A Theory of Mandatory Convertibles: Distinct Features for Large Repeated Financing

Susheng Wang

October 2016

Abstract: In recent years, mandatory convertibles are becoming a popular means of raising capital, especially for large projects. This paper is the first theoretical paper to investigate mandatory convertibles using the incomplete-contract approach. We show that mandatory convertibles can be an efficient instrument in sequential financing. Mandatory convertibles have some distinct features compared to other convertibles, such as mandatory conversion, a high dividend rate, and capped capital appreciation. We show in theory that these features are designed to achieve efficiency.

Keywords: Mandatory Convertibles, Sequential Financing, Mandatory Conversion, Capped Capital Appreciation

JEL Classification: G32, G31

1 We gratefully acknowledge the helpful comments and suggestions from two referees.

2 Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong. Email: s.wang@ust.hk. Phone: (852) 2358-7600.
1. Introduction

Convertibles are attractive to firms with many real investment options (Mayers, 1998, 2000). Such firms need to raise capital repeatedly as opportunities arrive. Convertibles with their debt features at early stages attract investors and with their convertibility in later stages allow firms to convert debt into equity so as to go back to the capital market for further funding. Among convertibles, mandatory convertibles (MCs) are the most equity-like convertibles and they represent a major recent innovation in securities. All MCs have three unique characteristics: (a) mandatory conversion to equity at a specified date, usually in three to five years, (b) a high dividend rate, much higher than the dividend rate on the underlying common stock, and (c) capped capital appreciation, usually on the upside. We will show that these features serve the purpose of achieving efficiency in sequential financing.

Although convertibles appear early in 1800s (Calamos, 2003), MCs appear only recently; they appeared in 1988. The risk-return profile of a MC is very different from standard convertibles (Calamos, 2003, p.28). Unlike standard convertibles, MCs are typically issued by larger issuers. Starting from small beginnings in 1988, MCs have quickly become a popular means of raising capital in recent years (Dinsmore et al., 2016). For example, they accounted for one quarter of the $20 billion convertible market in 1996 (Arzac, 1997) but they amounted to $38.1 billion in 2007 (Chemmanur et al., 2014). As pointed out by Ertugrul et al. (2004), over a short period of about 10 to 15 years, MCs have become a major source of financing for large firms like Citicorp, Anthem, Kaiser, Aluminum, Texas Instruments and Home Depot. In particular, in venture capital financing, over 90% of investment is done through convertibles and these convertibles are mandatorily convertible at IPO (Sahlman, 1990; Hellmann, 2006). Dinsmore et al. (2016) observe that the US convertible market is basically made up by three types of convertibles: standard convertibles, mandatory convertibles, and convertible preferred stock, with market shares of 74%, 17%, and 9% respectively. That is, MCs are now the second most popular type of convertibles. Dinsmore et al. (2016) also observe that the market share of MCs is growing. However, issuance of MCs as a share of the convertible market fluctuates a lot over business cycles, dipping dramatically in the financial crisis and rising robustly during certain years of the 2000s.

However, studies on MCs are very rare. Chemmanur et al. (2014) show that MCs provide a solution to the financial restructuring problem faced by highly leveraged companies; MCs also enable growth companies to signal confidence about their future, particularly by including features such as guaranteed appreciation. Their model is an extension of Stein (1992) by allowing MCs. While Chemmanur et al. (2014) are the first theoretical work on MCs using the asymmetric-information approach, our work is the first theoretical work on MCs using the incomplete-contract approach. Besides, we allow sequential financing as an additional mecha-
nism, which appeals to the real options theory, while they do not. We argue that one key advantage of MCs is that it can be an efficient financing instrument in sequential financing. These two different approaches lead to very different solutions. For example, in Chemmanur et al. (2014), only certain types of firms choose MCs and all their solutions are inefficient; while in our model, any firm can issue MCs and our solutions are efficient under sensible conditions.

Wang (2009) and we both use the incomplete-contract approach to study financial assets. While Wang (2009) studies callable convertibles, we study MCs. There are four key differences between the two. First, the sets of admissible contracts in the optimization problems are totally different in the two papers; in fact, these two sets are mutually exclusive — MCs are not allowed in Wang (2009) and callable convertibles are not allowed in our paper. Second, the conversion ratio in Wang (2009) is a pre-determined fixed constant, while the conversion ratio in our model is a function of the spot stock price (a floating conversion ratio). This floating feature is the key feature in MCs. In fact, the design of an efficient MC focuses on the design of a proper floating conversion ratio. Third, both the firm and investors have ex-post options in Wang (2009), while none of them has an ex-post option in our model. Fourth, our paper is the first agency theory on MCs.

Convertibles have been observed to go hand in hand with sequential financing in reality (Mayers, 1998, 2000). In particular, MCs have often been involved in large repeated financing. We suspect that convertibles with their special features may be able to resolve many of the associated problems in sequential financing. Based on the incomplete-contract approach, proposed by Grossman & Hart (1986), this paper theoretically investigates whether or not MCs can effectively deal with the problems inherent to sequential financing. Indeed, we find that sequential financing using MCs can be efficient.

Three key studies relate to our work. Mayers (1998, 2000) proposes the view that a convertible is a cost-saving instrument in sequential financing. The costs include issue costs and agency costs. This view is particularly relevant to companies with many real investment options. Chemmanur et al. (2014) find separating equilibria in which different types of firms choose different types of financial instruments. They identify certain types of companies that choose MCs, while others choose debt, equity or other convertibles. However, theoretical studies that relate convertibles to sequential financing are rare. Cornelli & Yoshia (2003) argue that, with sequential financing, the manager has the incentive to do window dressing in order to attract further investments. With convertibles, if the manager overstates the company’s value, a holder can exercise conversion. By this, window dressing is prevented. Wang’s (2008) theoretical study emphasizes the call feature and call protection in callable convertibles (MCs are not callable convertibles by definition). In contrast, we here provide a theory that focuses on the three distinct features of MCs. In particular, the design of a floating conversion ratio is crucial for a resolution of incentive problems in our model.
There are also other related studies. Baker & Wurgler (2002) suggest that managers act like fund managers who manage a firm’s financial assets like an investment portfolio. In our model, conversion to equity is not based on the need for the investors to become the owners of the firm; instead, it is based on a balance between investors’ returns and the manager’s incentives. As shown in (16), the conversion ratio depends on the firm’s performance ex post. This is consistent with Baker & Wurgler’s (2002) empirical findings, although our interpretation of their empirical findings is completely different from theirs. Lewis, Rogalski & Seward (1998) observe the importance of call features in their empirical study. They find that issuers even adjust call features over business cycles. Their findings are consistent with our theoretical findings. Our floating conversion ratio is indeed dependent on market conditions. Wang & Zhou (2004) discuss equity sharing in sequential financing, while we discuss MCs in sequential financing. They show that equity sharing is inefficient and it is approximately efficient for firms with low marginal costs, while we find that MCs can be efficient, especially when marginal costs are high (Propositions 2 and 3). We argue that, by raising funds in stages, convertibles give investors an opportunity to see how the project goes, to keep pressure on the manager, and to strike a balance in risk sharing. In particular, we show that the three observed common characteristics in MCs can serve the purpose of achieving efficiency in sequential financing.

This paper proceeds as follows. In Section 2, we define a model of corporate finance in an environment with uncertainty, moral hazards and sequential financing. To carry out sequential financing, we allow MCs as vehicles of investment. In Section 3, we solve for equilibrium MCs. Efficient MCs are found and these MCs happen to exhibit the three prominent features of MCs in reality. Section 4 discusses two most popular types of MCs in reality. Some real-world examples are also presented. Finally, Section 5 concludes the paper with a few remarks. All the proofs are in the Appendix.

2. The Model

2.1. The Project

Consider a firm that relies on investors for capital. The firm is run by a manager (she) on behalf of the existing shareholders by maximizing the firm’s expected profit. The investors’ objective is also to maximize their own expected profit. The project lasts two periods. The manager provides an investment $x$, called effort, and she needs to raise the necessary capital $K$ from the market. After a contract between the manager and investors is signed, the manager provides her effort $x$ at cost $c(x)$. This cost is personal and is paid by herself. The effort is applied throughout the two periods and the cost function is positive, increasing and convex. The funding $K$ is raised in sequence with an initial installment $k_1$ in the first period and a
planned second installment \(k_2\) in the second period, where \(k_1 + k_2 = K\). Given effort \(x\) from
the manager and investments \(k_1\) and \(k_2\) from the investors, the project generates value \(y_p\) at
the end of the second period. The firm’s output (earnings) \(y\) is the sum of the output \(y_p\) from
this new project plus the value \(y_0\) of its existing projects: \(y = y_p + y_0\). After the uncertainty is
realized at the end of the first period, the manager considers raising a second installment \(k_2\).
The manager may need to bargain/renegotiate with the investors in the market again for the
second installment.

The project output \(\tilde{y}_p\) is uncertain as of the beginning of the first period. Let \(F(y, x, k_1)\)
and \(f(y, x, k_1)\) be respectively the distribution and density functions of output \(\tilde{y}\) conditional on
investments \((x, k_1)\). As we focus on the financing problem of the new project, we will simply
assume that \(y_0\) is a deterministic constant so that \(F(y, x, k_1)\) captures the randomness of \(\tilde{y}_p\)
only.\(^3\)

Production takes two periods, and both the manager and investors are indispensable. If
the project is liquidated before completion, it is liquidated for a fraction \(\theta k_1\) of the initial
capital investment, where \(\theta \in [0, 1]\). The manager and investors share the final profit on the
basis of the existing contract at the end of the project.

### 2.2. Timing of Events

Some information becomes known at the end of the first period: the realization of output
is publicly revealed at the end of the first period, and the manager’s effort is observable at time
t = 1 but not verifiable. The amounts \(k_1\) and \(k_2\) are verifiable, but the decision on whether or
not to continue the second investment at \(t = 1\) is taken ex post and is not contractible ex ante.
This implies that the decision on the option will be conditional on the observation of the man-
ger’s input \(x\) and on the knowledge of the random shock. As some information becomes
available ex post, the manager and the existing contract holders are allowed to renegotiate the
contract.\(^4\)

The timing of the events is illustrated in Figure 1. At \(t = 0\), the manager offers a contract
to investors. If the contract is accepted, the investors provide \(k_1\) and the manager applies

\(^3\) Given the distribution function \(F\) of the project output \(\tilde{y}_p\), the distribution function \(F\) of output \(\tilde{y}\) is
\(F(y, x, k_1) = P(\tilde{y} \leq y) = P(\tilde{y}_p + y_0 \leq y) = P(\tilde{y}_p \leq y - y_0) = F_p(y - y_0, x, k_1)\), where \(P()\) is the probability of
an event. If \(y_0\) is random, all the results hold except that \(F\) is replaced by the joint distribution function of \((\tilde{y}_p, y_0)\). We
can also allow the randomness of \(\tilde{y}_0\) to be resolved earlier or later than \(\tilde{y}_p\).

\(^4\) We follow the incomplete-contract approach proposed by Grossman & Hart (1986). Refer to their paper on
the concepts of “a variable being ex ante nonverifiable but ex post observable”, “an ex post decision being ex ante
uncontractible”, “ex post renegotiation”, and “an efficient bargaining outcome”.

5
effort $x$ and incurs personal cost $c(x)$. At $t = 1$, the uncertainty is resolved and the two parties consider the options to quit or to continue. Depending on the nature of the contract, the two parties may renegotiate or decide on their options at $t = 1$ or $t = 2$. At $t = 2$, the project is finished and the two parties divide the output based on the existing contract.

<table>
<thead>
<tr>
<th>Contracting</th>
<th>Information Revelation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effort: $x$</td>
<td>Information Revelation</td>
</tr>
<tr>
<td>Investment: $k_1$</td>
<td>Investment: $k_2$</td>
</tr>
<tr>
<td>Ex ante</td>
<td>Ex post</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>End</td>
</tr>
</tbody>
</table>

Output: $y$

Figure 1. The Timing of Events

### 2.3. Financial Instruments

Admissible contracts in our model are mandatory convertibles. A mandatory convertible (MC) pays a guaranteed rate of return $r$ and mandatorily converts the investment into equity at a floating conversion ratio $\tau(p)$ at maturity based on the spot stock price $p$ of the firm. Given the spot stock price $p$, the conversion ratio $\tau(p) \in [0, 1]$ is the proportion of the firm’s equity that the MC holder gets after conversion. Our MCs include many types of MCs in reality, which distinguish themselves by their potential in capital appreciation, or more specifically, by their design of the floating conversion ratio $\tau(p)$.

In our two-period model, in particular, the firm issues a MC at $t = 0$ with principle $k_1$ that matures at $t = 1$, and the firm issues another MC at $t = 1$ (if the project is doing well) with principle $k_2$ that matures at $t = 2$. Investors invest in the form of MCs that pay a guaranteed dividend rate $r$ and mandatorily convert the investment into equity at a floating conversion ratio $\tau(p)$ at the end of each period (at $t = 1$ and $t = 2$).

The MCs are tradable financial assets. During the course of investment and production, the assets may change hands at any time. When it is traded, all the rights defined by the assets go with the assets to the new holder.

---

5 Adding another random shock in the second period will add a lot of technical complication. Such complication is unnecessary for the purpose of this paper.
2.4. Assumptions

More initial investment $k_1$ from investors and more effort $x$ from the manager should increase the chance of producing a high output. Since the probability of producing an output higher than $y$ is $1 - F(y, x, k_1)$, the distribution function $F(y, x, k_1)$ should be decreasing in $(x, k_1)$, in the sense of first-order stochastic dominance (FOSD).

**Assumption 1** (FOSD). $F_x(y, x, k_1) < 0$ and $F_{k_1}(y, x, k_1) < 0$ for all $(y, x, k_1)$, where $F_x$ and $F_{k_1}$ are the partial derivatives of $F$ with respect to $x$ and $k_1$, respectively.

**Assumption 2.** The support of $F(y, x, k_1)$ is independent of the investments $(x, k_1)$.

These two assumptions are standard in almost all agency models in the literature. The first assumption makes sense, meaning that the investments $x$ and $k_1$ improve the chance of a high output. The second assumption is unnecessary but it substantially reduces the complexity of agency theory.\(^6\) For example, suppose the output process has the following form:

$$\tilde{y} = h(x, k_1) + \tilde{e},$$  \hspace{1cm} (1)

where $h(x, k_1)$ is an increasing function and $\tilde{e}$ is a random variable with zero mean. Denote $g(\tilde{e})$ and $G(\tilde{e})$ as respectively the density and distribution functions of $\tilde{e}$. Then,

$$F(y, x, k_1) = G[y - h(x, k_1)], \quad f(y, x, k_1) = g[y - h(x, k_1)].$$

We have $F_x = -h_x g < 0$ and $F_{k_1} = -h_{k_1} g < 0$. If the support of $G$ is $(-\infty, \infty)$, the support of $F$ is also $(-\infty, \infty)$, which is independent of $(x, k_1)$.

We further assume risk neutrality in income and no discounting of future income for both parties.

2.5. Model Setup

The total required investment $K$ is not a choice variable; it is a given number, necessary for a planned business expansion. When allocating the two installments $k_1$ and $k_2$, with $k_1 + k_2 = K$, the two parties will consider the benefit from the resolution of uncertainty by investing ex post and the two parties may liquidate the project ex post and not raise the second installment. But this allocation may affect the manager’s incentives and hence the output. An

\(^6\) We will have Riemann-Stieltjes integrals of the form $\int_{y_1}^{y_2} v(y) dF(y, x, k_1)$, where $v(y)$ is a real-valued function and $y$ is the variable of integration. A sufficient condition for this integral to be well-defined is: $v(y)$ is continuous and bounded on interval $[y_1, \infty)$. We can also allow $v(y)$ to be discontinuous on a measure zero set if we treat $\int_{y_1}^{y_2} v(y) dF(y, x, k_1)$ as a Lebesgue-Stieltjes integral (Horst, 1984).
ex-post investment decision imposes a risk on the manager (due to the investors’ options to quit or to renegotiate), which may lead to lower effort. Therefore, the allocation of investment needs to be arranged properly in order to balance the manager’s incentive to expend effort and the investors’ benefit from late investment.

We do not restrict ourselves to the principal-agent setup, in which one of the parties is given the full bargaining power ex ante. Instead, the two parties in our model negotiate and bargain over the terms of a contract ex ante and possibly ex post. The possibility of renegotiation implies that the two parties will negotiate an agreement that ensures social welfare maximization ex ante as well as ex post, subject to incentive compatibility (IC) and individual rationality (IR) conditions. An agreement is a financial instrument (i.e., a MC) that defines their trade, rights and bargaining positions.

Specifically, let $\Pi_I$ and $\Pi_M$ be respectively the ex-ante payoffs to the investors and the manager. The investors’ initial investment $k_1$ is contractible, but the decision on whether or not to raise the second installment is not contractible ex ante. As usual, the manager’s investment $x$, called effort, is unverifiable. The contract is an outcome of negotiation. With the assumption of an efficient bargaining outcome, the ex-ante problem is

$$\max_{x, k_1, r, \tau(\cdot)} \Pi_M + \Pi_I$$

subject to

$$\frac{\partial \Pi_M}{\partial x} = 0, \quad (\text{IC condition})$$

$$\Pi_I > 0, \quad (\text{IR condition})$$

with choices $x, k_1, r$ and $\tau(\cdot)$, the manager’s IC condition, and the representative investor’s IR condition. Since investors are homogenous in our model, we can think that there is only one representative investor (he) and he buys MCs in both periods. Similarly, we can assume that the firm issues one MC at $t = 0$ with principle $sk_1$ or alternatively issues $k_1$ units of MCs at $t = 0$ with principle $\$1$ each.

Our tasks are (a) to identify an efficient mechanism (i.e., a MC) that is popular in practice and (b) to identify some special features in the financial instrument that are necessary for efficiency and also popular in practice.

**Remark 1.** In our model, we refer to one manager and many investors. The group of potential investors represents the market. The manager stays with the company from the beginning to the end, but MC holders can change any time. A MC is tradable. An investor who buys a MC at $t = 0$ may sell his MC to another investor later. Also, an investor who invests at $t = 0$ by purchasing a MC may or may not invest at $t = 1$; a second investor may choose to invest at $t = 1$.

**Remark 2.** Although we limit ourselves to MCs, since we can find an efficient solution, a larger set of admissible contracts is unnecessary.
3. Mandatory Convertibles

In this section, we investigate how sequential financing can be efficiently carried out by MCs. All the proofs are in the Appendix.

3.1. The First-Best Problem: the Benchmark

As a benchmark, we first consider the first-best problem, which is a sequential financing problem without agency problems. Specifically, the manager’s effort $x$ is contractible and the project continues at $t = 1$ with a second installment if and only if it is ex-post efficient to do so. Hence, at time $t = 1$, if and only if

$$y_p - k_2 \geq \theta k_1 \quad \text{or} \quad y \geq y_1 \equiv y_0 + \theta k_1 + k_2, \quad (3)$$

the second installment is provided and the project continues; otherwise, the project is liquidated with liquidation value $\theta k_1$ at $t = 1$ (Figure 2).

![Figure 2. The First-Best Problem](image)

Hence, with probability $F(y_1, x, k_1)$ that the project is liquidated at $t = 1$, the first-best problem can be written as:

$$V^* \equiv \max_{x, k_1 \geq 0} (y_0 + \theta k_1) F(y_1, x, k_1) + \int_{y_1}^{\infty} (y - k_2) dF(y, x, k_1) - k_1 - c(x),$$

or

$$V^* = \max_{x, k_1 \geq 0} y_1 F(y_1, x, k_1) + \int_{y_1}^{\infty} y dF(y, x, k_1) - K - c(x), \quad (4)$$

where $V^*$ is social welfare at the first best. In all integrals in this paper including the above two, the variable of integration is always $y$. The equations determining the solution are listed in the following proposition.

**Lemma 1** (First Best). The first-best solution $(x^*, k_1^*)$ is determined by two equations:

$$(1 - \theta) F(y_1, x, k_1) \int_{y_1}^{\infty} F_x (y, x, k_1) dy = 0, \quad (5)$$

$$\int_{y_1}^{\infty} F_x (y, x, k_1) dy + c'(x) = 0. \quad (6)$$
As shown in the proof of Lemma 1, (5) means that the marginal impact of initial investment $k_1$ on social welfare is zero, and (6) means that the marginal impact of the manager’s effort $x$ on social welfare is zero. The second-order conditions are discussed later in Remark 3.

### 3.2. The Second-Best Problem: Convertible Financing

In this section, we solve problem (2) by assuming that the manager issues MCs to carry out sequential financing.

In practice, the conversion ratio of a MC typically has the form in Figure 3. Two caps at the two ends are a key feature. When the stock price is sufficiently low, a maximum conversion ratio is guaranteed; when the stock price is sufficiently high, a minimum conversion ratio is offered; and when the stock price is in the middle range, the conversion ratio is floating according to the spot stock price. The difference between $p_2$ and $p_3$ may be considered as the normal range of capital appreciation. This difference is typically small in practice, about 15% to 20%.

Formally, we denote a MC as $(r, p_2, p_3, \tau(\cdot))$ with dividend rate $r$, trigger prices $p_2$ and $p_3$, and conversion ratio $\tau(p) \in [0, 1]$. As shown in (7) and Figure 3, it pays dividends at a guaranteed rate $r$ at the end of each period (at $t = 1$ and $t = 2$), mandatorily converts the investment into equity at the end of each period using a floating conversion ratio $\tau(p)$, and controls capital appreciation using two trigger stock prices $p_2$ and $p_3$. That is,

$$
\tau(p) = \begin{cases} 
\tau_2 & \text{if } p \leq p_2, \\
\tau_m(p) & \text{if } p_2 < p < p_3, \\
\tau_3 & \text{if } p \geq p_3,
\end{cases}
$$

(7)

where $\tau_1$ and $\tau_2$ are two constants, serving as two caps, and $\tau_m(p) \in [0, 1]$ is a function.

---

7 We are here restricting admissible MCs. However, since the optimal MC that we manage to find is the first-best, our restriction does not reduce the efficiency of our solution. That is, our solution is as efficient as a solution when there is no restriction on MCs, implying that our solution is an optimal solution if there is no restriction on MCs. Our restriction ensures that our solution has the key features of real-world MCs.  

8 MCs typically allow holders to convert before maturity. However, the conversion ratio that applies to early conversion is the lowest possible conversion ratio. Hence, holders typically do not convert early (definitely not so in our model). We will hence simply assume no early conversion.
The ex post stock price or firm value is output minus interest. Hence, the stock price is
\[ p = \begin{cases} p_1 & \text{if the project is liquidated,} \\ y - rK & \text{if the project continues,} \end{cases} \]
where \( p_1 \equiv y_0 + \theta k_1 - r k_1 \) is the stock price when the project fails, while \( y_1 \equiv y_0 + \theta k_1 + k_2 \) is the output level below which the project is liquidated. \( p_1 \) is the worst outcome from an investor’s point of view, and is the so-called initial price or liquidation preference in practice. Let \( y_2 \) and \( y_3 \) be the corresponding output levels at which the stock prices are the trigger prices \( p_2 \) and \( p_3 \), respectively, where \( y_3 \geq y_2 \geq y_1 \). That is, \( y_2 \) and \( y_3 \) are the trigger output levels, and \( p_2 = y_2 - rK, p_3 = y_3 - rK \).

<table>
<thead>
<tr>
<th>trigger output</th>
<th>trigger price</th>
<th>event</th>
</tr>
</thead>
<tbody>
<tr>
<td>( y_1 = y_0 + \theta k_1 + k_2 )</td>
<td>( p_1 = y_0 + (\theta - r)k_1 )</td>
<td>the project is liquidated</td>
</tr>
<tr>
<td>( y_2 = p_2 + rK )</td>
<td>( p_2 )</td>
<td>the first trigger price ( p_2 )</td>
</tr>
<tr>
<td>( y_3 = p_3 + rK )</td>
<td>( p_3 )</td>
<td>the second trigger price ( p_3 )</td>
</tr>
</tbody>
</table>

where the trigger prices \( p_2 \) and \( p_3 \) are defined in a MC, with \( p_3 > p_2 > p_1 \). If the investments are the first-best investments \( k_1^* \) and \( k_2^* = K - k_1^* \), we denote \( y_1 \) and \( p_1 \) as \( y_1^* \) and \( p_1^* \), respectively.

Let us now derive the two parties’ decisions. By ex-post renegotiation, the project will be liquidated if and only if \( y < y_1 \). That is, \( y_1 \) is the threshold for default. If \( y < y_1 \), the project is not worth continuing and is hence liquidated at \( t = 1 \). If \( y \geq y_1 \), the project continues to the next period and a second MC is issued for investment \( k_2 \) (Figure 2). Hence, when \( y < y_1 \), the investors receive \( r k_1 \) and \( \tau(p_1)\theta k_1 \) at \( t = 1 \) and the project is liquidated; when \( y \geq y_1 \), the investors receive \( r k_1 \) at \( t = 1 \) and \( r k_2 \) and \( \tau(p)(y - rK) \) at \( t = 2 \), where \( y - rK \) is the firm’s profit. On the other hand, when \( y < y_1 \), the firm receives \( [1 - \tau(p_1)]\theta k_1 \) and pays \( r k_1 \) at \( t = 1 \) and the project is liquidated; when \( y \geq y_1 \), the firm receives \( [1 - \tau(p)](y - rK) \) at \( t = 2 \). Taking
into account the manager’s cost $c(x)$ and the investors’ investments $k_1$ at $t = 0$ and $k_2$ at $t = 1$ if $y \geq y_1$, the ex-ante payoffs for the investors and the manager are respectively:\footnote{Here $\tau(p)$ is the total equity holding for all the investors who are holders of the two MCs issued at $t = 0$ and $t = 1$ (or equivalently, $\tau(p)$ is the representative investor’s equity holding). Hence, $\Pi_i$ is the total payoff to the new shareholders and $\Pi_{m}$ is the total payoff to the original shareholders. Although we do not explicitly mention dividends on common stock, the firm may pay them. No matter whether such dividends are paid or not, the investors’ total income from the equity share should be $\tau(p)(y - rK)$. Hence, a reader can simply assume no dividends on common stock. However, if there is a random stock in the second period, the timing of a dividend payment is relevant and it reflects risk sharing, which would force us to specify dividends on common stock explicitly.}

$$\Pi_i = [\tau(p_1)\theta k_1 + r k_1]F(y_1, x, k_1) + \int_{y_1}^{\infty} [\tau(y - rK)(y - rK) + rK - k_2]dF(y, x, k_1) - k_1$$

$$= [\tau(p_1)\theta k_1 + (1 - r)k_2]F(y_1, x, k_1) + \int_{y_1}^{\infty} \tau(y - rK)(y - rK)dF(y, x, k_1) - (1 - r)K, \tag{8}$$

and

$$\Pi_{m} = \{y_0 + [1 - \tau(p_1)\theta k_1 - r k_1]F(y_1, x, k_1)$$

$$+ \int_{y_1}^{\infty} [1 - \tau(y - rK)](y - rK)dF(y, x, k_1) - c(x). \tag{9}$$

Social welfare is

$$V \equiv \Pi_i + \Pi_{m} = y_1F(y_1, x, k_1) + \int_{y_1}^{\infty} ydF(y, x, k_1) - K - c(x). \tag{10}$$

We assume that social welfare is positive at the first best, i.e., the project is viable ex ante.

The second-best problem is problem (2) with choices $x$, $k_1$, $r$ and $\tau(p)$. Our question is: why do firms have the form of a conversion ratio in Figure 3? Our answer to this question is to show that such a conversion ratio can achieve efficiency in sequential financing.

\textbf{Proposition 1.} MCs can be efficient in sequential financing. Specifically, if the stock price $p_1^* = y_0 + (\theta - r)k_1^*$ at project failure is positive, where $k_1^*$ is the first-best initial investment, then there is an efficient (first-best) MC $(r, p_2, \tau(\cdot))$ with dividend rate $r \in (0, 1)$, trigger prices $p_2 > p_2 \geq p_1^*$ and a floating conversion ratio $\tau(\cdot)$ of the form in (7).

We have several remarks about Proposition 1. First, the efficient MC in Proposition 1 happens to have the three known characteristics: a high dividend rate, mandatory conversion, and capped capital appreciation. The proof of Proposition 1 requires the dividend rate to be sufficiently high.

Second, the optimal solution in Proposition 1 is not unique; there are many efficient MCs that induce the first-best investments $(x^*, k_1^*)$ determined by conditions (5) and (6).
Third, in all our propositions, Propositions 1-3, a high dividend rate is required in the proofs, although it is not a necessary condition for efficiency of our MCs. In practice, a MC typically lasts three to five years before it is mandatorily converted into equity at maturity. An annual dividend rate for a MC typically ranges from 6% to 15%. Hence, our $r$ can range from 0.18 to 0.75 (one period equals three to five years).

Fourth, in most cases, the firm is well established and the liquidation preference in a MC is typically stated as 100%. But, due to transaction costs, not all investors can actually get 100% back if the project is liquidated. Hence, we may argue that $\theta$ represents transaction costs of raising capital. If so, this $\theta$ should be less than 1 and, for large firms, close to 1.

Fifth, in our solution, $\tau_2$ can be an arbitrary number less and close to 1, and $\tau_m(y)$ can be an arbitrary decreasing function, but $\tau_3$ is defined by (16) depending on the choice $\tau_2$ and $\tau_m$. In practice, $\tau_2$ is typically 1.

Finally, in the proof of Proposition 1, $y_2$ needs to be sufficiently close to $y_3$ for efficiency. In many real-world MCs, the gap between $y_2$ and $y_3$ is typically small, normally being about 15% to 20%. There are two considerations in the design of a MC. One consideration is the manager’s incentives: the IC condition. As shown by the proof of Proposition 1, this can be properly controlled by a choice of the cap $\tau_3$. A low $\tau_3$ is intended to satisfy the manager. The other consideration is the investors’ returns: the IR condition. This can be partially controlled by $\tau_2$. A high $\tau_2$ is intended to satisfy the investors.

**Remark 3.** Conditions (5) and (6) are the first-order conditions (FOCs) of (4). As we know, we also need the second-order conditions (SOCs) to ensure that the solution is indeed correct. These SOCs require some minor technical conditions on the functional forms of $F(\cdot;\cdot,\cdot)$ and $c(\cdot)$. For example, for the output process in (1), if $h(x,k)$ is increasing and concave in $x$ and $k$ separately and $\bar{e}$ follows the normal distribution $N(0,\sigma^2)$, then when $\theta$ is close to 1 both SOCs are satisfied. Also, in problem (2), we can add a SOC: $\frac{d^2\Pi_M}{\partial x^2} < 0$. It turns out that the results in Propositions 1-3 are exactly the same if this SOC is imposed. There are an infinite number of efficient solutions; the SOC will reduce the size of this solution set.

### 4. Applications

This section focuses on the two most popular types of MCs: DECS and PRECS (Arzac, 1997). **DECS** stands for Debt Exchangeable for Common Stock, and **PRECS** stands for Preferred Equity Redemption Cumulative Stock.
4.1. DECS and PERCS

A popular type of MCs is DECS. A DECS can be denoted as \( (r, p_2, p_3, \tau(\cdot)) \) with dividend rate \( r \), trigger prices \( p_2 \) and \( p_3 \), and convertible ratio \( \tau(\cdot) \) that has the following form:

\[
\tau(p) = \begin{cases} 
1 & \text{if } p \leq p_2, \\
p_2/p & \text{if } p_2 < p < p_3, \\
p_2/p_3 & \text{if } p \geq p_3.
\end{cases}
\]  

This conversion ratio has the advantage of being easy to understand and easy to verify using publicly known information. It turns out that a DECS can be efficient, which is stated in Proposition 2.

**Proposition 2** (DECS). MCs of the DECS type can be efficient. Specifically,

(a) When \( y_0 < rk_1^* \), there is an efficient MC \( (r, p_2, p_3, \tau(\cdot)) \) of the DECS type with dividend rate \( r \) and trigger stock prices \( p_2 \) and \( p_3 \), \( p_3 > p_2 \geq p_1^* \), if and only if \( c'(x^*) > (y_0 - rk_1^*)F_x(y_1^*, x^*, k_1^*) \).

(b) When \( y_0 \geq rk_1^* \), any MC \( (r, p_2, p_3, \tau(\cdot)) \) of the DECS type with dividend rate \( r \) and trigger stock prices \( p_2 \) and \( p_3 \), \( p_3 > p_2 \geq p_1^* \), is inefficient.

The second popular type of MCs is PERCS. A PERCS can be denoted as \( (r, p_2, \tau(\cdot)) \) with dividend rate \( r \), trigger stock price \( p_2 \), and convertible ratio \( \tau(\cdot) \) of the following form:

\[
\tau(p) = \begin{cases} 
1 & \text{if } p < p_2, \\
p_2/p & \text{if } p \geq p_2.
\end{cases}
\]  

By this conversion ratio, a holder has 100% capital gain when the stock price is below \( p_2 \), but no capital gain when the stock price is higher than \( p_2 \). It turns out that a PERCS can be efficient, which is stated in Proposition 3.

**Proposition 3** (PERCS). MCs of the PERCS type can be efficient. Specifically,

(a) When \( y_0 < rk_1^* \), there is an efficient MC \( (r, p_2, \tau(\cdot)) \) of the PERCS type with dividend rate \( r \) and trigger price \( p_2 \), if and only if \( c'(x^*) > (y_0 - rk_1^*)F_x(y_1^*, x^*, k_1^*) \).

(b) When \( y_0 \geq rk_1^* \), any MC \( (r, p_2, \tau(\cdot)) \) of the PERCS type with dividend rate \( r \) and trigger price \( p_2 \), is inefficient.

We have several remarks on these two popular MCs. First, Propositions 2 and 3 show that, although PERCS have a simpler conversion ratio than DECS, they are equally efficient.

Second, for both DECS and PERCS, the condition \( y_0 < rk_1^* \) is necessary for efficiency. This condition implies that the project is sufficiently large to the firm so that the interest payment
on the initial investment is larger than the firm’s original firm value. It is consistent with the
fact that MCs are often used in large financial projects in practice.

Third, the MCs presented in all our propositions are the MCs issued at \( t = 0 \). These MCs
are interesting since they face uncertainty, moral hazards and information revelation. The
MCs issued at \( t = 1 \) are trivial, since they do not face information revelation in the future.

Fourth, condition \( c'(x^*) > (y_0 - r k_i^*) f^2 (y'_i, x', k_i^*) \) means that the manager’s marginal cost
of effort is sufficiently large. It induces the firm to offer investors a better deal, which leads to
efficiency.

Finally, Table 2 from Ertugrul et al. (2004, Table II) suggests that DECS are gaining pop-
ularity and PERCS are losing popularity over time. Chen et al. (1999) have also observed that
PERCS have tailed off since 1993 and DECS have gained popularity since 1993; in fact, they
argue that DECS were introduced to replace PERCS. Ertugrul et al.’s (2004) empirical finding
suggests that DECS are able to more accurately convey the firms’ future value through its more
flexible conversion ratio. Using graphic illustrations, Calamos (2003, Figures 1.14 and 1.15)
argue that DECS offer a better risk-reward profile than PERCS.

<table>
<thead>
<tr>
<th>Year</th>
<th>DECS</th>
<th>PERCS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>1992</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>1993</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>1994</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>1995</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>1996</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>1997</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1998</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>1999</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>2000</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>2001</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>2002</td>
<td>16</td>
<td>0</td>
</tr>
</tbody>
</table>

4.2. Case Studies

We now show two real-world examples.

Case 1

This first case is from Chemmanur et al. (2014). In June 2000, Valero Energy Corporation
issued $150 million worth of PEPS (Premium Equity Participating Securities) at $25 per unit.
The common stock was selling at price $29.125 at the issue time. The PEPS paid a quarterly dividend of 7.75%, while the dividend on the underlying common stock was 2.75%. The PEPS were mandatorily convertible to shares of common stock on August 18th, 2003. The shares of common stock received for each unit of MC depend on the stock price \( p \) at the conversion time according to the following formula:

\[
\begin{cases} 
0.85837 \text{ units if } p \leq 29.125, \\
\frac{25}{p} \text{ units if } 29.125 < p < 34.95, \\
0.71531 \text{ units if } p \geq 34.95.
\end{cases}
\]

(13)

Here, \( 0.85837 \times 29.125 = 25 \) and \( 0.71531 \times 34.95 = 25 \). In other words, when the stock price is between the initial price $29.125 and its 20% appreciation price $34.95, the holder receives no capital appreciation; when the stock price goes down, the holder suffers full capital depreciation; and when the stock price is above 20% appreciation, the holder receives full capital appreciation.

**Case 2**

This second case can be found on the internet. In March 2007, Freeport-McMoran Copper & Gold Inc. issued $2.5 billion worth of a MC at price $100 per unit. The common stock was selling at $61.25 per share at the issue time. The dividend rate is 6.75%. The convertible is mandatorily convertible on May 1, 2010. The shares of common stock for each unit of MC depend on the stock price \( p \) at the conversion time according to the following formula:

\[
\begin{cases} 
1.6327 \text{ units if } p \leq 61.25, \\
\frac{100}{p} \text{ units if } 61.25 < p < 73.50, \\
1.3605 \text{ units if } p \geq 73.50.
\end{cases}
\]

(14)

Here, \( 1.6327 \times 61.25 = 100 \) and \( 1.3605 \times 73.50 = 100 \). In other words, when the stock price is between the initial price $61.25 and its 20% appreciation price $73.50, the holder receives no capital appreciation; when the stock price goes down, the holder suffers full capital depreciation; and when the stock price is above 20% appreciation, the holder receives full capital appreciation.
5. Concluding Remarks

This paper investigates a popular type of convertibles: MCs. We focus on three common features in MCs: mandatory convertibility, a high dividend rate, and capped capital appreciation. We identify efficient MCs and we show that the three known features are necessary for efficiency.

Following Chemmanur et al.’s (2014) first theory on MCs using the asymmetric-information approach, we propose a theory using the incomplete-contract approach. We view an asymmetric-information model as having an exogenous type and an incomplete-contract model as having an endogenous type. This difference in endogeneity of information leads to very different results. First, all MCs in Chemmanur et al. (2014) are inefficient, while our MCs in Propositions 1-3 are efficient under sensible conditions. Second, only certain type of firms issues MCs in Chemmanur et al. (2014), while any firm can issue MCs in our model as long as it is willing to offer a high dividend rate. Third, large financing projects are more likely to issue MCs in our model since they tend to have a large $rk^i$.

Using a similar model as Stein’s (1992) with financial distress and information asymmetry, Chemmanur et al. (2014) find that firms facing higher probability of financial distress and smaller extent of information asymmetry will choose to issue MCs, while those facing lower probability of financial distress and greater extent of information asymmetry will choose to issue ordinary convertibles. However, empirical studies on this prediction are inconclusive. Chemmanur et al. (2014) find empirical support for this theoretical prediction. But, using a larger database that includes debt, equity, ordinary convertibles and MCs, Ertugrul et al. (2004) do not find support for this conclusion. In fact, Ertugrul et al. (2004) find that MC issuers are not more likely to have more financial distress than ordinary convertible issuers; on the contrary, MC issuers should typically have lower ex-ante firm risk. Regarding our theory, we presented two real-world examples that seem to be perfectly consistent with our theory. These two cases are typical among real-world MCs. However, there is so far no empirical study along the incomplete-contract approach and on our theory in particular.

Besides their efficiency property, MCs may be chosen for a few other possible reasons. Since MCs are treated as equity on the issuing company’s balance sheet, companies may choose to issue MCs in order to improve their capital structure. However, investors may be deterred by the fact that MCs have the same downside risk as the underlying common stock. Although MCs offer short-term protection on downside risk before conversion, they put a cap on capital appreciation. Unlike most convertibles that are traded over the counter, MCs are traded on the open market, which makes them more accessible to retail investors. Besides, MCs also have the greatest tax advantage among convertibles. Although these reasons may help explain the popularity of MCs, the efficiency property is an essential question to address.
Appendix

Proof of Lemma 1

With \( y_1 = y_0 + K + (\theta - 1)k_1 \) being dependent on \( k_1 \), the first-order condition for \( k_1 \) is

\[
0 = \frac{\partial V}{\partial k_1} = (\theta - 1)F(y_1, x, k_1) + y_1 F_{k_1}(y_1, x, k_1) + (\theta - 1)y_1 f(y_1, x, k_1) + \int_{y_1}^{\infty} ydF_{k_1}(y, x, k_1) - (\theta - 1)y_1 f(y_1, x, k_1)
\]

\[
= (\theta - 1)F(y_1, x, k_1) + y_1 F_{k_1}(y_1, x, k_1) + \int_{y_1}^{\infty} ydF_{k_1}(y, x, k_1)
\]

\[
= (\theta - 1)F(y_1, x, k_1) - \int_{y_1}^{\infty} F_{k_1}(y, x, k_1)dy.
\]

The first-order condition for \( x \) is

\[
0 = \frac{\partial V}{\partial x} = y_1 F_x(y_1, x, k_1) + \int_{y_1}^{\infty} ydF_{k_1}(y, x, k_1) - c'(x) = -\int_{y_1}^{\infty} F_x(y, x, k_1)dy - c'(x).
\]

Here, for simplicity of presentation, we have implicitly assumed that, for \( z = x, k_1 \),

\[
\lim_{y \to \infty} yF_x(y, x, k_1) = 0 \quad \text{and} \quad \int_{y_1}^{\infty} F_x(y, x, k_1)dy \quad \text{is well defined.} \tag{15}
\]

The first condition above holds if the mean output is mathematically well defined. The second condition means that \( F_x(y, x, k_1) \) is Lebesgue-integrable. These two are trivial requirements.

Proof of Proposition 1

Since the objective function in (10) does not depend on \( r \) and \( \tau(\cdot) \), we can solve problem (2) by the following two steps. We first let \( x \) and \( k_1 \) be the first-best investments \((x', k_1')\) in Lemma 1. This first-best solution maximizes the objective function in (10) when no conditions are imposed. Second, given \((x', k_1')\) and denoting

\[
k_2 \equiv K - k_1', \quad y_2' \equiv y_0 + \theta k_1' + k_2', \quad p_1' \equiv y_0 + (\theta - r)k_1',
\]

we need to find a set of variables \( \{r, y_2, y_3, \tau_2, \tau_3, \tau_m(\cdot)\} \) with \( r, \tau_2, \tau_3 \in [0, 1] \) and \( y_3 > y_2 \geq y_2' \) that ensure the IC and IR conditions in (2).

We consider the IR condition in (2) first. If \( r \to 1 \), we obviously have \( \Pi_i > 0 \) for \( \Pi_i \) defined in (8). That is, we can guarantee the IR condition by choosing a sufficiently large \( r \).

We now consider the IC condition in (2). We arbitrarily set \( y_2 = y_2' \). Then,
\[ \frac{\partial \Pi_m}{\partial x} = (y_0 + (1 - \tau_2)\theta k'_1 - r k'_1)E_x(y'_1, x', k'_1) + \int_{y'_1}^{\infty} [1 - \tau(y - r)K](y - r)KdF_x(y, x', k'_1) - c'(x') \]
\[ = (y_0 + (1 - \tau_2)\theta k'_1 - r k'_1)E_x(y'_1, x', k'_1) + \int_{y'_1}^{\infty} [1 - \tau_m(y - r)K](y - r)KdF_x(y, x', k'_1) \]
\[ + (1 - \tau_3) \int_{y'_3}^{\infty} (y - r)KdF_x(y, x', k'_1) - c'(x') \]
\[ = (y_0 + (1 - \tau_2)\theta k'_1 - r k'_1)E_x(y'_1, x', k'_1) - \int_{y'_1}^{\infty} \tau_m(y - r)KdF_x(y, x', k'_1) - c'(x') \]
\[ = (y_0 + (1 - \tau_2)\theta k'_1 - r k'_1 - (y'_3 - r)k'_3)E_x(y'_1, x', k'_1) \]
\[ - \int_{y'_1}^{\infty} \tau_m(y - r)KdF_x(y, x', k'_1) - \tau_3 \int_{y'_3}^{\infty} (y - r)KdF_x(y, x', k'_1), \]

where we have used condition (6) to eliminate \( c'(x) \). Then, the IC condition is satisfied by the first-best investments \((x', k'_1)\) if we choose the following \( \tau_3 \):

\[ \tau_3 = - \frac{\left[ \tau_2 \theta k'_1 + (1 - r)k'_2 \right]E_x(y'_1, x', k'_1) + \int_{y'_1}^{y_3} \tau_m(y - r)KdF_x(y, x', k'_1) \int_{y'_3}^{\infty} (y - r)KdF_x(y, x', k'_1)}{\int_{y'_3}^{\infty} (y - r)KdF_x(y, x', k'_1)}. \] (16)

We now need to ensure \( \tau_3 \in [0, 1] \). We have

\[ \int_{y'_3}^{\infty} (y - r)KdF_x(y, x', k'_1) = -(y'_3 - r)K \int_{y'_3}^{\infty} E_x(y, x', k'_1)dy > 0, \] (17)

where we have used the assumption in (15). Hence, if \( y_3 \to y'_1 \), we have \( \tau_3 > 0 \). On the other hand, condition \( \tau_3 < 1 \) is equivalent to

\[ - \left[ \tau_2 \theta k'_1 + (1 - r)k'_2 \right]E_x(y'_1, x', k'_1) - \int_{y'_1}^{y_3} \tau_m(y - r)KdF_x(y, x', k'_1) < \int_{y'_3}^{\infty} (y - r)KdF_x(y, x', k'_1). \]

By (17), we need

\[ \left[ \tau_2 \theta k'_1 + (1 - r)k'_2 \right]E_x(y'_1, x', k'_1) + \int_{y'_1}^{y_3} \tau_m(y - r)KdF_x(y, x', k'_1) > (y'_3 - r)K \int_{y'_3}^{\infty} E_x(y, x', k'_1)dy. \]

Then, by letting \( y_3 \to y'_1 \), we need

\[ \left[ \tau_2 \theta k'_1 + (1 - r)k'_2 \right]E_x(y'_1, x', k'_1) > (y'_3 - r)K \int_{y'_3}^{\infty} E_x(y, x', k'_1)dy, \]
or
\((y_0 + (1 - \tau_2)\theta k_1^* - r k_1^*)F_x(y_1^*, x^*, k_1^*) + \int_{y_1^*}^{\infty} F_x(y, x^*, k_1^*) \, dy < 0.\)

If we choose \(\tau_2\) sufficiently close to 0, the above is guaranteed if \(p_1^* = y_0 + (\theta - r)k_1^* \geq 0.\) That is, we can find \(\tau_2, \tau_3 \in [0, 1]\) to satisfy the IC condition if \(y_2 = y_1^*\) and \(\tau_3\) is close to \(y_2\) while \(\tau_m(p)\) can be an arbitrary function.

**Proof of Proposition 2**

For a MC of the DECS type with a conversion ratio defined in (11), our task is to find \(p_2, p_3\) and \(r\) such that the second-best problem in (2) achieves the first best. We now show the existence of an efficient MC of the DECS type directly without using Proposition 1.

Consider the IR condition in (2) first. By (8), we have

\[
\Pi_i = [\theta k_1 + (1 - r)k_2]F(y_1, x, k_1) + \int_{y_1}^{\infty} \tau(y - rK)(y - rK)dF(y, x, k_1) - (1 - r)K.
\]

We have \(\Pi_i > 0\) if \(r\) is close to 1. That is, the IR condition can be guaranteed by a sufficiently large dividend rate.

We now consider the IC condition in (2). For the first-best investments \((x^*, k_1^*)\), denote

\[
k_2^* \equiv K - k_1^*, \quad y_1^* \equiv y_0 + \theta k_1^* + k_2^*.
\]

We will find \(y_2, y_3\) and \(r\) satisfying \(y_3 > y_2 \geq y_1^*\) and \(r \in [0, 1]\) so that the IC condition is satisfied. By (9), with \(r(p)\) defined in (11), we have

\[
\frac{\partial \Pi_m}{\partial x} = (y_0 - r k_1^*)F_x(y_1^*, x^*, k_1^*) + \int_{y_2}^{y_3} (y - rK)dF_x(y, x^*, k_1^*) - c'(x^*)
\]

\[
= (y_0 - r k_1^*)F_x(y_1^*, x^*, k_1^*) + (y_3 - y_2)F_x(y_3, x^*, k_1^*) - \int_{y_2}^{y_3} F_x(y, x^*, k_1^*)dy
\]

\[
- \frac{y_3 - y_2}{y_3 - rK}\int_{y_3}^{\infty} (y - rK)dF_x(y, x^*, k_1^*) - c'(x^*)
\]

\[
= (y_0 - r k_1^*)F_x(y_1^*, x^*, k_1^*) - \int_{y_2}^{y_3} F_x(y, x^*, k_1^*)dy
\]

\[
- \frac{y_3 - y_2}{y_3 - rK}\int_{y_3}^{\infty} F_x(y, x^*, k_1^*)dy - c'(x^*)
\]

\[
= (y_0 - r k_1^*)F_x(y_1^*, x^*, k_1^*) - \int_{y_2}^{y_3} F_x(y, x^*, k_1^*)dy
\]

\[
- \frac{y_3 - y_2}{y_3 - rK}\int_{y_3}^{\infty} F_x(y, x^*, k_1^*)dy
\]

where we have used (6). Then,
\[
\frac{\partial \Pi_M}{\partial x} = (y_0 - rk'_1)F_x(y'_1, x', k'_1) - \int_{y_2}^{y_3} F_x(y, x', k'_1)dy - \int_{y_2}^{\infty} F_x(y, x', k'_1)dy
\]
\[
- \frac{y_3 - y_2}{y_3 - rK} \int_{y_3}^{\infty} F_x(y, x', k'_1)dy + \int_{y_1}^{y_2} F_x(y, x', k'_1)dy
\]
\[
= (y_0 - rk'_1)F_x(y'_1, x', k'_1) + \frac{y_2 - rK}{y_3 - rK} \int_{y_3}^{\infty} F_x(y, x', k'_1)dy + \int_{y_1}^{y_2} F_x(y, x', k'_1)dy.
\] (18)

If \(y_0 \geq rk'_1\), all three terms in the above are negative for any \(y_3 > y_2 > rK\), implying \(\frac{\partial \Pi_M}{\partial x} < 0\) under any circumstances. Hence, all MCs of the DECS type are inefficient.

However, if \(y_0 < rk'_1\), when we take \(y_2 = y'_1\) and \(y_3 \to \infty\), by (18), we have \(\frac{\partial \Pi_M}{\partial x} > 0\). On the other hand, with \(y_2 = y'_1\), if we let \(y_3 \to y'_1\), by (18), we have
\[
\frac{\partial \Pi_M}{\partial x} = (y_0 - rk'_1)F_x(y'_1, x', k'_1) + \int_{y_1}^{y_3} F_x(y, x', k'_1)dy = (y_0 - rk'_1)F_x(y'_1, x', k'_1) - c'(x').
\]

Hence, we have \(\frac{\partial \Pi_M}{\partial x} < 0\) if \(c'(x') > (y_0 - rk'_1)F_x(y'_1, x', k'_1)\). Therefore, by the continuity of \(\frac{\partial \Pi_M}{\partial x}\) in \(y_3\), there exists \(y_3 \in (y'_1, \infty)\) such that the IC condition \(\frac{\partial \Pi_M}{\partial x} = 0\) is satisfied.

But, when \(y_0 < rk'_1\), if \(c'(x') \leq (y_0 - rk'_1)F_x(y'_1, x', k'_1)\), by (18), we have
\[
\frac{\partial \Pi_M}{\partial x} \geq (y_0 - rk'_1)F_x(y'_1, x', k'_1) + \int_{y_1}^{y_3} F_x(y, x', k'_1)dy + \int_{y_1}^{y_2} F_x(y, x', k'_1)dy
\]
\[
> (y_0 - rk'_1)F_x(y'_1, x', k'_1) + \int_{y_1}^{y_2} F_x(y, x', k'_1)dy + \int_{y_1}^{y_2} F_x(y, x', k'_1)dy
\]
\[
= (y_0 - rk'_1)F_x(y'_1, x', k'_1) - c'(x') \geq 0.
\]

That is, when \(y_0 < rk'_1\), condition \(c'(x') \leq (y_0 - rk'_1)F_x(y'_1, x', k'_1)\) ensures \(\frac{\partial \Pi_M}{\partial x} > 0\) for any \(y_3 > y_2 > rK\) so that the IC condition cannot be satisfied. If so, all MCs of the DECS type are inefficient.

**Proof of Proposition 3**

For a MC of the PERCS type with a conversion ratio defined in (12), our task is to find \(p_2\) and \(r\) such that the second-best problem in (2) achieves the first best. We now show the existence of an efficient MC of the PERCS type directly without using Proposition 1.

Consider the IR condition in (2) first. By (8), we have
\[
\Pi_i = [\theta k_1 + (1 - r)k_2]F(y_1, x, k_1) + \int_{y_1}^{\infty} \tau(y - rK)(y - rK)dF(y, x, k_1) - (1 - r)K.
\]

We have \(\Pi_i > 0\) if \(r\) is sufficiently large. That is, the IR condition in problem (2) can be guaranteed by a sufficiently large dividend rate.

We now look into the IC condition in problem (2). By (9), we have
\[ \Pi_M = (y_0 - rk_1)F(y_1, x, k_1) + \int_{y_2}^{\infty} \left( 1 - \frac{y_2 - rK}{y - rK} \right) (y - rK) dF(y, x, k_1) - c(x). \]

Then, with \( y_2 \geq y_1 \),
\[ \frac{\partial \Pi_M}{\partial x} = (y_0 - rk_1)F_x(y_1, x, k_1) + \int_{y_2}^{\infty} (y - y_2) dF_x(y, x, k_1) - c'(x) \]
\[ = (y_0 - rk_1)F_x(y_1, x, k_1) - \int_{y_2}^{\infty} F_x(y, x, k_1) dy - c'(x). \]

At the first-best \((x', k'_1)\), with \( k'_2 = K - k'_1 \) and \( y'_1 = y_0 + \theta k'_1 + k'_2 \), we have
\[ \frac{\partial \Pi_M}{\partial x} = (y_0 - rk'_1)F_x(y'_1, x', k'_1) - \int_{y'_2}^{\infty} F_x(y, x', k'_1) dy - c'(x'). \] (19)

Using (6), with \( F_x < 0 \), we have
\[ \frac{\partial \Pi_M}{\partial x} < (y_0 - rk'_1)F_x(y'_1, x', k'_1) - \int_{y'_1}^{\infty} F_x(y, x', k'_1) dy - c'(x') = (y_0 - rk'_1)F_x(y'_1, x', k'_1). \]

If \( y_0 \geq rk'_1 \), then \( \frac{\partial \Pi_M}{\partial x} < 0 \) for any \( y_2 > y'_1 \). This means that MCs of the PERCS type are inefficient.

If \( y_0 < rk'_1 \), by (19), we have \( \frac{\partial \Pi_M}{\partial x} > 0 \) when \( y_2 \) is close to \( y'_1 \). On the other hand, if we let \( y_2 \to \infty \), then by (19)
\[ \frac{\partial \Pi_M}{\partial x} = (y_0 - rk'_1)F_x(y'_1, x', k'_1) - c'(x'). \]

Hence, \( \frac{\partial \Pi_M}{\partial x} < 0 \) when \( y_2 \) is large enough if
\[ c'(x') > (y_0 - rk'_1)F_x(y'_1, x', k'_1). \]
If so, by the continuity of \( \frac{\partial \Pi_M}{\partial x} \) in \( y_2 \), we can find a \( y'_2 \in (y'_1, \infty) \) such that \( \frac{\partial \Pi_M}{\partial x} = 0 \), implying the IC condition. This means that there exists a MC that satisfied the IC and IR conditions in problem (2) for the first-best investments \((x', k'_1)\). This MC is efficient.

But, when \( y_0 < rk'_1 \), if \( c'(x') \leq (y_0 - rk'_1)F_x(y'_1, x', k'_1) \), by (19), we have
\[ \frac{\partial \Pi_M}{\partial x} \geq (y_0 - rk'_1)F_x(y'_1, x', k'_1) - c'(x') \geq 0. \]
That is, when \( y_0 < rk'_1 \), condition \( c'(x') \leq (y_0 - rk'_1)F_x(y'_1, x', k'_1) \) ensures \( \frac{\partial \Pi_M}{\partial x} > 0 \) for any \( y_2 > rK \) so that the IC condition cannot be satisfied. If so, all MCs of the PERCS type are inefficient.
References


