamic envelope formula.

In order to fully characterize the capital structure we solve for the Pareto optimal contract. Each point on the Pareto frontier corresponds to a specific capital structure. As the contract evolves, it endogenously chooses points according to the history of technology shocks. In the illiquid region, the cash-strapped constraint binds and the big firm only provides *working capital* to the small firm. Through a sequence of consecutive low cost realizations, the small firm has to earn its way into liquidity. In the liquid region, the big firm promises to *take-over* small firm on the realization one more low cost type for a determinable strike price. Thereafter, the small firm operates in-house, producing the efficient quantity.

The value of the meta firm is increasing in the share of the agent. As the agent assumes a greater stake of the total surplus, incentives align with bargaining power reducing agency frictions and increasing the size of pie towards its efficient value. A less obvious observation is that along the optimal contract, the value of the meta firm is also increasing in the utility spread. Why does an increase in information rent increase the economic surplus? Each "good shock" decreases the optimal distortion which in turn increases both the information rent and surplus. A "bad shock" on the other hand, increases optimal distortions which reduces the information rent of the agent and *downsizes* the meta firm by reducing the economic surplus.

The efficient contract represents a mature meta firm that has been able to overcome financial frictions- both incentive and cash-strapped constraints are slack. However, in contrast to the iid model the value of the mature firm is Markovian. The total economic surplus fluctuates between two values depending on the last period technology shock. Obviously the corresponding value for the "good shock" is higher than that for a bad one.

---

$^7$In the corporate finance view of our model, the Modigliani-Miller Theorem does not hold since capital structure matters for the value of the firm.
**Foundation for Limited Liability.** We also provide a conceptual foundation for a productive role of limited liability in dynamic contracts. We ask: when should the positivity of stage utility of the small firm be interpreted as a limited liability constraint? The answer depends on whether the agent does better in the benchmark model without financial constraints or in our model with financial constraints.

Consider the principal profit maximizing contract on the Pareto frontier, that is the big firm has all the bargaining power. The ex ante expected utility of the small firm from the contract is determined endogenously as part of the optimum. We show that for a large measure of parameters (and almost all economically relevant ones), the ex ante expected utility of the agent is higher in our model than in the benchmark. Thus, being "protected by limited liability" helps the agent in a utilitarian sense and is a meaningful interpretation.

**Literature.** This paper connects two related literatures- dynamic mechanism design and dynamic financial contracting. Our work can be seen as an attempt to bridge the two strands. Dynamic mechanism design has predominantly been focused on (the challenging task of) modeling dynamic incentives. We arguably require more models in pushing the envelope on what Myerson calls feasibility constraints. Feasibility in our model of course manifests in the agent being strapped for cash. Incentives and feasibility interact every period, allowing us to uncover hitherto unexplored nature of dynamic distortions.

To dynamic financial contracting, we bring some newly developed techniques from the former literature. We pose familiar questions, but move them away from the moral hazard and cash flow diversion models to frameworks of screening and adverse selection. It allows us to precisely calculate the short-run properties of the optimal contract (through equation (•)), and hence the capital structure of the "meta firm". While the Kiyotaki proposal motivates our modeling choice, we stay fairly loyal to the seminal work by Clementi and Hopenhayn [2006] in our execution of it. Our paper can viewed as the Markovian dynamic screening counterpart to their iid cash flow diversion model.

This marriage of ideas in the two literatures can potentially produce many other interesting applications. Two immediately come to mind- optimal taxation and double auctions. It might be reasonable to assume, for example in the presence of incomplete credit markets, that a government cannot prevent citizens from forgoing consumption beyond a certain limit in any given work period, hence allowing a version of our cash-strapped constraint to interact with the standard private information of labor productivity in the dynamic Mirrlees model of taxation. In economic transactions captured by repeated double auctions, it might be realistic to assume that either party cannot commit to posting large upfront capital that will be honored over time. Again the cash-strapped constraint will interact with the willingness to buy or sell. We discuss these formally in the conclusion as potentials for future research.
References

[1] [partial]


