A Back-up Quarterback View of Mezzanine Finance

March 6, 2019

Abstract

A significant number of real estate transactions involve the use of mezzanine finance. Traditional arguments describe the purpose of mezzanine debt as completing the market, specifically as plugging a financing gap. We describe a completely different yet equally fundamental role for mezzanine debt which creates economic value even in a world where all investors are homogenous and risk-neutral. According to our theory, mezzanine financiers serve as substitute operating agents -- back-up quarterbacks of sorts -- ready to replace the original operator when the project underperforms. The presence of skilled mezzanine investors in the capital structure makes it cheaper to provide incentives to incumbent operators.

Keywords: Mezzanine Financing; Capital Structure
JEL codes: D47; D82

1 Introduction

"If you've never owned and operated properties, you probably shouldn't be a mezzanine lender, because you're really not well positioned to take over properties." *Bruce Batkin, CEO of Terra Capital Partners.*

A significant number of real estate transactions involve the use of mezzanine finance. As documented for instance by Stein (1997) or Rubock (2007), mezzanine loans appear to have gained importance at the expense of junior mortgages in the United States over the past three decades, and were particularly popular during pre-crisis years. One key difference between mezzanine loans and junior mortgages is that the former contract typically provides for the expeditious assumption of ownership rights by mezzanine lenders when incumbent owners are unable to meet their payment obligations. Second, as we document in this paper, mezzanine investors tend to be skilled investors, unlike the financial institutions that typically provide mortgage financing. These fundamental differences motivate the view of mezzanine finance we develop in this paper.

In the theory we propose, mezzanine investors provide an optimal blend of capital and operating skills to commercial real estate investments. Property operators have the opportunity to divert or hide rents. In two party arrangements between senior lenders and operators, the only option to mitigate the resulting moral hazard frictions is for senior lenders to threaten operators with a foreclosure when the property underperforms. Unlike senior lenders, mezzanine investors are skilled operators who can operate properties when they remain viable as an ongoing concern despite poor early performance. Including them in the capital structure makes it cheaper to provide the necessary incentives to incumbent operators. Mezzanine investors must be sufficiently skilled to play an essential role. We show that it is also optimal to require a capital contribution from them because doing so gives these investors built-in skin-in-the-game when they take over operations.

This complementarity between the provision of capital and managerial expertise has also been emphasized by Holmstrom and Tirole (1997) in a different context.[1] They build a model where firms with low net worth feature junior financiers that help monitor the firm. Whereas Holmstrom and Tirole focus on monitoring skills, our mezzanine investors provide back-up operating skills.

Our model also shares several features with the work of DeMarzo and Fishman (2007), DeMarzo and Sannikov (2006) and Biais et al. (2007). These papers combine the firm's choice of capital structure with a contract providing incentives for a manager to operate a project with the interests of the investors in mind. In DeMarzo and Fishman (2007) and DeMarzo and Sannikov (2006) the optimal contract between the principal and the manager can be implemented with straight debt, equity, and a line of credit used for temporary liquidity shortages. Biais et. al. (2007) show that cash reserves can play the same role as the line of credit in those papers. A key feature of this class of models is that with limited commitment on the part of the manager the principal may decide to liquidate the project after a sequence of bad earnings report. This occurs even though all parties clearly see that the manager's reports are genuine and liquidation is ex-post inefficient. Commercial real estate features all the key frictions that motivate this literature. In particular, mortgage laws afford borrowers myriad protections making the foreclosure process especially onerous. Our main point is that this makes for an environment where adding skilled operators to the capital structure is likely to be particularly valuable.

The amount of funds provided by the mezzanine layer is often quite modest relative to the project's total funding requirements. Yet, it is critical to the project's viability. The mezzanine investor's capital at inception is not motivated by the need to plug a financial gap. Senior lenders could finance the venture on their own. Neither is complicating the capital structure an attempt to cater to investors with different preferences as in Allen and Gale

[1]Hart and Moore (1994) also emphasize the inalienability of human capital provided by entrepreneurs as a key element of the optimal capital structure of a business venture.

(1988): all investors in our setup are risk neutral. The capital contribution makes it cheaper to provide the right incentives to the standby operator if and when he takes over the project. Simply put, incentives, not capital shortages, justify the presence of mezzanine investors in the capital structure of the project.

The corporate finance literature contains many papers that study the role of junior financing in firm's capital structures. Among recent papers, Repullo and Suarez (2004) consider a project that requires the simultaneous effort of an entrepreneur and the advice of a venture capitalist. With unobservable effort and advice, double moral hazard shapes the design of the venture capital contract. Casamatta (2003) analyzes the optimal capital structure when the efforts of the venture capitalist and the entrepreneur are substitutes. While these models deal with the problem of security design when entrepreneurs and capitalists with different preferences and information work together in a team, in our model the capitalist does not interfere at all as long as the manager performs satisfactorily. Our focus is on an optimal contract that purposely includes a backup manager to ensure both better pre-turnover performance and smooth continuation when turnover occurs. Roberts and Sufi (2009) and Rauh and Sufi (2010) show that firms tend to have elaborate capital structures, with different layers of debt and quasi-equity securities and find that riskier firms have more complex debt structures. Our model provides one broad reason for capital structure complexity, and does so while explaining the logic for mezzanine finance, an intermediate source of financial capital that is inalienable from the unique role human capital plays in the project.

Despite its growing importance, mezzanine finance has received surprisingly little attention to date in the real estate literature. As Pagliari (2017) explains, a likely reason for this gap is the lack of data on the volume and risk-return characteristics of mezzanine debt. Mezzanine debt typically trades privately and the fact that contracts are tailored to each deal and set of investors (as our theory implies they should be) makes aggregation difficult even when the data is available. Growing demand from investors has led to the recent introduction of indices that track the performance of junior commercial real estate debt, including mezzanine loans (see, e.g., Giliberto, 2018). As more data on mezzanine performance become available, we expect this glaring gap in the literature to be quickly filled and view our contribution as a step in that direction.

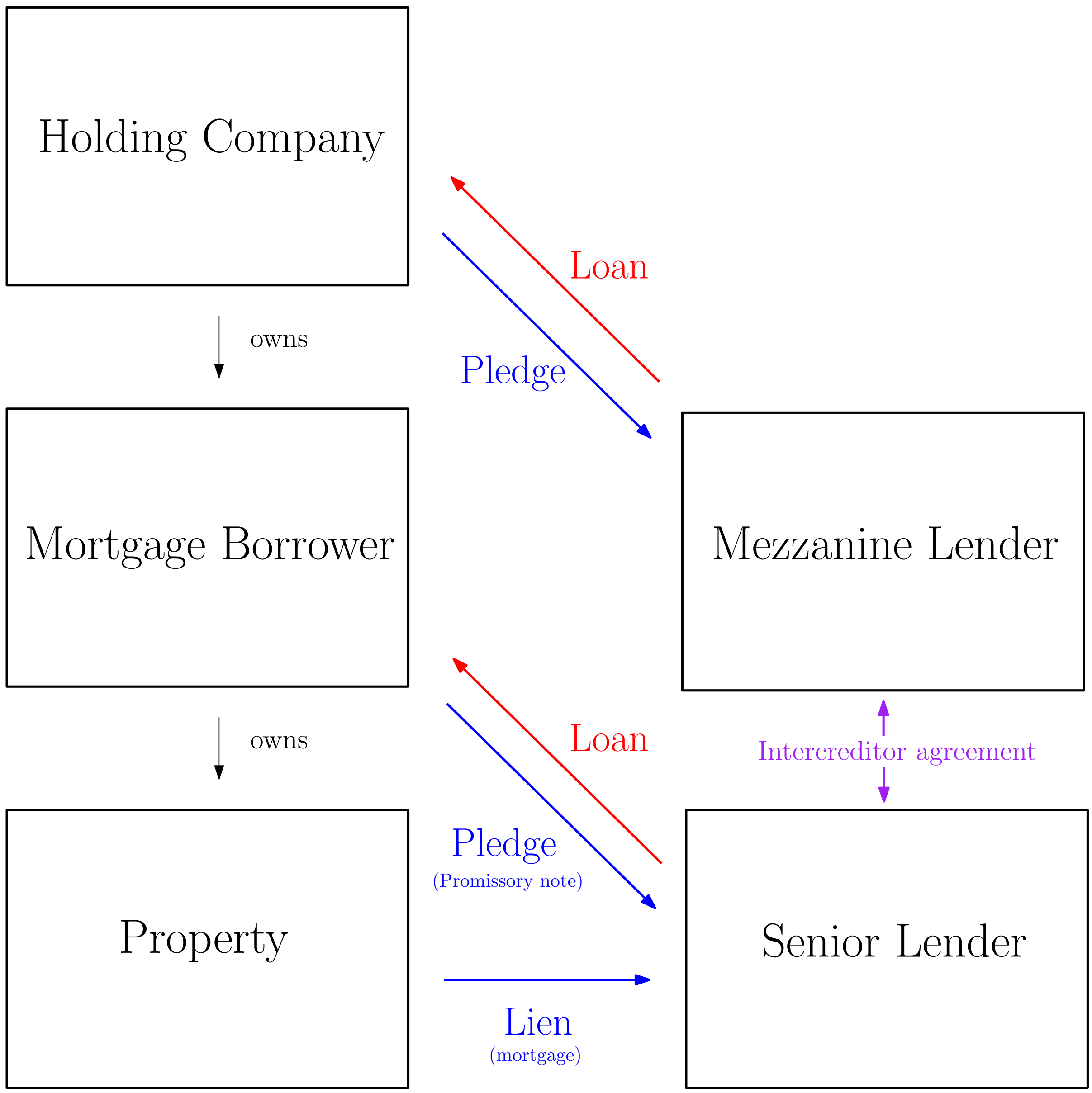
The rest of the paper is organized as follows: Section 2 documents the distinguishing characteristics of mezzanine contracts and mezzanine investors in commercial real estate. Section 3 describes an environment that features the fundamental frictions that characterize commercial real estate transactions, namely moral hazard on the part of operators and costly termination options for senior lenders. Section 4 characterizes the optimal bilateral arrangement between senior lenders and operators. Section 5 establishes the essential role of mezzanine investors. Section 6 discusses various natural extensions as well as the main testable implications of our theory. Section 7 concludes.

2 Mezzanine finance in commercial real estate

Commercial real estate (CRE) transactions feature significant asymmetric information such as unobservable effort on the part of owners and operators, not to mention their ability to divert or hide rents. In addition, the foreclosure process that protects the rights of senior lenders is slow, onerous, prone to disputes and usually results in heavy ex-post losses. Taking over and operating properties is an option for senior lenders but they tend to be institutions with limited expertise and operating capacities. Our primary point in this paper is that this combination of frictions create an environment where skilled, junior contributors to the capital structure play an essential role.

Figure 2 shows a schematic representation of a property purchase in real estate that involves mezzanine finance. Capital in this scenario comes from three different sources: equity from a mortgage borrower and owner of the property, a loan from a senior lender to the borrower, and a mezzanine loan. The mezzanine loan is issued not to the mortgage borrower but to a holding company that owns the mortgage borrower. The relation between the senior lender and the mortgage borrower is governed by two distinct documents: a promissory note and a mortgage (or deed of trust). The promissory note stipulates loan payments and all

Figure 1: Mezzanine Finance in Commercial Real Estate



subsidiary obligations of the borrower, such as commitments to keep the property in good shape or enter into insurance contracts for standard property hazards. Default occurs when any of the contracting clauses is violated.

In the event of default, the deed of trust contains acceleration clauses that give the senior lender the right to demand the entire loan balance by initiating a foreclosure sale process. Foreclosure on commercial real estate collateral is governed by mortgage laws that differ across states, but typically provide for mandatory redemption periods and other borrower protections that make the acceleration process costly. As Gertler et al. (2007, pp 398-99) discuss, it is not unusual for the foreclosure process to exceed one year, legal expenses alone can reach ten percent of the loan balance, and the borrower has limited incentives to spend on maintenance during the lengthy foreclosure process causing the property to deteriorate at a fast rate. These direct costs alone can amount to over thirty percent of the outstanding loan balance at the time of default.

Mezzanine contracts in real estate usually stipulate specific payment obligations but, unlike mortgages, they are secured by an equity interest in the entity or holding company that owns the mortgage borrower.[2] One key implication is that the mezzanine lender's collateral is usually treated as personal property rather than a general tangible claim under the relevant law, which results in the mezzanine lenders taking possession of the collateral under article 8 of the Uniform Commercial Code, an action that is markedly more expeditious (the process usually takes a few weeks at the most) and less costly than foreclosures under state mortgage laws. Furthermore, an intercreditor agreement between mezzanine lenders and senior lenders usually stipulates that in the event of payment difficulties, mezzanine lenders have the option to take over the property as long as they commit and manage to meet the remaining payments owed to senior lenders.

The efficiency of collateral repossession distinguishes mezzanine finance not just from senior mortgages but also from other forms of intermediate claims. For example, preferred

[2]See Berman (2013) for a detailed discussion of the legal framework that governs mezzanine finance in real estate.

equity owners do not have any foreclosure rights or specific collateral claims. In their case, recourse is so limited that a promised dividend can be suspended by a simple vote of the board of directors without any risk that the holding company will be sued. Remedies for junior mortgage owners, for their part, fall under the same onerous mortgage laws as senior mortgages. Much worse from the viewpoint of senior lenders, junior lenders can trigger the highly costly foreclosure process when the borrower is delinquent regardless of the status of the first loan. Their presence further complicates and lengthens the default process and renders non-litigious dispute resolutions more difficult to achieve and subject to holdup problems.[3] From that point of view, mezzanine lenders are highly preferable since their foreclosure rights do not alter the senior lenders' resolution rights in any way.

Precisely because of the associated risks for senior loans, rating agencies penalize the presence of subordinated debt such as second mortgages more heavily than they penalize the presence of mezzanine financing when rating first mortgages.[4] Not surprisingly, as junior mortgage volumes have fallen, mezzanine volumes have risen as the importance of mortgage-backed securitization has grown. Whereas different mortgage liens interact and affect the value of senior lenders' claims to collateral assets, the foreclosure rights of mezzanine loans and senior loans do not intersect.

In our model, the optimal contract calls for junior claim-holders with the ability to operate the property if the original owner underperforms. As the paper's opening quote illustrates, operating capacities are in fact viewed as a *sine qua non* feature of mezzanine finance providers. To document this more systematically, we compiled a list of the most prominent private providers of mezzanine financing in the United States.[5] Since these are private corporations it is not possible to know for sure that the list contains all the largest private providers of

[3]See Stein (1997) for a detailed discussion.

[4]As Rubbock (2007) explains, "this reduced penalty [on mezzanine loans] became for lenders one incentive to cast additional leverage as mezzanine debt rather than as a B-note, so as to minimize the negative credit impact on the rated senior loan. CRE mezzanine debt placed into CRE CDOs increased from $25.6 million in 2004 to approximately $3.22 billion in 2006, and is on a dramatic upward swing."

[5]The set of publicly traded providers of mezzanine finance in real estate comprises mostly listed Real Estate Investment Trusts – REITs – and those obviously have operating capacities. In fact, by law, most REIT assets must be real estate assets the lion share of their income must come from real estate.

Table 1: Prominent private mezzanine lenders in the United States

Firm	Real Estate specialist	Owns properties	Sponsors equity funds	Operating capacity	Top Management has operating background
AEW Capital Management	✓	✓	✓	✓	✓
Apollo Commercial	✓		✓		✓
ARC Realty Finance Trust	✓				✓
Ares	✓		✓		✓
Artemis Realty Capital	✓				✓
Clarion Partners	✓	✓	✓	✓	✓
Cornerstone Real Estate Advisers	✓	✓	✓	✓	✓
Dominion Mortgage Corporation	✓				✓
Everwest Real Estate Partners	✓	✓	✓	✓	✓
Federal Capital Partners	✓	✓	✓	✓	✓
George Smith Partners	✓				✓
Harbor Group	✓	✓	✓	✓	✓
KKR-Real Estate	✓	✓	✓		✓
Ladder Capital	✓				✓
LEM Capital	✓	✓	✓	✓	✓
LoanCore Capital Markets	✓		✓		
Lowe Enterprises Investors	✓	✓	✓	✓	✓
Mack Real Estate Group	✓	✓		✓	✓
Mesa West Capital	✓				✓
NorthStar Realty Finance	✓	✓	✓	✓	✓
Pearlmark Real Estate Partners	✓	✓	✓	✓	✓
Related-Real Estate Fund Management	✓	✓	✓	✓	✓
Redwood-Kairos	✓	✓		✓	✓
Rockwood Capital	✓	✓	✓	✓	✓
Square Mile Capital Management	✓	✓	✓		✓
Stonebeck Capital	✓				✓
Starwood Property Trust	✓	✓	✓		✓
Strategic Realty Capital LLC	✓	✓	✓	✓	✓
Terra Capital Partners	✓				✓
Torchlight Investors	✓				✓
W Financial	✓				✓
Witkoff Group	✓	✓	✓	✓	✓
Wrightwood Financial	✓	✓			✓

mezzanine loans but conversations with top managers at prominent mezzanine firms suggest that the list does in fact cover the immense majority of private mezzanine lending in the United States.[6]

The key question for our purposes is whether these firms tend to have in-house or easy access to operating capacities. To answer that question we searched through the documentation those private firms make available online for direct evidence that 1) they are real estate specialists, unlike the typical financial intermediary that provides senior funding, 2) they own properties or 3) sponsor equity funds, which implies that, at the very least, they have relationships with operating partners, 4) they directly operate properties or, finally, 5) that they are managed by top executives that have some experience in operations.[7] As table 1 shows, all mezzanine providers on our list are real estate specialists and all but one are managed by executives that have some experience in operations. Most own properties directly or sponsor equity funds and a majority of mezzanine lenders actually provide operating services to other investors. The bottom line is that mezzanine lenders, unlike senior lenders, are highly skilled investors, as our model predicts they should be.

3 The environment

We consider an environment where, as is often the case in commercial real estate transactions, 1) senior lenders lack operating skills but have the ability to provide the bulk of the necessary financing, 2) invoking foreclosure rights is costly and, 3) there are agency frictions between operators and capital providers. To capture those frictions in a parsimonious fashion, consider an economy with three dates $t = 0, 1, 2$. A property can be built or purchased at date 0 at a capital cost we normalize to one unit of the unique good.

There are two potential operators for the property each endowed with a quantity $\omega < \frac{1}{2}$

[6] We are especially grateful to Tom McCahill, Managing Principal for Mezzanine Finance at EverWest Real Estate Partners, for his help in this respect.

[7] The online appendix (available for download at (blinded) provides specific documentation for each lender in the table.

of the unique good at date 0. When the property is managed by the first operator, it yields rents R with probability π at date 1. With probability $1-\pi$ the property is vacant and yields no rent.[8] After rents are observed at date 1 the property can be sold to outside investors or repurposed for a net value of S. Since we think of this value as including any and all transaction costs associated with early termination, S could be negative and all our results allow for this possibility. If the property remains managed by operator 1 at date 2, it once again yields rents R with probability π, and nothing otherwise. Rent outcomes are i.i.d across periods, although no result of substance would change if property performance was correlated across periods.

The second operator also has the ability to run the property but is not as efficient at generating rents. For one interpretation of this gap in productivity, it may be that operator 2 is less familiar with local market conditions and less well-connected than operator 1. Formally, when operator 2 is at the helm, the project yields θR instead of R when successful, where $\theta \in [0,1]$. One could instead model operator heterogeneity as differences in success rates without affecting our key findings. Both operators have linear preferences and neither discounts the future. They can store their endowment and earn a risk-free net return $r \geq 0$ at date 2. When they do not operate the property, the value of the best alternative use of each operator's time is $V_o \geq 0$.

Because $2\omega < 1$, operators cannot finance the purchase without outside financing even it they pool their resources. But the economy contains a risk-neutral agent with no operating skills but deep pockets – the senior lender, henceforth. She has access to a risk-free storage technology that earns return $r \geq 0$ at date 2 and pins down the opportunity cost of her funds. In order to obtain the necessary financing from the senior lender, operators must offer her a contract which covers this opportunity cost in expected terms. We will first consider the case in which operator 1 must offer a two-party contract in which operator 2 plays no role. We will then consider three-party contracts.

Contracting between senior lenders and operators is limited by several frictions. First,

[8] Assuming a more complex support for rents would not change any of our results.

only the operator in charge of the property observes rents. She has the option to divert those rents at a proportional cost $\phi \in [0, 1]$.[9] The cost proxies for the time and resources the operating manager has to spend in diverting funds. Furthermore, operators cannot commit ex-ante to managing the project at date 2 hence they must expect at least the value $V_o \geq 0$ of their outside option in remaining payoff from any arrangement. Lenders for their part, can commit to any two-period arrangements, including ex-post inefficient actions if needed.

4 Optimal two-party contracts

Consider first the case where operator 1 is limited to contracting only with the senior lender. A two-party contract must stipulate a mortgage loan $m \leq 1$ contributed by the senior lender and an amount $e = 1 - m \leq 1$ of equity contributed by the operator. Second, it must specify a mortgage payment schedule $\{q(h) \leq R\}$ for all possible histories h of rent messages at dates 1 and 2. Finally, the contract states termination probabilities $(\tau(0), \tau(R))$ that depend on the two possible rent realizations at date 1.

It will soon be clear that if termination ever happens, it happens only when the property underperforms at date 1. In effect, the contract features an acceleration clause which says that if the operator is unable to make mortgage payment $q(R)$, the senior lender then has the right to request that the property be sold. The termination process is therefore triggered by failure to pay on the part of the borrower, as is the case in practice. Typically, at the optimal contract, it will turn out that $\tau(0) < 1$ so that failure to pay is necessary but not sufficient to trigger termination.

Formally, termination is an action selected by the lender. The operator may choose/agree to include acceleration/termination clauses in the contract he proposes because their presence optimally mitigates moral hazard frictions hence make it easier to satisfy the lender's participation constraint. Because the lender cannot observe rents directly, she must rely on

[9] Instead of assuming that the operator can hide rents at a cost, one could assume that unobservable effort is necessary to generate rental income. The resulting environments are isomorphic for our purposes.

reports from the operator. A standard appeal to the revelation principle tells us that we can concentrate our attention on direct revelation contracts without any loss of generality.

We require mortgage payments to be below rents which is equivalent to assuming that all capital contributions to the project by the operator are made at date 0. This is also without loss of generality since both parties are equally patient and the lender has the ability to commit to any payment arrangement, including the exchange of actuarially fair intertemporal transfers. Given a contract, let

$$V_2(y) = (1 - \tau(y)) \left[\pi(R - q(y, R)) + (1 - \pi)((R - q(y, 0))\right] + \tau(y)V_0$$

denote the utility expected by the operator as of date 2 when rent message $y \in \{0, R\}$ is issued at date 1. The mortgage payment at date 2 may depend on the two rent messages issued by the operator by that time. This expression for V_2 reflects the fact that if the operator does not manage the property, he enjoys his outside option, and nothing more. Turning now to the onset of the project, the value of the contract to operator 1 is:

$$V_1 = \pi \left[R - q(R) + V_2(R)\right] + (1 - \pi) \left[R - q(0) + V_2(0)\right] + (\omega - e)(1 + r), \tag{4.1}$$

where, as stated in our definition of a contract, $q(y)$ is the mortgage payment at date 1 and depends on the first message $y \in \{0, R\}$. Operator 1's objective is to make V_1 as high as possible subject to incentive compatibility (direct revelation) and the fact that the contract must cover the senior lender's opportunity cost.

For direct revelation at date 1, we need:

$$R - q(R) + V_2(R) \geq V_2(0) + (1 - \phi)R. \tag{4.2}$$

For all $y \in \{0, R\}$ such that $\tau(y) < 1$, remaining expected payoffs when the property is continued at date 2 must meet participation constraints:

$$V_o \leq V_2^c(y), \tag{4.3}$$

and incentive compatibility constraints

$$V_2^c(R) \geq V_2^c(0) + (1 - \phi)R \tag{4.4}$$

where for $y \in \{0, R\}$,

$$V_2^c(y) = \pi(R - q(y, R))$$

is the payoff to the operator given continuation.

The senior lender's expected payoff, in long form, equals

$$W_1 = \pi\left[q(R) + \tau(R)S\right] + (1-\pi)\tau(0)S + \pi^2(1-\tau(R))q(R, R) + (1-\pi)\pi(1-\tau(0))q(0, R) - m(1+r).$$

The first four terms of the above expression correspond to each of the nodes at which the contract calls for a mortgage payment from the operator to the senior lender, and are weighted by the corresponding probabilities. Participation on the part of the senior lender boils down to

$$W_1 \geq 0. \tag{4.5}$$

The operator seeks to maximize his expected payoff subject to conditions (4.1 − 4.5). Since this amount to maximizing an objective (4.1) which varies continuously with contract terms over a compact set of feasible contracts, a solution to this problem exists. Let V^N be the resulting expected utility for the operator, where the superscript stands for no mezzanine. This value varies continously with the set (R, π, ϕ, r) of parameters that define our environment. In fact, it should be clear that V^N rises continuously with R and π but and fall continuously as ϕ or r rise. A profitable contract exists – hence an investment in the property is made –

if and and only if

$$V^N(R, \pi, \phi, r) \geq 2V_o + \omega(1 + r).$$

To characterize optimal contracts in details, we will proceed recursively. Under any contract, the operator enters the final period with an expected utility level $V_2 \geq V_o$. Given V_2, the maximum payoff the senior lender can expect is:

$$W_2^c(V_2) = \max_q \pi q$$

subject to:

$$\pi(R - q) = V_2 \text{ (promise keeping)},$$

$$R - q \geq (1 - \phi)R \text{ (truth telling)},$$

and

$$q \leq R \text{ (limited liability)}.$$

In principle, the operator could attempt to build a contract that delivers them more than V_2 in expected utility as of date 2. But the senior lender has the ability to commit to any promises and actions at date 2. Contracts that commit the senior lender to potentially ex-post inefficient actions can make it cheaper to provide the right incentives to the manager over the life of the arrangement, as we will establish below. The operator may agree to stipulations which they know they would like to eventually renegotiate because doing so potentially makes the set of contracts that satisfy conditions $(4.1 - 4.5)$ larger. In particular, there may exist no contract that satisfy lender participation without termination at date 0, even though termination is ex-post inefficient.[10]

This statement of the second period problem anticipates on the fact that it is optimal

[10] In practice, this corresponds to financial institutions invoking their foreclosure rights even in cases where transaction costs are so high that renegotiation would clearly be the profitable option ex post. If operators know that foreclosure is too costly to be invoked ex post and without commitment, the threat of foreclosure has no effect on incentives. When moral hazard frictions are strong enough, commitment to ex post inefficient actions becomes the only option.

for the senior lender not to make a transfer to the operator when the project fails. Making such a transfer would make it more costly to give the operator the incentives they need to report successful outcomes. This immediately implies that $q = R - \frac{V_2}{\pi}$ is optimal provided this makes q low enough to meet the truth-telling constraint, which holds as long as

$$\frac{V_2}{\pi} \geq (1-\phi)R.$$

Otherwise, the feasible set is empty. This appears to suggest that if he wishes to continue the project, the operator cannot propose contracts that delivers him less than $\pi(1-\phi)R$ in remaining utility following period 1's rent announcement. But, in fact, he has a broader set of options.

To see this, note that when the project is terminated at the end of period 1, the expected remaining value to the lender is $W_2^S(V_2) = S + V_o - V_2$. Indeed, the operator must receive a payoff of $V_2 - V_o$ since he expects V_2 under the contract but only gets V_o from his outside option. If $S + V_o \geq \pi\phi R$ interruption is always optimal for both parties at date 2. Henceforth, we will focus on the more interesting case where $S + V_o < \pi\phi R$, so that termination is ex-post inefficient. In that case, exercising the early termination option only makes sense if the contract delivers less than $\pi(1-\phi)R$ to the operator.

But how can a contract feasibly deliver less than $\pi(1-\phi)R$ in expected terms when that promise is the lowest possible promise consistent with incentive compatibility at date 2? The optimal solution is to randomize between termination and no termination when $V_2 \in (V_o, \pi(1-\phi)R)$. More precisely, in the closure of that interval, the optimal termination probability is:

$$\tau(V_2) = \frac{\pi(1-\phi)R - V_2}{\pi(1-\phi)R - V_o},$$

while it is zero everywhere else. The operator gets a payoff of V_o if the project is terminated but $\pi(1-\phi)R$ otherwise. As a result, the overall remaining payoff to the senior lender under the contract proposed by the operator and following the first output announcement is:

$$W_2(V_2) = \tau(V_2)S + (1 - \tau(V_2))W_2^c\left(\max\left\{V_{2,}\pi(1-\phi)R\right\}\right).$$

This value function is concave, strictly increasing in the range $[V_o, \pi(1-\phi)R]$ and thereafter strictly decreasing with a slope of -1. The fact that W_2 decreases in the range $[V_o, \pi(1-\phi)R]$ means that scrapping is ex-post inefficient. Yet, termination may well be part of the ex-ante optimal contract because threatening the operator with termination may be necessary to provide him with the right incentives in period 1.

Formally, consider the senior lender's expected value as of date 1 given an expected utility V_1 for the operator under the proposed contract. Given the initial investment $m \geq 0$, the senior lender's payoff is, at best:

$$W_1(V_1|m) = \max_{q,V_2^L,V_2^H} \pi\left[q + W_2(V_2^H)\right] + (1-\pi)W_2(V_2^L) - m(1+r)$$

subject to:

$$\pi\left[R - q + V_2^H\right] + (1-\pi)V_2^L \geq V_1 \text{ (promise keeping)},$$

$$R - q + V_2^H \geq V_2^L + (1-\phi)R \text{ (truth telling)},$$

$$q \leq R \text{ (limited liability)},$$

and

$$V_2^L, V_2^H \geq V_o \text{ (lower bound on agent payoff at date 2)},$$

where (V_2^L, V_2^H) are the payoffs the contract delivers the operator at date 2, as a function of whether the output realization is low or high in the first period. The following result provides a complete characterization of the two-party contracts that maximize the operator's surplus given the senior lender's participation constraint.

Proposition 4.1. *Optimal two-party contracts satisfy the following properties:*

1. *If and only if* $W_1(V_1|1-\omega) \geq 0$ *for some* $V_1 \geq 2V_o + \omega R$ *then the property is purchased and an optimal contract exists such that* $e = \omega$, $m = 1-\omega$ *and* $W_1(V_1|1-\omega) = 0$ *;*

2. *The contract is terminated with positive probability at date 1 if and only if*

$$V^N < 2\pi(1-\phi)R;$$

3. *If* $V^N < 2\pi(1-\phi)R$ *then all optimal contracts satisfy* $e = \omega$ *and* $m = 1-\omega$

Proof. It is at least weakly optimal for the operator to commit his endowment to the project since the senior lender can (if nothing else) store that endowment at the same rate as operators. The first item of the proposition follows. To establish the second item of the proposition, assume that $V^N < 2\pi(1-\phi)R$ but, by way of contradiction, that the probability of termination is zero. This requires that $V_2^L \geq \pi(1-\phi)R$. But truth telling then implies that

$$R - q + V_2^H > \pi(1-\phi)R + (1-\phi)R,$$

so that, in turn, the operator's expected payoff must satisfy:

$$V^N > \pi\left[R - q + V_2^H\right] + (1-\pi)V_2^L \geq \pi(1-\phi)R + \pi(1-\phi)R = 2\pi(1-\phi)R,$$

which is the contradiction we sought. If on the other hand $V^N \geq 2\pi(1-\phi)R$ then the contract can deliver an expected payoff higher than $\pi(1-\phi)R$ so that the project can be operated in both period without violating any incentive compatibility constraint.

To establish the final item simply note that if $V^N < 2\pi(1-\phi)R$ but that, by way of contradiction, $e < \omega$, a higher equity commitment could be required from the operator which would enable the contract to make V_2^L higher without lowering the senior lender's surplus. Since this lowers the probability of termination strictly without changing the senior lender's payoff, the operator's payoff must rise. □

To see that the case $V^N < 2\pi(1-\phi)R$ is possible, so that termination can be part of optimal two-party contracts, assume by way of example that $\omega = V_o = 0$. At expected payoff $V_1 = 2\pi(1-\phi)R$ for the operator, the senior lender's payoff is

$$\pi R + \pi R + (1-\pi)S - 2\pi(1-\phi)R - (1+r).$$

Indeed, the senior lender must provide all the needed financing but, on the other hand and as established in the previous proposition, the property can be operated with probability one in both periods since $V^N \geq 2\pi(1-\phi)R$ making gross expected surplus from operating the property equal to $\pi R + \pi R$.

Instead, a contract can always be written so that the property is shut down no matter what following a bad performance in period 1. The senior lender's surplus, in that case, is

$$\pi R + \pi^2 R + (1-\pi)S - \pi(1-\phi)R - (1+r). \tag{4.6}$$

To understand this expression, note that in that case the lender can set $V_2^L = 0$ so that truth telling in period one only requires setting

$$R - q(R) + V_2^H = (1-\phi)R \tag{4.7}$$

which can be done by setting $q(R) = R$ and $V_2^H = (1-\phi)R$. This level of V_2^H is enough to meet truth telling in the second period following success in period 2. Putting all this together implies that $V^1 = \pi(1-\phi)R$ at this alternative contract.

As long as π is sufficiently close to one, payoff (4.6) is lower than payoff (4.7). In fact, we can choose parameters so that payoff (4.6) is negative while payoff (4.7) is positive. In other words, incentive compatibility requires too high a payoff for the operator to be compatible with participation by the lender unless the threat of termination at date 1 is part of the contract. When π is close to one, that threat is cheap to invoke (little total surplus is destroyed in ex ante terms) but effective in discouraging misreporting by the operator in period 1. When (4.6) is negative while (4.7) is positive, we must have $V^N \in [\pi(1-\phi)R, 2\pi(1-\phi)R)$ and, therefore, termination must be a positive probability event at the optimal contract.

5 The essential role of mezzanine investors

Termination is an ex-post inefficient outcome which operators may make a part of the optimal contract because moral hazard frictions cannot be sufficiently mitigated otherwise. The same incentives can be provided more cheaply by inviting a mezzanine investor with operating skills into the capital structure, as we establish in this section. We will also show that when mezzanine investors are part of the optimal contract, it is optimal to request a capital contribution from them. Mezzanine finance, in this sense, is an optimal blend of capital and operating skills.

To formalize this, consider contracts that stipulate that, if no rents are generated in the first period, the incumbent operator may be replaced by operator 2 with positive probability. This tri-party contract must now stipulate an original contribution to the capital structure by both operators and, like in the previous section, mortgage payment schedules for every possible history.[11] Denoting the capital contributions by the two operators by e_1 and e_2, respectively, it is at least weakly optimal to set $e_1 = e_2 = \omega$. Recall that that $2\omega < 1$ so that even if they team up and pool their funds, the two operators still need the senior lender to get the project started. We will assume that when drafting a tri-party contract operator 1 retains all the bargaining power so that they only need to make sure that operator 2 gets an expected payoff at least as high as the value of his outside option.

The optimal two-party contract we described in the previous section provides the senior lender with an acceleration clause that allows them to foreclose the project if no payment is made in period 1. In the three-party contract we now describe, the mezzanine lender is given the right to assume ownership of the property when no cash-flows are generated in period 1. In practice, there may be direct costs associated with complicating the contract

[11] Bringing a third agent into the contract allows in principle for other contingencies as well but if mezzanine investors create value, it must be by taking over following a bad announcement in the first period.

in this fashion, or simply with searching for mezzanine investors. We will write $c \geq 0$ for those costs. Including mezzanine investors in the capital structure is only profitable if doing so raises operator 1's surplus by more than this cost, as we will discuss in the next section.

An indirect cost of resorting to a back-up operator is that operator 2 may not be as productive as operator 1. Our main result below shows that the gap in operating skills between the two operators is a key determinant of whether tri-party contracts create value. For instance, if θ is such that $\phi\pi\theta R \leq S$, then involving manager 2 does not expand the set of options for operator 1 and the senior lender. More generally, expected output when the back-up option is used is

$$\pi R + \pi^2 R + (1-\pi)\pi\theta R,$$

which is below the expected output $\pi R + \pi R$ the project generates when the senior lender commits to continuing the project with the initial operator. Our next result provides conditions under which operator 1's gross surplus increases when they involve operator 2 in the financing contract.

Proposition 5.1. *The maximal payoff V^M operator 1 can generate with a mezzanine investor in place exceeds the maximal payoff V^N he can generate with two-party contracts if and only if:*

1. $V^N < \pi(1-\phi)R + \pi(1-\phi)R$, *and*

2. θ *is sufficiently close to* 1.

Proof. When $V^N > \pi(1-\phi)R+\pi(1-\phi)R$ the senior lender can commit to letting the property run for two periods, and incentive compatibility constraints have slack. If, on the other hand,

$$V^N < \pi(1-\phi)R + \pi(1-\phi)R. \tag{5.1}$$

then the optimal contract features termination with positive probability. A contract which instead of selling the property following bad performance in period 1 replaces the original operator with operator 2 gives the senior lender a surplus of $\phi\pi\theta R$ while operator 2 expects $(1-\phi)\pi\theta R$. As long as θ is sufficiently close to one this gives the senior lender a strictly higher expected payoff than the optimal two party contract. In turn, this means that all mortgage payments can be strictly lowered without violating participation by the senior lender or incentive compatibility for the operator, which means that operator 1's expected payoff goes up strictly as well, and completes the proof. □

Unless incentive compatibility constraints are never binding at the optimal two-party contract, the incumbent operator is better off adding a back-up operator to the capital structure at date 0, provided the operating skills of the second manager are sufficiently high. In other words, making the termination option cheaper to invoke for the senior lender enables operator 1 to extract more surplus from the contract. While this may seem surprising at first glance, the intuition is simple. When a better termination option becomes available, the original contract remains incentive feasible but the senior lender's surplus increases. This means that at the original contract, the senior lender's participation constraint has slack. Operator 1 can turn this slack into lower mortgage payments at every contingency when they control the property and owe a payment.

The proposition states that gross surplus increases with a back-up operator in the contract. However, bringing a third party into the contract is only profitable provided this increase in gross surplus exceeds the cost of adding a new contractual party, i.e $V^M - V^N > c$. The next section will discuss conditions under which this condition is more likely to be met.

A critical question proposition 5.1 leaves unanswered is the timing of operator 2's involvement. Can the senior lender wait to discover whether a back-up operator will be needed after operator 1 fails, or is it essential that the second operator's capital contribution to the contract be made before that uncertainty is realized? Our next result says that it is in fact essential that the mezzanine investor be part of the original capital structure.

Proposition 5.2. *If $\omega > 0$ then all contracts with a back-up operator involve $e_2 > 0$. Fur-*

thermore, if and only if

$$V_o + \omega(1+r) < \pi(1-\phi)R$$

then a strictly positive fraction of the capital commitment e_2 *must take place BEFORE date* 1 *uncertainty is resolved.*

Proof. A back-up operator is part of the optimal contract if and only if

$$2V_o + \omega(1+r) < V^N < 2\pi(1-\phi)R.$$

The first inequality is the participation constraint for operator 1 while the second is the condition under which back-up operators create value. The first condition implies $V_o < \pi(1-\phi)R$. Should $\epsilon_2 = 0$ then operator 2's participation constraint has slack so that the contract can request a strictly positive commitment of capital from manager 2 at date 0 without changing any of the subsequent payoffs. If $V_o + \omega(1+r) < \pi(1-\phi)$, then, even if they provide $e_2 = \omega$ once uncertainty is resolved, the participation constraint still has slack. In that case, the senior lender is strictly better off requesting at least part of e_2 before uncertainty is resolved and exchanging those promises for a higher payoff if and only if operator 2 is called upon. Since this raises the senior lender's surplus, operator 1 can then propose a contract that lowers mortgage payments strictly. □

Proposition 5.2 establishes the indissociability of operator 2's human capital and financial capital. From the point of view of this argument there is a key distinction between the two elements of operator 2's outside option, V_o and $\omega(1+r)$. Operator 1 would like operator 2 to commit both elements early to the contract so that payments can be optimally allocated to the node where the participation constraint has slack. But the first part is inalienable and cannot be transferred to the senior lender early on, whereas operator 2's endowment of capital can. The optimal contract takes full advantage of that second portion.

In practice, the option to foreclose on a mezzanine claim to ownership typically lies with the mezzanine investor. One implementation of the optimal contract this section has described stipulates a arbitrarily small payment from the original operator to the mezzanine lender at date 1 and an acceleration clause which the mezzanine investor can invoke when missed. When $R = 0$, the operator misses that payment and operator 2 is then granted the right to assume ownership.

6 Key assumptions and testable implications

This section considers the role of several key assumptions we have maintained so far and makes explicit some of our theory's main testable implications.

6.1 Operator vs. lender control

We have assumed that operator 1 is in charge of designing the contract and gets to make a take-it-or-leave-it offer to the the senior lender and to operator 2. In many ways, this is without loss of generality. Operator 1 maximizes their surplus subject to participation constraints for the other two agents. Standard duality results imply that the contracts we describe would remain optimal if the senior lender selected a contract subject to the restriction that operator 1 must expect a payoff V^N in the two-party case and V^M in the tri-party contract.

Likewise, while the contract proposed by operator 1 specifies contingencies that involve actions by other agents (termination, in particular), this does not require that they have control over that decision once that contingency is reached. It is enough for the contract to stipulate an acceleration clause that gives the senior lender the right to foreclose in the event that no mortgage payment is received at date 1, as mortgage contracts do in practice. It is then optimal for lenders to commit to exercise this option even if it turns out to be ex-post inefficient. The key assumption – as in all models in the spirit of De Marzo and Fishman (2007) – is that lenders have the ability to commit. In practice, Black et al. (2017) estimate that the frequency of CRE loans extension between 2012 and 2014 in the United States was under 2 percent for securitized loans and around one third for portfolio loans (loans issued and kept by banks.) Commercial mortgage lenders do invoke their foreclosure rights in the majority of default cases.

The fact that the operator initiates the contract offer rather than the senior lender does have one important implication, however. When the senior lender makes a take-it-or-leave it proposal to operator 1, hiring the services of a back-up operator may be beneficial even when the optimal two-party contract features no termination. Indeed, at the optimal two-party contract the termination option may be so costly to invoke that, instead of using that option, the senior lender may choose to propose a contract such that the participation constraint of operator 1 has slack. With a back-up operator in place, the senior lender may be able to reduce that slack. Mezzanine investors create value for senior lenders, therefore, even in cases where the original contact features zero risk of inefficient termination. The working paper version of this paper[12] provides a full treatment of the case where the senior lender initiates the contract offer. Formally, we show:

Proposition 6.1. *When senior lenders make take-it-or-leave it offers to the first operator, the optimal two-party contract features termination with positive probability if and only*

1. $2V_o + \omega R < 2\pi(1-\phi)R$, *and,*

2. $\pi - (1-\pi)\frac{\phi\pi R - S}{\pi(1-\phi)R - V_o} > 0$.

Furthermore, the maximal payoff the senior lender can generate with a mezzanine investor in place exceeds all payoffs she can generate with two-party contracts if and only if:

1. $2V_o + \omega(1+r) < \pi(1-\phi)R + \pi(1-\phi)R$, *and*

2. θ *is sufficiently close to* 1.

Parameters may be such that $2V_o + \omega(1+r) < \pi(1-\phi)R + \pi(1-\phi)R$ but, at the same time, that $\pi - (1-\pi)\frac{\phi\pi R - S}{\pi(1-\phi)R - V_o} < 0$ say because S is low. In that case, the optimal two-party contract does not feature any chance of termination and yet, bringing in a back-up

[12]The paper is available for download at (blinded)

operator raises the lender's surplus because the threat of termination becomes cheaper to invoke. Therefore, mezzanine contracts may create value even when the two-party contract does not contain any chance of early termination.

6.2 Relative skills of mezzanine investors

When $\theta < 1$ operator 1 is uniformly more productive than operator 2 so that any optimal contract features operator 1 at the helm. If $\theta > 1$ on the other hand, the property generates more rents when operator 2 is in charge. When the senior lender gets to design the contract, it is immediate that she would then choose to have operator 2 run the property while operator 1 serves the role of mezzanine investor.

The situation becomes more complicated when operator 1 has control of the investment opportunity and gets to design the contract. When $\theta > 1$ it may be optimal to delegate operations to operator 2 from the start and compensate them sufficiently to induce their participation. However, contracts of this type where operator 2's participation constraint binds – i.e. contracts that give operator 2 a surplus exactly equal to $2V_0 + \omega(1 + r)$ – may not be compatible with the senior lender's participation constraint. This potentially forces operator 1 to give up more surplus to operator 2 when they relinquish operational control. But eventually – as θ gets large – operator 1 must be better off with operator 2 at the helm. Indeed, the upper bound on the surplus operator 2 needs to be given is $2(1 - \phi\pi\theta R)$ so that operator 1 can secure at least a surplus of

$$2\pi\theta R - 2(1 - \phi\pi\theta R) - (1 + r).$$

As θ gets large enough, that payoff must dominate what any contract with operator 1 initially in charge can generate.

In summary, when operator 2 is significantly more productive than operator 1, it is socially and privately optimal to put operator 2 in charge of the property at first, while operator 1 plays the role of the mezzanine investor.

6.3 Different discount rates

The case where all agents discount cash flows at the same rate has the simplifying feature that consumption by operator 1 can be efficiently postponed to date 2. This simplifies the analysis of the optimal contract because we can assume without loss of generality that the operator retains no rents ($q(R) = R$) in the first period. This occurs because with equal discount rates there is no benefit to early consumption. As De Marzo and Fishman (2007, see section 3.3) discuss in the context of a model that nests our environment without back-up operators, the main consequence of making the operator more impatient than the senior lender is that the two-party contract may now call for positive consumption in date 1, i.e. $q(R) < R$. The contract may even call for a positive payment to the operator even when the property yields no rent. But our key results, that ex-post inefficient termination of the property may take place in two-party contracts and, in turn, that including operator 2 as a a back-up operator is optimal as long as operator 2 is sufficiently skillled, are unchanged. In fact and if anything, since the contract now has to fund early consumption by operator 1, this increases the odds that continuation utilities as of date 2 will be low enough that putting positive probability on termination becomes necessary.[13]

6.4 Testable implications

Including operator 2 in the contract is optimal when $V^M - V^N > c$. Particularly interesting is the case where $V^M - c > 0 \geq V^N$ i.e. the case where owning the property only becomes a positive NPV investment once a back-up operator in place. To see that this can happen, simply assume that ϕ is low enough (moral hazard is sufficiently severe) that

$$\pi\phi R + \pi\phi R < (1 - \omega)(1 - r).$$

[13]The details of the argument are identical to those provided by De Marzo and Fishman (2007) in section 3.3.

In that case, meeting the senior lender's participation constraint is not possible without termination being a positive probability. With S sufficiently close to $\pi\phi R$, the project may be viable, but if S is sufficiently low (recall that we allow it to be negative) the project is not. As long as θ is sufficiently close to 1 but S is sufficiently low, the project is viable with and only with a back-up operator in place.

These considerations yield several testable implications for the type of projects for which mezzanine contracts are more likely to create value. First, as long as the cost c of including a back-up operator in the capital structure is partially fixed, the property has to be sufficiently large (y must be sufficiently high) to justify bearing the fixed cost. All else equal therefore larger projects should be more likely to feature a financial structure that includes mezzanine lenders as we model them.

Second and quite intuitively, operators and investments subject to more moral hazard frictions (a lower ϕ) benefit more from the presence of back-up operators. Indeed, V^N increases as ϕ does while the threshold $\pi(1-\phi)R + \pi(1-\phi)R$ past which back-up operators do not create value declines. Similarly, the benefit of including operator 2 in the contract increases if his ϕ is higher. Simply put, low reputation operators stand to benefit the most from the presence of mezzanine investors, especially when the reputation of those mezzanine investors is high.

The effect of π, the probability of success, on the value of mezzanine contracts is ambiguous. On the one hand, a higher π implies higher surplus so that it becomes more likely all else equal that the benefits of mezzanine contracts will offset the fixed part of their cost. This is a version of the size effect we discussed above. A higher π also makes it more likely that a termination clause will be part of the original contract, since it is cheaper to invoke. But, on the other hand, termination occurs less often so that the expected gain associated with including back-up operators is small.

The easiest way to see that the effect of π is ambiguous is to note that when $\pi = 0$, $V^N = V^M = 0$ since the property has no value, while when $\pi = 1$ failure never happens so that the senior lender's and operator 1's payoffs are independent of whether or not operator 2 is part of the contract. Mezzanine investors cannot create value when there is no uncertainty. When the property is risky however ($\pi \in (0,1)$) termination can be part of optimal two-party contracts and mezzanine becomes potentially valuable.

7 Conclusion

The presence of mezzanine investors in the capital structure can transform negative-NPV projects into positive-NPV projects by making it cheaper for senior lenders to provide incentives to property owners. With approval from the senior lender, skilled mezzanine investors have the option to take over and operate properties without experiencing the steep costs associated with the mortgage foreclosure process.

Mezzanine investors, in our theory, optimally provide a blend of capital and back-up operating skills. Even though senior lenders could fund the property alone, the optimal arrangement calls for a contribution to the capital structure by mezzanine investors to make it easier to provide them with the necessary incentives if and when they take over. This provide a natural explanation for the fact that mezzanine investors often provide a small sliver of capital. That sliver may amount to a small fraction of the needed fund but provides essential skin-in-the-game for mezzanine investors.

In our model, mezzanine lenders receive their highest payoff when the property underperforms early but remains viable as an ongoing concern. Cases where backup managers serve the purpose described in this paper should be thought of as properties that disappoint early but remain profitable. When properties experience shocks so severe that continuation is no longer possible, all stake-holders – including mezzanine investors – experience low returns. As more data on mezzanine returns become available, it will become possible to test this and our theory's other predictions.

References

Allen, F., and D. Gale, 1988. Optimal Security Design, Review of Financial Studies, 1, 229-263.

Baum, A., and D. Hartzell 2012. Global Property Investment, Wiley-Blackwell, Oxford, UK.

Berman, A., 2013. Mezzanine Debt and Preferred Equity in Real Estate, in Alternative Investments: Instruments, Performance, Benchmarks, and Strategies, H. Kent Baker and Greg Filbeck Editors, John Wiley & Sons, Hoboken, NJ.

Biais, B., G. Plantin, and J. Rochet, 2007. Dynamic security design: Convergence to continuous time and asset pricing implications, Review of Economic Studies, 74, 345-390.

Black, L., Krainer, J, and J. Nichols, 2017. From Origination to Renegotiation: A Comparison of Portfolio and Securitized Commercial Real Estate Loans, Journal of Real Estate Finance and Economics, 55, 1-31.

Casamatta, C., 2003. Financing and Advising: Optimal Financial Contracts with Venture Capitalists, Journal of Finance, 58, 2059-2086.

De Marzo P., and M. Fishman, 2007. Optimal Long-term Financial Contracting, Review of Financial Studies, 20, 2079-2128.

De Marzo P., and Y. Sannikov, 2006. Optimal Security Design and Dynamic Capital Structure in a Continuous Time Agency Model, Journal of Finance, 61, 2681-2724.

Gertler, D., Miller, N., Clayton, J., and P. Eichholtz, 2007. Commercial Real Estate Analysis and Investment, Cengage Learning, Mason, OH.

Giliberto, M., 2018. At Long Last, the CRE Market Has a Mezzanine-Loan Index. CRE Finance World, Winter 2018.

Hart O., and J. Moore, 1994. A Theory of Debt Based on the Inalienability of Human Capital, Quarterly Journal of Economics, 109, 841-879.

Holmstrom, B., and J. Tirole, 1997. Financial Intermediation, Loanable funds and the Real Sector, Quarterly Journal of Economics, 62, 663-691.

Pagliari, J. L., 2017. High-Yield Lending: It's Good Until It's Not, Journal of Portfolio Management, 43, 138-161.

Repullo, R., and J. Suarez, 2004.Venture Capital Finance: A Security Design Approach, Review of Finance, 8, 75-108.

Rubock, D., 2007. US CMBS and CRE CDO: Moody's Approach to Rating Commercial Real Estate Mezzanine Loans, Moody's Investors Service Report, March 29.

Stein, J. 1997. Subordinate Mortgage Financing: The Perils to the Senior Lender, Real Estate Review, Fall issue, 3-9.