Collateral-Based Asset Pricing*

Roberto Steri
roberto.steri@unil.ch

December 7, 2015

Abstract

Recent corporate finance studies show that hedging is a first-order driver of corporate decisions. I use firms' hedging behavior to build a novel asset pricing model, the Corporate CAPM. I propose a dynamic contracting framework in which collateral constraints induce a tradeoff between hedging and immediate needs for funding. Firms hedge by transferring resources to future states that are most important for firm's value. In the model, firms' hedging behavior is informative of the shareholders' stochastic discount factor, which measures the value of each state. As a consequence, discount rates can be inferred from firm's observed investment, financing, and hedging policies. On the corporate finance side, a calibrated version of the model is broadly consistent with observed corporate policies of US listed firms. On the asset pricing side, the Corporate CAPM is successful in pricing different test assets, also in comparison to leading asset pricing models.

JEL Classification: C61, C63, D21, D24, G10, G12, G31, G32, G35.
Keywords: Dynamic Contracting, Expected Returns, Hedging, Capital Asset Pricing Model.

*I am indebted to Lukas Schmid, my main advisor at Duke University, who supported, encouraged, and mentored me during my PhD. I owe special thanks to Domenico Ferraro and Pino Lopomo for the incredible amount of time and support in all stages of this work. I am also grateful to Hengjie Ai, Alessandro Arlotto, Ravi Bansal, Ian Dew-Becker, Alex Belloni, Frederico Belo, Wayne Ferson, Massimo Guidolin, Shiyang Huang, Filippo Ippolito, Dino Palazzo, Adriano Rampini, David Robinson, S. Vish Viswanathan, seminar participants to the Fuqua Lunch Seminar at Duke University, University of Pennsylvania (Wharton), University of Minnesota, University of Southern California, Ohio State University, Boston University, HEC Lausanne, INSEAD, Norwegian School of Economics, Warwick Business School, Universite Laval, the 2014 European Winter Meeting of the Econometric Society, the 2015 Annual Meeting of the Society for Economic Dynamics, the 2015 China International Conference in Finance. Roberto Steri is affiliated with the University of Lausanne and the Swiss Finance Institute, Batiment Extranef, 1015, Lausanne, Switzerland.

1. Introduction

Stochastic discount factors are the cornerstone of modern asset pricing. They allow to compute asset prices as the expected discounted value of future cashflows. Asset pricing theory ordinarily focuses on the demand of securities, and derives a stochastic discount factor from the optimizing behavior of an investor who decides over consumption and portfolio allocations. In this paper, I instead focus on the supply of securities. I build upon corporate finance theory to identify a stochastic discount factor from firms' policies. This leads to a novel asset pricing model, the Corporate CAPM.

My approach mainly relates to the papers of Jermann (2010) and Belo (2010), who recover a stochastic discount factor from firm's real investment decisions on the production side of the economy. Instead, I recover a stochastic discount factor from firms' hedging behavior. Hedging is not only a pivotal economic mechanism in corporate finance, but also a fundamental channel through which firms transfer resources across states of the world. As previous studies show, hedging is a quantitatively first-order corporate policy. Hedging is practically implemented with combinations of traditional debt instruments and other financial instruments like lines of credit and financial derivatives. In particular, credit lines appear to be a prominent implementation of hedging. Sufi (2009) reports that credit lines constitute more than 80 percent of bank debt for public firms in the US. Colla, Ippolito, and Li (2013) report that the drawn part of credit lines accounts for 22 percent of their total debt.[1] A firm that hedges a state reveals information on the importance of that state for its own value. The value of each state is also measured by the owners' stochastic discount factor. Therefore, the stochastic discount factor can be identified through observed firms' decisions, and used to price the assets in the economy. The concept of hedging I entertain here draws on the close connection between collateralized financing and risk management recognized by Rampini and Viswanathan (2010), and Rampini and Viswanathan (2013). Hedging and financing both involve promises to pay from the firm to external lenders in some states of the world. Firms can contract different payments across states, and implement hedging with state-contingent promises

[1] In addition, Rampini and Viswanathan (2010), Bolton, Chen, and Wang (2011), and Rampini and Viswanathan (2013), highlight the importance of hedging to understand firm's growth, investment, and financing policies. Nikolov, Schmid, and Steri (2013), and Li and Whited (2013) document the quantitative importance of hedging for firms' policies.

to pay that induce a distribution of preserved debt capacity. Collateral constraints arising from limited enforcement restrict such promises, and hence the amount of resources firms can effectively transfer across states.

The Corporate CAPM expresses the stochastic discount factor in terms of firms' characteristics, and can be approximated as a linear two-factor model. The factors are a "hedging" factor, which equals the change in firms' net worth[2], and a "profitability" factor, which is associated to the change in firms' productivity. The two factors have an intuitive interpretation. They measure the amount of resources available to the firm in a certain state. These resources (net worth) can be either actively transferred through hedging, or originate from existing profitable investments. I implement asset pricing tests with the Generalized Method of Moments (GMM) to assess the empirical performance of the model. As the recent empirical literature recommends (Lewellen, Nagel, and Shanken (2010), Daniel and Titman (2012)), I consider different test assets in empirical tests, namely the Fama-French 25 portfolios sorted by size and book-to-market equity, the 30 Fama-French industry portfolios, and 25 portfolios sorted by market and HML beta as in Yogo (2006). Overall, the Corporate CAPM finds support in the data. The model prices the test assets well, and delivers low pricing errors even in comparison to leading asset pricing models, as the CAPM, the Consumption CAPM, and the Fama and French three-factor model. Historically, asset pricing models obtained from consumption-based stochastic discount factors have not succeeded in accounting for the variation of expected returns across stocks. One important reason for their empirical failure is the smoothness of consumption data. This prevents expected returns to line up with covariances with consumption aggregates, as these models predict. On the contrary, the Corporate CAPM gets traction because it links the stochastic discount factor to firms' characteristics, which exhibit larger fluctuations.

My theoretical framework is a dynamic contracting model. Hedging is in fact an inherently dynamic process. Firms engage in hedging to transfer resources from today to future times and states when they are more valuable. For instance, a firm might hedge specific future states to finance profitable investment opportunities, or to pay out more dividends in bad times. In the model, firms have valuable investment opportunities that

[2]As standard in the dynamic contracting literature, net worth is the firm's counterpart of household's wealth, and captures how constrained a company is with respect to funds to allocate to investment, and distributions.

arise stochastically over time. However, they have limited funds, and they sign contracts with external lenders to aid external financing of profitable investments. Contracts have limited enforcement. The entrepreneur has the option to renege the contract and divert capital for their own private benefit. In equilibrium, this limited commitment problem endogenously imposes a collateral constraint, and firms implicitly borrow constrained against their equity value. In this context, value maximization provides a rationale to hedge more valuable states, in a tradeoff with their funding needs for current investment and distributions. Firms' debt capacity is limited, and firms can preserve it for specific future states by optimally contracting state-contingent repayments with the lender. A firm can therefore hedge any future state by promising a low repayment in the case that state occurs. Hence, firms can in effect transfer resources (net worth) across states. In this setting, the stochastic discount factor reflects which state must have led a firm to optimally make its observed decisions, and can be backed out from the firms' state-by-state first-order conditions with respect to debt repayments. Conditional on how financially constrained they are, firms implement investment and financing policies to transfer resources to most important states, where the stochastic discount factor is high.

On the corporate finance side, I solve the model numerically and I find that a calibrated version is quantitatively consistent with basic stylized facts about corporate investment and financing, and with key aggregated asset pricing moments. To solve the model, and to determine the properties of the optimal contract, I formulate the contracting problem recursively as an infinite-horizon dynamic programming problem. The problem has a nonstandard topological structure because of the presence of the objective function, the firm's equity value, in the borrowing constraint. I use Knaster-Tarski (Tarski (1955)) fixed point theorem[3] to prove the problem has a well-defined equilibrium. In addition, the number of decision variables is high because of state-contingent hedging decisions. To deal with this issue, I introduce an equivalent mixed-integer programming representation of the dynamic programming problem. The equivalent problem is a natural extension of the extant linear programming methods for dynamic programming to the specific topological structure of the model. These methods have been introduced in finance by Trick and Zin (1993), and then extended to large state spaces by Nikolov, Schmid, and Steri (2013). As in Nikolov, Schmid, and Steri (2013), I exploit a separation

[3]See Aliprantis and Border (2006), and Kamihigashi (2012).

oracle, an auxiliary linear programming problem, to achieve computational efficiency.

This paper lies at the intersection of three lines of research. First, it relates to the large literature that develops quantitative production models to investigate the cross-section of equity returns. Recent contributions include Zhang (2005), Livdan, Sapriza, and Zhang (2009), Gomes and Schmid (2010), Garlappi and Yan (2011), Obreja (2013), and Bazdrech, Belo, and Lin (2013). With respect to these papers, my focus is to obtain a stochastic discount factor, instead of rationalize observed spreads in returns with respect to specific firms' characteristics. Second, the paper builds upon the literature on hedging and dynamic contracting in corporate finance, that refers to Rampini and Viswanathan (2010), Rampini and Viswanathan (2013), Rampini, Sufi, and Viswanathan (2013), and whose quantitative implications have been examined in Li and Whited (2013), and Nikolov, Schmid, and Steri (2013). In this context, this paper analyzes the asset pricing implications of contracting models of hedging. Finally, this work is closely related to the literature that attempts to identify a stochastic discount factor in production models from firms' policies and data, Cochrane (1993), Jermann (2010), and Belo (2010). The key difference with these works is the economic mechanism that allows to identify the stochastic discount factor from firms' decisions.[4]

This work has potential implications for future research. As Cochrane (2011) discusses, research in asset pricing ultimately aims at understanding how asset returns and consumption are jointly determined in general equilibrium. In this perspective, the identification of a stochastic discount factor from the production side of the economy imposes additional restrictions that may provide further guidance for modeling the consumption side of the economy, rather than representing a competing approach. On the empirical side, new testable hypotheses for cross-sectional differences in returns can be developed from the present framework, especially from the observation that variables that describe firms' policies enter the stochastic discount factor directly.

The paper is organized as follows. Section 2 develops the key intuition of the paper in a two-period example. Section 3 presents the dynamic contracting model, and

[4]Cochrane (1993) and Belo (2010) rely on a representation of production sets in which firms can affect idiosyncratic productivity shocks, and Jermann (2010) investigates the equity premium by taking advantage of state-contingent technologies. Here, the relevant state-contingent action that allows to identify the stochastic discount factor is based on the corporate finance theory of hedging, in the context of dynamic contracting.

describes the numerical solution method. Section 4 introduces the key asset pricing result of the paper, the Corporate CAPM. Section 5 assesses the quantitative performance of the calibrated model for providing a reasonable description of corporate investment and financing decisions. Section 6 presents the empirical tests of the Corporate CAPM. Section 7 concludes.

2. A Two-Period Example

The goal of this section is to convey the main idea of this work with a simple example. Typical production models do not lead to an explicit expression for the stochastic discount factor, but only to pricing equations for asset returns. Here I show that when firms transfer resources across states of nature through risk management, the stochastic discount factor can be instead backed out from firms' optimization conditions. The argument proceeds as follows. I first illustrate in a two-period model why when firms cannot implement risk management the stochastic discount factor cannot be obtained from the firm's problem. I then show why introducing hedging decisions allows to do so.

Consider a model with two periods: today, and tomorrow. Three states of nature, rainy, foggy, and sunny, can possibly occur tomorrow, with probabilities π_R, π_F, and π_S respectively. Consider a firm with an initial wealth endowment w that has access to a production technology. The production technology delivers a stochastic output $A(S)f(k) > 0$ in the sunny state tomorrow, $A(F)f(k) > 0$ in the foggy state tomorrow, and $A(R)f(k) > 0$ in the rainy state tomorrow, with $A(S) > A(F) > A(R)$. $f(\cdot)$ is a production function, and k denotes investment in real capital. The economy ends tomorrow: capital fully depreciates, and a liquidating dividends $d(S)$, $d(F)$, and $d(R)$ are distributed in the sunny, foggy, and rainy states respectively. The firm can borrow from a competitive, risk neutral, and deep-pocket lender at a constant rate R.[5] The firm's problem is to decide over capital k and debt a repayment b to maximize the expected

[5]Section 3 discusses this assumption. Supplementary Appendix A reports the lender's problem.

discounted value of its profits, that is

$$
\begin{aligned}
U(w) &= \max_{k,b} d + \pi_S M(S)d(S) + \pi_F M(F)d(F) + \pi_R M(R)d(R) && (1)\\
& s.t. \\
w + b &= d + k && (2)\\
d(s) &= A(s)f(k) - Rb \qquad s \in \{S, F, R\} && (3)\\
d &\geq 0 && (4)\\
d(s) &\geq 0 \qquad s \in \{S, F, R\} && (5)
\end{aligned}
$$

$M(S)$, $M(F)$, and $M(R)$ are the realizations of the owners' stochastic discount factor in the sunny, foggy, and rainy states[6], Equation (1) is the budget constraint today, and simply equates sources and uses of funds, where d is today's dividend. Equations (4), and (5) rule out negative dividends. Equation (4) states that the firm has access to no other external funds, while Equations (5) guarantee debt is actually riskfree and is repaid tomorrow in all states.[7] Equations (4) and (5) determine limits on the amount the firm can borrow. b must therefore lie in the closed interval $[k - w, A(R)f(k)]$. Equations (5) can be interpreted as collateral constraints, which states that the firm can borrow up to the cash flow it obtains whatever tomorrow's weather is.[8]

Denote by λ the Lagrange multiplier on constraint (4), and by $\pi_s \lambda_s$ the Lagrange multipliers on constraints (5). The first-order conditions of this problem lead to the usual pricing equations for the return on real capital and for the loan interest rate:

$$E[(M(s) + \lambda_s)R^k(s)] = 1 + \lambda \tag{6}$$

$$E[(M(s) + \lambda_s)R] = 1 + \lambda \tag{7}$$

where $s \in \{S, F, R\}$ is an index for the state, and $R^k(s) \equiv A(s)f_k(k)$. Two points are worth noting. First, the pricing equations contain additional terms related to the Lagrange multipliers on the constraints. This reflects the fact that the typical assumption

[6]This objective only requires that a stochastic discount factor exists. This is the case in the absence of arbitrage opportunities. The objective therefore captures the idea that physical assets and riskfree debt are priced consistently with other securities that investors can trade.

[7]Because $d(S) \geq d(F) \geq d(R)$, the constraints in Equation (5) for $s \in \{S, F\}$ are never active.

[8]In the full model, the collateral constraint arises endogenously as an outcome of dynamic contracting.

of free portfolio formation is violated (see Cochrane (2001), Chapter 4).[9] Intuitively, the firm trades real capital and loans.[10] If, for example, the collateral constraint in the rainy state in (5) is binding, the firm cannot freely tilt its portfolio of assets by increasing its debt stock and leaving its capital stock unchanged. The presence of Lagrange multipliers accounts exactly for this restriction. In fact, when the constraints are not binding, Equations (6) and (7) reduce to $E[M(s)R^k(s)] = 1$ and $E[M(s)R] = 1$. Second, and most important, the firm's optimality conditions do not allow to get an expression for the stochastic discount factor in each state. This happens because the firm cannot transfer resources across the rainy, foggy, and sunny states, or equivalently from today to one future state only. As Equations (3) show, by changing its capital and debt decisions in the feasible set, the firm *jointly* increases its payoff in all three states. A unit more of capital generates more output in both states in proportions determined by $A(S)$, $A(F)$, and $A(R)$, and a unit more of debt reduces the payoff by R in both states. A simple algebraic manipulation of Equation (3) indeed shows that the payout in the sunny state can be rewritten as a fixed function of the payoffs in the foggy and rainy states as:

$$d(S) = d(R) + \frac{A(S) - A(R)}{A(F) - A(R)}(d(F) - d(R)) \tag{8}$$

Panel A of Figure 1 makes this idea clear. The solid lines represent the possibility set for the firm's equity payoffs in the sunny and in the rainy state tomorrow for different choices of capital and debt. For simplicity I keep the payoff in the foggy state fixed, although this result holds for every other pair of states. It is immediate to notice that the feasible sets for the payoffs have a kink. In the consumption side of the economy, the condition that these Leontief-type payoffs must be tangent to an indifference curve form the familiar relation $p = E\left[\beta \frac{u'(c(s))}{u'(c)} d(s)\right] = E[M(s)d(s)]$, where p is the price of the firm's equity, c is today's consumption, $u(\cdot)$ is the investor's utility function, and β is his time discount factor. The indifference curves are related to the marginal rate of substitution between today's and tomorrow's consumption, and their slope allows to identify $M(s)$. However, the dashed lines show that any point the firm is willing to choose is consistent

[9] This is very common in models with financial constraints, and the "corrected" discount factor is sometimes denoted as "the firm's discount factor". See for example Mendoza (2000), and Rampini and Viswanathan (2013).

[10] For simplicity assume these assets cannot be traded by the household directly.

with many indifference curves.

[Insert Figure 1 Here]

Consider now the same problem in which the firm is allowed to hedge by setting different debt repayments $b(S)$, $b(F)$ and $b(R)$ for the sunny, foggy, and rainy states. The firm's problem becomes:

$$\begin{aligned}
U(w) &= \max_{k,b} d + \pi_S M(S)d(S) + \pi_F M(F)d(F) + \pi_R M(R)d(R) && (9)\\
& s.t. \\
w + b &= d + k && (10)\\
d(s) &= A(s)f(k) - Rb(s) \qquad s \in \{S, F, R\} && (11)\\
d &\geq 0 && (12)\\
d(s) &\geq 0 \qquad s \in \{S, F, R\} && (13)
\end{aligned}$$

where the amount of debt financing raised today from the risk-neutral lender is $b = E[b(s)]$.[11] The first-order conditions with respect to k, $b(s)$, $s \in \{S, F, R\}$, are:

$$E[(M(s) + \lambda_s)R^k(s)] = 1 + \lambda \tag{14}$$

$$M(s) = \frac{1 + \lambda - \lambda_s R}{R} \tag{15}$$

Equation (14) is the familiar pricing equation for capital, while Equation (15) provides an expression for the stochastic discount factor that must have let to the observed firm's policy. Notice that the difference in the discount rates of lenders and borrowers does not imply the presence of arbitrage opportunities. The stochastic discount factor in fact adapts such that equity claims are priced consistently with the presence of a risk-neutral lender that allows the firm to implement limited risk sharing. This is apparent comparing the stochastic discount factor in Equation (15) with the one for the case without collateral constraints, that is $M(s) = \frac{1}{R}$. In this case the firm guarantees full insurance to the owners, their marginal utility across states is equalized, and equity claims are prices as if the firm were risk neutral. Supplementary Appendix A discusses this case.

[11] As I discuss in Section 3.2, in the complete model lenders offer an elastic supply of credit at all future times and dates at the riskfree rate.

Because the stochastic discount factor is higher in most valuable states, the firm trades off dividend distributions today (with a higher λ) in order to pay out in most important states tomorrow (with a lower λ_s), even though the latter reduces the payout in other states or makes it overall more volatile. With risk-averse investors, most important states are those where aggregate consumption is low and firms are less productive, such as the rainy state in this example. Contingent claims that pay out more in those states are therefore more valuable for investors. In addition, because the solution of the firm's problem depends on its wealth w, two firms with different initial wealth in general implement different policies. This does not mean that there is a stochastic discount factor for each firm. Instead, the firm changes its investment and financing policy in a state-contingent way, depending on whether the state is either sunny, foggy, or rainy. As a consequence, in principle, both firms' policies (and data) could be used as a reference point to back out the stochastic discount factor and to price other assets. In Section 4, I refer to this result as the *relativity property*.

Panel B of Figure 1 illustrates why the stochastic discount factor can be recovered in the presence of hedging. Firms are able to set $b(S)$ and $b(R)$ and determine their payout profile in both the rainy and the sunny states. Contingent claim hyperplanes are therefore differentiable (linear), and indifference curves must be tangent to them at the decision point.

3. The Dynamic Limited Enforcement Model

This section develops a discrete-time dynamic agency model in a neoclassical environment. Entrepreneurs make investment and financing decisions with an infinite time horizon. This ensures they take into account the expected consequences of current actions for the feasibility of future decisions. Dynamic financing is subject to limited enforcement constraints.[12] Firms borrow constrained against their equity value from competitive lenders, and implement state contingent debt repayments up to their debt capacity. The state contingent nature of the contract allows firms to transfer resources to states and times where they are more valuable. In Subsections 3.1 and 3.2, I detail the technology

[12]Related contracting problems are proposed, for example, by Albuquerque and Hopenhayn (2004), Rampini and Viswanathan (2010), Rampini and Viswanathan (2013), Li and Whited (2013), and Schmid and Steri (2013).

and the industry environment, and the financial contracting problem. Despite its conceptual simplicity, this problem has two nonstandard features. First, conventional dynamic programming results do not apply because the equity value enters the enforcement constraint. Second, the presence of state contingent debt repayments as decision variables makes the problem virtually intractable with conventional iterative numerical methods. In Subsection 3.3 I address these two issues. Using a fixed point argument, I first show the existence and uniqueness of the value function as the solution of a dynamic programming problem with an appropriate initial condition. Then, I extend the linear programming techniques in Trick and Zin (1993), Trick and Zin (1997), Nikolov, Schmid, and Steri (2013), and Schmid and Steri (2013), and propose a computationally efficient solution method based on mixed-integer programming. Finally, in Subsection 3.4, I characterize the solution illustrate the qualitative firm optimal investment and financing policies.

3.1. Technology and Competitive Environment

A continuum of perfectly competitive firms operates in an industry. Each firm produces a homogeneous product, whose price is normalized to one. In period t, a fraction ϕ of new firms randomly enters the industry. Existing firms become unproductive and exit with probability ϕ, so that the total mass of operating firms is unchanged over time.

An entrant i arrives with some initial capital stock $k_{i,0}$. Entrants engage in a long-term contract with lenders to obtain external financing. Firms have access to a production technology that generates a stochastic stream of profits

$$\Pi(k_{i,t}, s_{i,t}) \equiv A(s_{i,t})k_{i,t}^{\alpha}$$

where $s_{i,t}$ is a shorthand for the state $\{x_t, z_{i,t}\}$, $k_{i,t}$ is the capital input of firm i at time t, $\alpha \in (0, 1)$ is the curvature parameter of the production function, which exhibits decreasing returns to scale, and $A(s_{i,t})$ is a stochastic process describing productivity. Here $A(s_{i,t}) = x_t z_{i,t}$, where x_t and $z_{i,t}$ are respectively aggregate and firm-specific technology shocks. The idiosyncratic shock $z_{i,t}$ is the driving force of firm-level heterogeneity, and generates a nontrivial cross-section of firms, while the aggregate shock x_t describes the overall technological level of the economy. $z_{i,t}$ and x_t follow Markov processes with finite support

Z and X, and stationary transition functions $Q_z(z_{i,t+1}|z_{i,t})$ and $Q_x(x_{t+1}|x_t)$ as follows:

$$\log(z_{i,t+1}) = \rho_z \log(z_{i,t}) + \sigma_z \epsilon^z_{i,t+1} \tag{16a}$$

$$\log(x_{t+1}) = (1-\rho_x)\mu_x + \rho_x \log(x_t) + \sigma_x \epsilon^x_{t+1} \tag{16b}$$

where $\epsilon^z_{i,t}$ and $\epsilon^z_{j,t}$ are uncorrelated for every $i \neq j$, and ϵ^x_t is uncorrelated with $\epsilon^z_{i,t}$ for every i. $\epsilon^z_{i,t}$ and ϵ^x_t are truncated iid standard normal variables. The capital stock $k_{i,t}$ obeys the law of motion

$$k_{i,t+1} = (1-\delta)k_{i,t} + i_{i,t+1}$$

where δ is the depreciation rate and $i_{i,t+1}$ denotes corporate investment.

3.2. The Contracting Framework

Upon arriving in the industry, the firm enters a long-term contractual relationship with an outside lender. The contract not only provides initial funding, but also financing over the firm's lifecycle. Following several previous studies, lenders are risk neutral and have "deep pockets", that is they offer an elastic supply of credit in all times and states. The risk neutrality assumption is intended to capture the different exposure to risks of lenders and risk-averse firms. Thus, both parties have motives to trade and implement, albeit imperfect, risk sharing. This assumption can be interpreted as a reduced form for lenders having a very large amount of funds to achieve a sufficient diversification and being insensitive to risks arising from granting *individual* loans.[13] The risk neutrality assumption is also convenient because it allows not to put additional structure on the lenders' possible stochastic discount factor. This allows to avoid to explicitly model lenders' decisions and ownership structure, such as bankers' decisions over portfolios of loans and deposits.

Entrepreneurs are risk averse and discount future dividend payouts with a stochastic

[13]In expected utility theory, the risk neutrality assumption captures the evidence for which wealthy individuals behave as if they were risk neutral with respect to small risks (Rabin (2000)). Indeed, in models with large investors, the latter are typically modeled as risk neutral or as agents with CARA utility.

discount factor process $\{M(x_{t+\tau})\}_{\tau=0}^{\infty}$.[14] Risk-neutral lenders' discount factor is instead $R_t \equiv E_t[M(x_{t+1})]$.

The timing of events over a firm's lifecycle is as follows. As soon as a firm enters the industry, it signs a long-term contract with the lender to obtain initial funding. Then, at the beginning of each period, the firm first faces the exogenous exit shock, and the state $s_{i,t}$ realizes. There are no information asymmetries because $s_{i,t}$ is publicly known. The entrepreneur has limited liability, and the firm defaults if its value after observing the shock goes to zero. Second, firm's decisions and operations occur: inputs are purchased, production takes place, revenues are collected, transfers to and from lenders are made, and dividends are distributed. Third, the firm chooses either to renege the contract or to continue operations. This limited enforcement problem is discussed in more detail below. Panel A of Figure 2 summarizes the intra-period timing.[15]

As detailed in Supplementary Appendix B, the contracting problem can be formulated recursively using net worth as a state variable, in line with Abreu, Pearce, and Stacchetti (1990). Realized net worth in a future state $s_{i,t+1}$ determines the amount of resources that are available to the firm in a certain state, net of liabilities. Intuitively, net worth is the corporate counterpart of households wealth, and captures how constrained a company is in terms of resources to allocate to investment, and distributions. The equivalent contracting problem is the following:

[14] The effective discount factor accounts for the probability that the firm exits the industry, that is:

$$M(x_{t+\tau}) = \widehat{M}(x_{t+\tau})(1-\phi)$$

[15] In this setup, the contract has one side commitment. While there is a limited commitment problem on the firm's side, the lender honors the long-term contract. This feature becomes apparent in the recursive formulation in Supplementary Appendix B, where a lender's promise-keeping constraint is part of the problem.

$$
\begin{aligned}
V(w_{i,t}, s_{i,t}) &= \max_{\{d_{i,t}, k_{i,t+1}, b(s_{i,t+1})\}} d_{i,t} + E_t\left[M(x_{t+1})V(w(s_{i,t+1}), s_{i,t+1})\right] && (17)\\
& \quad s.t. && \\
d_{i,t} &\geq 0 && (18)\\
w_{i,t} &\geq d_{i,t} + k_{i,t+1} - E_t[b(s_{i,t+1})] && (19)\\
w(s_{i,t+1}) &\leq \Pi(k_{i,t+1}, s_{i,t+1}) + (1-\delta)k_{i,t+1} - R_{t+1}b(s_{i,t+1}) \qquad \forall s_{i,t+1} && (20)\\
\theta k_{i,t+1} &\leq E_t\left[M(x_{t+1})V(w(s_{i,t+1}), s_{i,t+1})\right] && (21)\\
b_{i,0} &\geq 0 && (22)
\end{aligned}
$$

The collateral constraint (21) arises from the limited enforcement problem. The lender anticipates that the borrower has limited commitment and can renege the contract. Thus, the lender is willing to lend up to the point the borrower's diversion value $\theta k_{i,t+1}$ does not exceed the borrower's continuation value $E_t\left[M(x_{t+1})V(w(s_{i,t+1}), s_{i,t+1})\right]$. This imposes an enforcement constraint on the firm's side, and makes reneging the contract never optimal.

The recursive formulation in terms of net worth not only improves the computational efficiency of the numerical solution because of the smaller state-space, but is also convenient to introduce the notion of hedging. As I discuss in more detail in Subsection 3.4, the firm has a limited borrowing capacity because of the enforcement constraint. In this formulation, the firm has the possibility to choose state-contingent promised utility (debt repayments) $b(s_{i,t+1})$ for each state.[16] The firm can therefore choose to hedge a specific state s at time $t+1$ by choosing a lower debt repayment $b(s)$. Other conditions equal, hedging a state has three effects. First, the firm saves debt capacity by relaxing the enforcement constraint. Second, as Equation (20) shows, the firm increases its net worth in state s at time $t+1$, by lowering its required repayment. As a result, more resources are available for investments and distributions in state s. Third, as Equation (19) illustrates, a lower repayment in some future state implies a lower amount of external debt raised at time t, and less net worth available for today's investment and distributions. In sum, the firm implements hedging by transferring net worth from today to specific future states tomorrow. Because the firm's debt capacity is limited by the borrowing constraint, the

[16]As Rampini and Viswanathan (2010), Rampini and Viswanathan (2013), and Nikolov, Schmid, and Steri (2013) discuss, state-contingent debt can be implemented using credit lines, forward, and futures.

company faces a tradeoff between raising funds today, and preserving them for specific states that may occur tomorrow.

3.3. Model Solution

Because the objective function itself appears on the right-hand side of the enforcement constraint, the dynamic programming problem in (17)-(22) is not a standard convex optimization problem. In particular, verifying the discounting property of Blackwell's sufficient conditions would require the knowledge of the solution to be determined. The solution of the functional equation may therefore not be unique. However, a different approach based on Knaster-Tarski fixed-point theorem allows to establish two results. First, the value function is the unique fixed point of the Bellman operator in a restricted functional space. The lower boundary of this functional space is the zero function, while the upper boundary is the solution to a planner's problem in which the enforcement constraint is removed. Second, the sequence of functions obtained by iterating the Bellman operator from the lower bound converges pointwise to such a fixed point. This leads to the following lemma:

Lemma 1 (Fixed Point) *Assume $M(x_{t+1}) = \beta M_0(x_{t+1})$, with $\beta < 1$, and*

$$\lim_{n\uparrow\infty} \beta^n E_t \left[M_0(x_{t+1})V(w(s_{i,t+1}), s_{i,t+1})\right] = 0 \tag{23}$$

Let T be the Bellman operator associated with the problem (17)- (22), $V^{UB}(w_{i,t}, s_{i,t})$ the solution of the same problem without constraint (21), and $V^{LB}(w_{i,t}, s_{i,t})$ a function over the same domain of $V(w_{i,t}, s_{i,t})$ such that $V^{LB}(w_{i,t}, s_{i,t}) \leq V(w_{i,t}, s_{i,t})$. Then:
i) The value function is the unique fixed point of T in the order interval $[V^{LB}(w_{i,t}, s_{i,t}), V^{UB}(w_{i,t}, s_{i,t})]$.
ii) The sequence of functions $\{T^n V^{LB}(w_{i,t}, s_{i,t})\}_{n=1}^{\infty}$ converges to $V(w_{i,t}, s_{i,t})$ pointwise.

The previous lemma provides an operating procedure to solve for the equilibrium contract. The solution can be obtained by value function iteration from the any initial condition $V^{LB}(w_{i,t}, s_{i,t}) \leq V(w_{i,t}, s_{i,t})$, such as the null function. Assumption (23) is guaranteed if the first-best solution $V^{UB}(w_{i,t}, s_{i,t})$ is bounded, and as long as the time-discount factor β in $M(x_{t+1})$ is less than one, and $M_0(x_{t+1})$ is finite. The last two conditions are generally guaranteed in common specifications of the stochastic discount factor.

Unfortunately, because of the large number of control variables (capital, and one debt variable for each future state), the previous iterative solution strategy is plagued by a severe curse of dimensionality, and cannot be practically implemented.[17] In particular, the maximization step is critical. For each iteration, determining the combination of control variables that maximizes the sum of distributions and the continuation value for each state would imply to search over a grid of $nk \cdot nb^{nx \cdot nz}$ points, where nk, nb, nx, and nz are respectively the number of grid points for capital, promised utility, the aggregate, and idiosyncratic shocks. To deal with this computational issue, I start from the linear programming representation of dynamic programming problems with infinite horizon (Ross (1983)). I then propose an equivalent mixed-linear programming representation of the dynamic programming problem. On this representation, I find a numerical solution by extending the constraint generation algorithm in Trick and Zin (1993).[18] Specifically, I take advantage of a separation oracle, an auxiliary linear programming problem, to deal with large state spaces and achieve computational efficiency. In Appendix B, I derive the key results on which the solution method is based, and I provide details on the implementation of the computational procedure adopted.

3.4. Optimal Policies

In this section, I characterize the optimal policy of the firms in the model through their first-order conditions.[19] The optimality conditions show how investment, financing, hedging, and payout policies are intimately related, and illustrate the qualitative mechanisms that drive firm's decisions. Because the problem has no closed-form solution, the following analysis is based on the economic interpretation of the Lagrange multipliers as shadow values.

Before introducing the optimality conditions, the numerical illustration in Figure 3

[17] As Rust (1996) discusses, a possible alternative to new computational methods for the solution of large-scale dynamic programming problems is massively parallel policy iteration. However, hardware requirements for massive parallel computation are enormous.

[18] As Denardo (1970) discusses, when discounting is present, Howard (1960) policy iteration corresponds exactly to block pivoting in the *full* equivalent linear program. Constraint generation considers sequences of smaller problems to obtain the solution.

[19] In a similar framework, Thomas and Worrall (1994) prove that the value function is differentiable. Their result extends to this model.

summarizes a few key properties of the firm's value and policy functions.[20] In Figure 3 the model is solved numerically under the baseline parametrization in Table 1. All policies, unless otherwise specified, refer to the middle state for both the aggregate and the idiosyncratic shocks. Panel A depicts firm's value as a function of current net worth. The value function is increasing and weakly concave in net worth. In particular, it is strictly concave up to a cutoff value w^C, then it becomes linear. As typical in the contracting literature, w^C defines two regions. For $w_{i,t} \geq w^C$, an additional unit of net worth translates into a one-for-one increase in equity because the total real value of the contract is not affected. If $w_{i,t} < w^C$, instead, additional net worth alters the entrepreneur incentives, and the equity value increases with a slope greater than one. Panel B shows the payout policy of the firm. The firm pays no dividends up to w^C, then the payout function is linear in net worth. Notice that the value function is strictly concave precisely in the region where no dividends are paid. Panels C and D present the investment policy $k_{i,t+1}$ and the amount of debt raised $E_t[b(s_{i,t+1})]$. From the threshold w^C onwards, the firm is reaching a "first-best" optimal level of capital. Instead, for $w_{i,t} < w^C$, the firm is constrained in its investment, because the sum of its net worth and the raised debt finance does not suffice to achieve the "first-best" capital stock. Finally, Panels E and F depict the hedging policy of the firm with respect to aggregate and idiosyncratic states. The solid lines represent the repayments the equilibrium contract specifies for the middle state, the dashed red lines refer to one state down, and the dash-dotted green lines to one state up. In general, which states the firm hedges depend on the parameter values in the model, and especially on the persistence of the autoregressive processes in Equations (16b) and (16a). Under the baseline parametrization, Panel E shows that the firm is implementing a lower repayment in the lower state, where the stochastic discount factor is high. On the contrary, Panel F shows that firms have an incentive to hedge more profitable idiosyncratic states, because of the persistence of investment opportunities over time. When aggregate states are concerned, this effect is instead dominated by the one on discount rates.

[Insert Figure 3 Here]

I now define $\mu_{i,t}$, $\nu_{i,t}$, and $\lambda_{i,t}$ as the Lagrange multipliers on the dividend non-negativity constraint (18), on the budget constraint at time t (19), and on the borrowing

[20] Some properties can be established also analytically, and are rather standard in the hedging literature. I omit them, and refer the reader to Rampini and Viswanathan (2013) for the details.

constraint (21). Denote by $\nu(s_{i,t+1})$ the marginal value of net worth in the state $s_{i,t+1}$ at time $t+1$, that is $\nu(s_{i,t+1}) \equiv V_w(w(s_{i,t+1}), s_{i,t+1})$. Notice that by the envelope condition, the marginal value of net worth at time t equals $V_w(w_{i,t}, s_{i,t})$.

The first-order condition with respect to dividends $d_{i,t}$ is:

$$\overbrace{\nu_{i,t}}^{\text{Marginal Benefit of Payouts}} = \overbrace{1+\mu_{i,t}}^{\text{Marginal Cost of Payouts}} \tag{24}$$

The payout policy of the firm balances the cost and the benefits of allocating an additional unit of current net worth to dividend distributions. The investment policy $k_{i,t+1}$ can be illustrated with the corresponding first-order condition:

$$\overbrace{E_t[M(x_{t+1})\nu(s_{i,t+1})(\Pi_k(k_{i,t+1,}s_{i,t+1}) + (1-\delta))]}^{\text{Marginal Benefit of Investment}} = \overbrace{\frac{\nu_{i,t}+\theta\lambda_{i,t}}{1+\lambda_{i,t}}}^{\text{Marginal Cost of Investment}} \tag{25}$$

The left-hand side of Equation (25) represents the marginal benefit of an additional unit of capital. Investing one unit more increases realized net worth in every future state by the return on physical capital $\Pi_k(k_{i,t+1,}s_{i,t+1}) + (1-\delta)$. The marginal benefit of investment is the expectation of these returns, accounting for the different importance of future states. Here, the effective discount factor for cash flows from invested capital is $M(x_{t+1})\nu(s_{i,t+1})$. The first component $M(x_{t+1})$ is the stochastic discount factor of the owners, while the second component $\nu(s_{i,t+1})$ relates to the concavity of the value function. The latter term is familiar in models of financial constraints. Specifically, it accounts for the different marginal value of firm's net worth across future states, and effectively renders the firm more risk averse. The right-hand side is instead the effective marginal cost of increasing the capital stock by one unit. In addition to the shadow cost $\nu(s_{i,t+1})$ of reducing net worth at time t, there are two correction terms, $\theta\lambda_{i,t}$ and $1+\lambda_{i,t}$, that reflect the presence of the borrowing constraint. Increasing investment has an effect on both sides of the enforcement constraint (21). First, it makes it more tight by increasing the diversion value of capital on the left-hand side, with a shadow value of $\theta\lambda_{i,t}$ for the firm. Second, it increases future net worth and, because the value function is increasing in it, also the continuation value on the right-hand side of (21) raises. This lowers the shadow value of investing for the firm, as the term $1+\lambda_{i,t}$ at the denominator of (25) captures.

Finally, the first-order conditions with respect to state-contingent debt $b(s_{i,t+1})$ in the contract describes the firm financing and hedging policies:

$$\overbrace{R_t \nu(s_{i,t+1}) M(x_{t+1})}^{\text{Marginal Benefit of Hedging}} = \overbrace{\frac{\nu_{i,t}}{1+\lambda_{i,t}}}^{\text{Marginal Cost of Hedging}} \tag{26}$$

Equation (26) illustrates the key tradeoff between raising less external resources today and hedging a specific future state $s_{i,t+1}$ by contracting, and implementing, a lower state-contingent repayment $b(s_{i,t+1})$. For this reason, Equation (26) highlights how financing and hedging policies are profoundly related. Specifically, the left-hand side represents the marginal benefit of hedging a specific state $s_{i,t+1}$ by reducing the corresponding repayment $b(s_{i,t+1})$, where R_t is the interest rate charged by the risk-neutral lender. As in Equation (25), the effective value of the state for the firm is $M(x_{t+1})\nu(s_{i,t+1})$. The right-hand side instead measures the cost of reduced current net worth. The shadow value of the lower amount of resources available for investment and financing is measured by $\nu_{i,t}$. The term $1+\lambda_{i,t}$ reflects a less tight borrowing constraint because of the increased continuation value, as a consequence of hedging the state $s_{i,t+1}$. In fact, a lower repayment $b(s_{i,t+1})$ increases net worth $w(s_{i,t+1})$, and in turn relaxes the borrowing constraint.

4. The Corporate CAPM

This section introduces the key asset pricing results of this paper. I first derive the stochastic discount factor in terms of firm's policies and characteristics. This leads to an asset pricing model, which I refer to as the Corporate CAPM. Finally, I discuss the aggregation properties of the asset pricing model and, in particular, a property I dub as the relativity property. The latter is an irrelevance results which states that any subset of firms in the economy can be used to back out the stochastic discount factor. Operatively, this property allows to choose different benchmark sets with respect to which stock prices and returns can be computed.

Proposition 1 (The Corporate CAPM) *i) The stochastic discount factor can be backed* *out from the firm's optimality conditions as follows:*

$$M(x_{t+1}) = \frac{1}{R_t}\frac{1}{1+\lambda_{i,t}}\frac{V_w(w_{i,t}, x_t, z_{i,t})}{V_w(w(s_{i,t+1}), x_{t+1}, z_{i,t+1})} \tag{27}$$

ii) The stochastic discount factor can be approximated as a linear function of observable firm-level variables, and quantities that are predetermined at time t, that is:

$$\log M(x_{t+1}) \approx \mu^M_{i,t} - \overline{a}_{i,t}(w(s_{i,t+1}) - w_{i,t}) - \overline{b}_{i,t}\left(\frac{\rho_{i,t+1}}{\rho^A_{t+1}} - \frac{\rho_{i,t}}{\rho^A_t}\right) - \overline{c}_{i,t}\left(\rho^A_{t+1} - \rho^A_t\right) \tag{28}$$

where $\rho_{i,t}$ and $\rho^A_{i,t}$ relate to idiosyncratic and aggregate productivity respectively:

$$\begin{aligned} \rho_{i,t} &\equiv z_{i,t} = \frac{\Pi(k_{i,t}, s_{i,t})}{k^\alpha_{i,t}} \\ \rho^A_{i,t} &\equiv x_{i,t} = \frac{\Pi^A(k_{i,t}, s_{i,t})}{\left(k^A_{i,t}\right)^\alpha} \end{aligned}$$

and $\mu^M_{i,t} \equiv \log\frac{1}{R_t} + \log\frac{1}{1+\lambda_{i,t}}$, $\overline{a}_{i,t}$, $\overline{b}_{i,t}$, and $\overline{c}_{i,t}$ are predetermined variables at time t, with

$$\mu^M_{i,t} \equiv \log\frac{1}{R_t} + \log\frac{1}{1+\lambda_{i,t}} \tag{29}$$

$$\overline{a}_{i,t} \equiv \frac{V_{ww}(w_{i,t}, s_{i,t})}{V_w(w_{i,t}, s_{i,t})} \tag{30}$$

$$\overline{b}_{i,t} \equiv \frac{V_{wz}(w_{i,t}, s_{i,t})}{V_w(w_{i,t}, s_{i,t})} \tag{31}$$

$$\overline{c}_{i,t} \equiv \frac{V_{wx}(w_{i,t}, s_{i,t})}{V_w(w_{i,t}, s_{i,t})} \tag{32}$$

The first part of the proposition obtains a stochastic discount factor from firms' decisions. Equation (27) is the counterpart of Equation (15) in the two-period example of Section 2. This result reflects the key intuition of the paper, which I develop in Section 2, and that Panel B of Figure 1 illustrates. The possibility to negotiate state-contingent debt repayments with the lenders allows firms to transfer resources across states. Firms have a rationale for hedging because of the endogenous collateral constraint, and have a motive to transfer net worth to most important states, where the stochastic discount factor is high. It is important to notice that in the absence of state-contingent debt, the stochastic discount factor cannot be recovered. This is the case in Panel A of Figure 1, in which firms cannot implement state-contingent decisions. The resulting first-order condition would not deliver a stochastic discount factor for each state, but only one equation containing an expectation over all future states, along the lines of (7).

Specifically, the stochastic discount factor relates to the firm's policy through the Lagrange multiplier $\lambda_{i,t}$ on the borrowing constraint, and the growth rate of the marginal value of net worth. The left-hand side is the stochastic discount factor, which essentially measures the value of an aggregate state for equity pricing. The right-hand side instead illustrates how the optimal decisions of heterogeneous firms adapt to the aggregate state to maximize the value for their shareholders. Backing out the stochastic discount factor therefore amounts to investigate what state must have led a firm to optimally make its observed investment and financing decisions. In the absence of state-contingent financing, realized net worth in individual future states could not instead be influenced by firm's decisions, but would vary across states only because of exogenous shocks. Firms' decisions would not therefore be informative of the stochastic discount factor.

The economic mechanism driving the result in Equation (27) relates to firms' hedging behavior. Firms have a motive to transfer resources (net worth) to states that are most important for their shareholder value. This policy would lower the marginal value of net worth in those states. However, investors' risk aversion implies that most important states are "bad times", in which marginal utility of consumption in high, and consumption is low. The term $\frac{1}{1+\lambda_{i,t}}$ accounts for firms being financially constrained. The more financially constrained they are, the higher the shadow value $\lambda_{i,t}$ of extra borrowing, the less their effective ability to transfer resources to most important states, in spite of their hedging motives. This is consistent with the models of Rampini and Viswanathan (2010), and Rampini and Viswanathan (2013), and the evidence in Rampini, Sufi, and Viswanathan (2013) and Nikolov, Schmid, and Steri (2013), according to which more constrained firms hedge less.

It is important to notice that *all* the state variables of the problem determine the policies of firms, and in turn affect their hedging abilities and the needs. From an empirical viewpoint, this result implies that firms' characteristics enter the stochastic discount factor directly. This mechanism is similar to the way, on the consumption side of the economy, the state variables of the representative household's problem enter the stochastic discount factor in the intertemporal CAPM of Merton (1973). The second part of the proposition provides an approximated linear representation of the Corporate CAPM, in terms of observable variables and quantities that are pre-determined at time t. Such an approximation delivers the following result:

Proposition 2 (Expected Return-Beta Representation) *The expected excess return on a security $E_t[R_{i,t+1} - R_t^f]$ is given by the following expression:*

$$E_t[R_{i,t+1} - R_t^f] \approx \widetilde{\lambda}_{j,t}^T \beta_{i,t} \tag{33}$$

where R_t^f is the riskfree return (or a riskfree equivalent), and the parameters $\widetilde{\lambda}_{j,t}^T$ and $\beta_{i,t}$ are given by

$$\widetilde{\lambda}_{j,t} = \left[\begin{array}{ccc} \bar{a}_{j,t} & \bar{b}_{j,t} & \bar{c}_{j,t} \end{array} \right] \sigma_{j,t}$$

$$\beta_{i,t} = \sigma_{j,t}^{-1} \left[\begin{array}{c} Cov_t(w(s_{j,t+1}) - w_{j,t}, R_{i,t+1} - R_t^f) \\ Cov_t\left(\frac{\rho_{j,t+1}}{\rho_{t+1}^A} - \frac{\rho_{j,t}}{\rho_t^A}, R_{j,t+1} - R_t^f\right) \\ Cov_t(\rho_{t+1}^A - \rho_t^A, R_{i,t+1} - R_t^f) \end{array} \right]$$

and $\sigma_{j,t}$ is the covariance matrix of $\left[\begin{array}{ccc} w(s_{j,t+1}) - w_{j,t} & \frac{\rho_{j,t+1}}{\rho_{t+1}^A} - \frac{\rho_{j,t}}{\rho_t^A} & \rho_{t+1}^A - \rho_t^A \end{array} \right]^T$.

Proposition 2 is an equivalent expected return/beta representation of the Corporate CAPM. This formulation emphasizes how expected excess equity returns are determined by the covariance with three factors: the "hedging" factor $w(s_{j,t+1}) - w_{j,t}$, and "idiosyncratic profitability" factor $\frac{\rho_{j,t+1}}{\rho_{t+1}^A} - \frac{\rho_{j,t}}{\rho_t^A}$, and the "aggregate profitability" factor $\rho_{t+1}^A - \rho_t^A$. As usual, $\beta_{i,t}$ can be interpreted as price of risk, and $\widetilde{\lambda}_{j,t}$ as quantity of risk. In the proposition, the index j refers to a benchmark firm with respect to which the factors are computed. The presence of two profitability factors denotes that in some states the j-th firm may be able to generate more resources either because all firms are more profitable (high aggregate productivity), or because it is more profitable with respect to the average (high idiosyncratic productivity). In both cases, firm's realized net worth increases in the state, and this affects the firm's hedging policy. Despite this result, in empirical tests it is convenient to aggregate firms to avoid the measurement problems that arise from separating idiosyncratic and aggregate productivity. The next proposition shows how firms can be conveniently aggregated to implement empirical tests of the model.

Proposition 3 (Aggregation) *Consider an arbitrary subset Ω of N firms in the cross-section.*

i) The expression of the stochastic discount factor in Equation (28) and its covariance representation can be restated in terms of averages across firms in Ω as follows:

$$\log M(x_{t+1}) \approx \frac{1}{N}\sum_{j\in\Omega}\left[\log \mu_{j,t}^M - \overline{a}_{j,t}(w(s_{j,t+1}) - w_{j,t}) - \overline{b}_{j,t}\left(\frac{\rho_{j,t+1}}{\rho_{t+1}^A} - \frac{\rho_{j,t}}{\rho_t^A}\right) - \overline{c}_{j,t}(\rho_{t+1}^A - \rho_t^A)\right] \tag{34}$$

and

$$E_t[R_{i,t+1} - R_t^f] \approx \widetilde{\lambda}_{j,t}^T \beta_{j,t} \tag{35}$$

with

$$\widetilde{\lambda}_{j,t} = -\sigma_{\Omega t}$$

$$\beta_{i,t} = \sigma_{\Omega t}^{-1}\begin{bmatrix} Cov_t\left(\frac{1}{N}\sum_{j\in\Omega} -\overline{a}_{j,t}(w(s_{j,t+1}) - w_{j,t}), R_{i,t+1} - R_t^f\right) \\ Cov_t\left(\frac{1}{N}\sum_{j\in\Omega} -\overline{b}_{j,t}\left(\frac{\rho_{j,t+1}}{\rho_{t+1}^A} - \frac{\rho_{j,t}}{\rho_t^A}\right), R_{i,t+1} - R_t^f\right) \\ Cov_t\left(\frac{1}{N}\sum_{j\in\Omega} -\overline{c}_{j,t}\left(\rho_{t+1}^A - \rho_t^A\right), R_{i,t+1} - R_t^f\right) \end{bmatrix}$$

and $\sigma_{\Omega t}$ is the covariance matrix of $\begin{bmatrix} \frac{1}{N}\sum_{j\in\Omega}\overline{a}_{j,t}(w(s_{j,t+1}) - w_{j,t}) & \frac{1}{N}\sum_{j\in\Omega}\overline{b}_{j,t}\left(\frac{\rho_{j,t+1}}{\rho_{t+1}^A} - \frac{\rho_{j,t}}{\rho_t^A}\right) & \frac{1}{N}\sum_{j\in\Omega}\overline{c}_{j,t}\left(\rho_{t+1}^A - \rho_t^A\right) \end{bmatrix}^T$

ii) If $N \to \infty$, then:

$$\log M(x_{t+1}) \approx \frac{1}{N}\sum_{j\in\Omega}\left[\log \mu_{j,t}^M - \overline{a}_{j,t}(w(s_{j,t+1}) - w_{j,t}) - \overline{c}_{j,t}(\rho_{t+1}^A - \rho_t^A)\right] \tag{36}$$

with the following expected return/beta representation

$$\beta_{i,t} = \sigma_{\Omega t}^{-1}\begin{bmatrix} Cov_t\left(\frac{1}{N}\sum_{j\in\Omega} -\overline{a}_{j,t}(w(s_{j,t+1}) - w_{j,t}), R_{i,t+1} - R_t^f\right) \\ Cov_t\left(\frac{1}{N}\sum_{j\in\Omega} -\overline{c}_{j,t}\left(\rho_{t+1}^A - \rho_t^A\right), R_{i,t+1} - R_t^f\right) \end{bmatrix}$$

and $\sigma_{\Omega t}$ is the covariance matrix of $\begin{bmatrix} \frac{1}{N}\sum_{j\in\Omega}\overline{a}_{j,t}(w(s_{j,t+1}) - w_{j,t}) & \frac{1}{N}\sum_{j\in\Omega}\overline{c}_{j,t}\left(\rho_{t+1}^A - \rho_t^A\right) \end{bmatrix}^T$

The first part of the proposition provides a theoretical irrelevance result, to which I refer as the relativity property. This can be illustrated as follows. In the model, all firms maximize the value for the owner, in that they use the same the stochastic discount factor to discount expected future profits. In the model, the left-hand side of Equation (34) is therefore constant, and the stochastic discount factor can be backed out by averaging out the right-hand side for any subset of firms Ω in the economy (e.g. an industry).[21] The property essentially states that the reference set of firms with respect to which expected returns are evaluated can be arbitrarily chosen. This differs from macro models with a representative agent, that dictate that factors must necessarily be aggregated quantities. The second part of the proposition provides an aggregation result when the number of firms used to construct the stochastic discount factor is large enough. In this case, the idiosyncratic productivity factor zeros out because of averaging out a large number of firms in the cross-section. This result is useful in implementing empirical tests of the model. Using individual firms to back out the stochastic discount factor can be problematic for two main reasons. First, as in any economic model, there are omitted forces that can affect individual firms much more than sample averages, such as product market competition, labor market frictions, or investment adjustment costs. Second, testing the model in a small sample of firms would pose the challenge of measuring and disentangling aggregate and idiosyncratic productivities. Such a task would lead to technical difficulties, and is subject to misspecification errors, as discussed for example in Burnside, Eichenbaum, and Rebelo (1996) and Ábrahám and White (2006).

5. Quantitative Analysis

I resort to calibration to evaluate the quantitative ability of the model to rationalize firm's observed policies. Calibration restricts some structural parameter values to replicate some key quantities in the data. Ideally, a one-to-one mapping between parameters and moments provides a sufficient condition for identification. Such a close mapping is hard to accomplish in any economic model, because firm's investment and financing decisions are intertwined, and the model parameters affect all the data moments.

To identify the key parameters in the model, I break them down into three groups. The

[21]In case Ω is a weighted portfolio of firms, all the sample averages are replaced by weighted averages.

first group includes parameters whose value can be restricted from existing quantitative works or mapped directly into data moments. The second group refers to parameters that can be identified using some aggregate asset pricing moments. The third group includes parameters that I set to obtain a match between the simulated data moments from the model, and the actual data moments. Panel C of Table 1 reports parameter values, while Panel A and B respectively show simulated and actual moments that pertain to corporate policies, and to aggregate asset pricing quantities. All data are described in Supplementary Appendix D.

[Insert Table 1 Here]

In the numerical solution of the model, I follow the recent literature on cross-sectional asset pricing and specify an exogenous process for the stochastic discount factor (Berk, Green, and Naik (1999), Zhang (2005), Gomes and Schmid (2010)). Since the goal of this section is to provide evidence that the model is quantitatively successful on the corporate side for a sensible choice of a pricing kernel, this strategy seems reasonable. All calibrations are based on annual data, consistent with the quantitative corporate finance literature. I follow Zhang (2005) and I specify the pricing kernel as follows:

$$\log M(x_{t+1}) = \log \beta + [(\gamma_0 + \gamma_1(x_t - \mu_x)](x_{t+1} - x_t) \tag{37}$$

where β, $\gamma_0 > 0$, and $\gamma_1 < 0$ are constant parameters.[22]

The parameters that pertain to the first group are the depreciation rate δ, the persistence ρ_x and the volatility σ_x of the aggregate shock process, and the exit rate ϕ. The depreciation rate is set to 0.15, to approximately match the depreciation rate for US listed firms in my sample. This value is the same used in Hennessy and Whited (2007), and DeAngelo, DeAngelo, and Whited (2011). ρ_x and σ_x are set to 0.95^4 and $0.007 \cdot 4$ to correspond, on an annual frequency and with the autoregressive specification in (16b), to the quarterly values of 0.95 and 0.007 in Cooley and Prescott (1995). As in Gomes and Schmid (2010), I set the fraction of incumbents ϕ to 0.02, in line with the study of Covas and Den Haan (2012).

[22]For an in-depth discussion of this assumption and of the properties of the pricing kernel see Berk, Green, and Naik (1999) and Zhang (2005).

The second set of parameters consists of those in the stochastic discount factor, β, γ_0, and γ_1. I pin down their value, using the strategy in Zhang (2005), to match three aggregate moments: the mean and volatility of the real interest rate, and the average Sharpe ratio. The parametrization in Equation (37) for the pricing kernel is convenient in that the real interest rate R_t^f and the maximum Sharpe ratio S_t are:

$$R_t^f = \beta^{-1} e^{-(\mu_m + \frac{1}{2}\sigma_m^2)} \quad (38)$$

$$S_t = \frac{\sqrt{e^{\sigma_m^2}(e^{\sigma_m^2} - 1)}}{e^{\frac{\sigma_m^2}{2}}} \quad (39)$$

with

$$\mu_m = [(\gamma_0 + \gamma_1(x_t - \mu_x)](x_t - \mu_x)(1 - \rho_x) \quad (40)$$

$$\sigma_m = [(\gamma_0 + \gamma_1(x_t - \mu_x)]\sigma_x \quad (41)$$

This strategy yields $\beta = 0.94$, $\gamma_0 = 12.5$, and $\gamma_1 = -120$, and gives a real interest rate of 2.99% per year, an annual interest rate volatility of 3.75%, and a Sharpe ratio of 0.35. These values are close to the corresponding data moments of 2.2%, 4.35%, and 0.41.

Finally, I pick 13 moments to match the remaining 5 parameters in the third group. I roughly categorize these moments as representing firm's investment, financing, and equity returns. On the investment side, I choose moments that relate to operating income, investment, and Tobin's Q. On the financing side, I consider the mean, variance, and serial correlation of leverage. On the asset pricing side, I pick the mean and the average of market excess return, and the average volatility of individual stock returns. The resulting parameter values appear to be reasonable. The curvature α is 0.76, in the range of values reported by Hennessy and Whited (2005), Hennessy and Whited (2007), and DeAngelo, DeAngelo, and Whited (2011) on annual data. The persistence and volatility ρ_z and σ_z of idiosyncratic productivity shocks are within one standard error of the estimates in Hennessy and Whited (2007), in which there are no capital adjustment costs as in the present framework. The parameter μ_x is a scale parameter, that determines the scale of the simulated economy and the steady-state capital stock. Finally, there is little guidance for the value of θ, which represents the fraction of capital that the entrepreneur effectively diverts in the case of liquidation. I set $\theta = 0.3$, which is in line with values of related quantities in existing models, such as DeAngelo, DeAngelo, and Whited (2011) and Nikolov, Schmid, and Steri (2013).

Panels A and B of Table 1 show that the model is broadly successful in matching both aggregate asset pricing moments, and moments that relate to corporate investment and financing. The model performance may further improve by adding other frictions and considering additional moments. However, the absence of these frictions like capital adjustment costs and fixed operating costs considerably simplifies the analysis. Because the focus of this work is to derive a stochastic discount factor from an optimal contracting framework, I privilege model parsimony over an improvement of the quantitative fit of the model.

[Insert Table 1 Here]

6. The Corporate CAPM: Empirical Evaluation

In this section, I test the implications of the Corporate CAPM in the data. Because the focus of this work is on differences in risk premia across assets, I examine the implications of the model for cross-sectional expected excess returns. To do so, I test the following restrictions on the pricing errors of a vector of excess returns R^e_{t+1}:

$$E_t[M(x_{t+1}R^e_{t+1})] = 0 \tag{42}$$

where $M(x_{t+1})$ is defined in Equation (36). The model with excess returns does not identify the intercept $\mu^M_{i,t}$ of the stochastic discount factor in Proposition 3. The intercept is in fact predetermined at time t, and can be normalized in empirical tests (Cochrane (2001), Yogo (2006), Belo (2010)). I implement empirical tests by GMM using yearly data from 1965 to 2010. Estimation is by two-step GMM, with the initial weighting matrix attaching equal weights to all assets. Supplementary Appendix E provides details on the estimation procedure, and replicates the empirical tests with an alternative measure of the productivity factor based on Fernald (2009). The latter analysis controls for possible misspecifications in measuring aggregate productivity ρ^A_t as a Solow residual, as discussed by Burnside, Eichenbaum, and Rebelo (1996). All data are described in Supplementary Appendix D.

The test assets are: (i) the 25 Fama-French portfolios sorted by size and book-to-market equity, (ii) the 30 Fama-French industry portfolios, (iii) 25 portfolios sorted by market and HML beta, and (iv) all the previous portfolios together. The 25 Fama-French portfolios are chosen because they capture the value and the size premia, which have received considerable attention in the literature. As in Lewellen, Nagel, and Shanken (2010), I include the 30 Fama-French industry portfolios to relax the tight factor structure of the 25 Fama-French portfolio. At Lewellen, Nagel, and Shanken (2010) document, the 30 industry portfolios represent a challenging test for all leading asset pricing models. Following Yogo (2006), I also include the beta-sorted portfolios, in order to address the critique in Daniel and Titman (2012).

As Equation (36) shows, if the number of firms with respect to which the factors are computed is large enough, the Corporate CAPM reduces to a two-factor conditional model. In other words, the coefficients $\overline{a}_{i,t}$ and $\overline{c}_{i,t}$ are time varying and depend on firms' characteristics. In the next two subsections, I therefore implement both unconditional and conditional tests. Unconditional tests treat $\overline{a}_{i,t}$ and $\overline{c}_{i,t}$ as constant parameters. Unconditional tests are reported for comparability with previous studies. In conditional tests, I instead use a model-based identification strategy. More precisely, I use the quantitative policy function of the model from Section 5 to find a parsimonious functional form for the time-varying coefficients in terms of constant parameters and observable variables. As aggregation properties in Proposition 3 illustrate, in order to implement empirical tests a level of aggregation must be specified. For comparability with previous studies that use aggregate data, in both Subsections 6.1 and 6.2 I aggregate data at the market level. In Subsection 6.4 I instead carry out empirical tests using the five Fama-French industries (consumer goods, manufacturing, hi-tech, healthcare, other) as references.

6.1. Unconditional Tests

If $\overline{a}_{i,t}$ and $\overline{c}_{i,t}$ are constant terms, Proposition 3 leads to a two factor model where the net worth and profitability factors are averaged across all firms. Table 2 presents the estimation results. Coefficient estimates for the two factors and the corresponding HAC standard errors are reported. The table also reports the following goodness-of-fit measures based on first-stage inference: the mean absolute pricing error (MAE), and the cross-sectional R^2 of a regression of realized average excess returns on predicted average excess returns, computed as in Campbell and Vuolteenaho (2004). As a measure of model mis-specification I report the Hansen-Jagannathan (HJ) distance (Hansen and Jagannathan (1997)). The HJ-distance can be interpreted as the minimum distance between the proposed stochastic discount factor and the set of correct stochastic discount factors for a given set of test assets. Finally, the table includes two formal tests of the model: the J-test of overidentifying restrictions (Hansen and Singleton (1982)), and a test of the null hypothesis of zero HJ-distance (Jagannathan and Wang (1996)). Although several studies[23] document the statistical power of both tests is low in the context of asset pricing tests, and their small-sample properties vary to a great extent with the sample size and the test assets, I report them for comparability with previous studies.

Unconditional tests suggest that the Corporate CAPM finds support in the data. The first two rows of Table 2 report GMM estimates of the coefficients on the net worth and profitability factor for all the test assets. Although conditional tests are a more appropriate setting to discuss the sign restrictions on the coefficients, the unconditional estimates are overall in line with the predicted signs for $\overline{a}_{i,t}$ and $\overline{c}_{i,t}$ from the model. Column 4 shows that when all test assets are considered, the coefficients on net worth and profitability factors have a negative and a positive sign respectively, as the model predicts. As Columns 1-3 show, the coefficient on the profitability factor is positive even for all test assets individually. In addition, while the 25 portfolios sorted by size and book-to-market and the risk-sorted portfolios do not individually lead to statistically significant estimates of the coefficients for the net worth factor, the estimates for the 30 industry portfolios clearly identify a negative coefficient. Such a negative coefficient remains significant when all portfolios are considered together, with a point estimate of -6.334, more than four standard errors from zero. This result supports the recommendation in Lewellen, Nagel, and Shanken (2010) to include the Fama-French industry portfolios in tests of asset pricing models.

The Corporate CAPM appears to capture most of the variation in expected returns across the test assets. Mean absolute pricing errors range from 0.676% to 0.838% per annum. Cross-sectional R^2 are also high, ranging from 0.771 for the industry portfolios, to 0.923 for the 25 size/book-to-market portfolios. Remarkably, the model is successful

[23] See, for example, Ferson and Foerster (1994), Ahn and Gadarowski (2004), and Lewellen, Nagel, and Shanken (2010).

in pricing the Fama-French 30 Industry portfolios. In fact, as Lewellen, Nagel, and Shanken (2010) document, these test assets represent a challenge for all leading asset pricing models. Finally, although the results of formal tests should be interpreted with extreme caution for the reasons above, both the tests based on the HJ distance and the J statistic cannot statistically reject the model.[24]

[Insert Table 2 Here]

Figure 4 provides a visual summary of the performance of the model. Panels A through D report predicted versus realized average returns for the four sets of test assets. If priced correctly, the portfolio should lie along the 45-degree line. The figure clearly shows that the pricing performance of the Corporate CAPM is more than satisfactory.

[Insert Figure 4 Here]

6.2. Conditional Tests

In this section I implement conditional tests of the Corporate CAPM. Because the change in aggregate profitability does not vary across firms, Equation (36) leads to the following specification for the stochastic discount factor:

$$\log M(x_{t+1}) \approx \mu^M - \frac{1}{N} \sum_{j \in \Omega} \left[\overline{a}_{j,t} (w(s_{j,t+1}) - w_{j,t}) \right] - \overline{c}_t (\rho^A_{t+1} - \rho^A_t) \tag{43}$$

where $\overline{c}_t \equiv \frac{1}{N} \sum_{j \in \Omega} \overline{c}_{j,t}$.

As the theoretical argument in Hansen and Richard (1987) remarks, testing conditional models is conceptually difficult because they inherently depend on the information structure of the agents in the economy. In empirical work, the most common testing strategy is to specify the conditional parts of the model as linear functions of some set of observable variables, such as the default and term spreads, the consumption-to-wealth ratio, and the aggregate dividend yield. Other approaches make use of higher frequency data, such as the MIDAS techniques in Ghysels, Santa-Clara, and Valkanov (2004).

[24]The values of the HJ distance for the case of all portfolios together is not reported because, as Cochrane (1996) discusses, the cross-moment matrix of returns is nearly singular when the number of test assets is large.

In the implementation of conditional tests, I use the policy function of the model to specify a parsimonious functional form for $\overline{a}_{j,t}$ and $\overline{c}_t$. I adopt a model-based identification strategy for three reasons. First, the annual data frequency of my sample is not well-suited to implement methods that take advantage of high frequency data. Second, the coefficients $\overline{a}_{j,t}$ and $\overline{c}_t$ depend on the state variables of the model, rather than on the observable variables usually considered in conditional tests based on macroeconomic factors. Third, as Brandt and Chapman (2006) discuss, a linear approximation for the functional forms of the coefficients in the model may result in large misspecifications. Admittedly, the information set investors access in the real world is larger than the state variables of the contracting model. However, as Hansen and Richard (1987) show, by the law of iterated expectations a conditional model can be tested by "conditioning down" finer information sets to coarser ones.[25]

Panels C and F of Figure 5 plot the building blocks for the conditional tests in this section, namely the coefficients $\overline{a}_{i,t}$ and $\overline{c}_{i,t}$ for the firm i. The coefficient $\overline{a}_{i,t}$ is negative and increasing in current net worth, and its graph is highly nonlinear, especially for firms with low net worth. The negative sign of $\overline{a}_{i,t}$ follows directly from the shape of the value function. Panels A and B depict respectively the denominator and the numerator of $\overline{a}_{i,t}$, as defined in Proposition 1. The graph in Panel A is the marginal value of net worth, which is positive because the value function is increasing in net worth. The graph in Panel B is its derivative with respect to net worth, which is negative because of the concavity of the value function. Analogously, the coefficient $\overline{c}_{i,t}$ is approximately linear and decreasing in the current aggregate shock x_t which, as Proposition 1 shows, can be measured in the data as the Solow residual ρ_t^A. Panels D and E depict the denominator and the numerator of $\overline{c}_{i,t}$ under the baseline calibration in Table 1.

[Insert Figure 5 Here]

To carry on conditional tests, I look for an approximation of $\overline{a}_{i,t}$ and $\overline{c}_t$ in terms of observable variables. To do so, I run regressions on the model solution to identify a functional form for $\overline{a}_{i,t}$ and $\overline{c}_t$ in terms of net worth $w_{i,t}$ and ρ_t^A. While both coefficients in principle depend on all the state variables of the model, my goal is to find a parsimonious

[25] As Cochrane (2001) points out, all the moments computed with respect to the coarser information set must exist.

functional form for them, which possibly involves only a subset of the state variables. Table 3 reports the estimates for a nonlinear regression of $\overline{a}_{i,t}$ on the function $a_0 \frac{1}{1+a_1 w_{i,t}}$, where a_0 and a_1 are constant parameters, and the estimates for a linear regression of $\overline{c}_t$ on ρ_t^A. The nonlinear regression is implemented with the algorithm in Levenberg (1944) and Marquardt (1963), as described in the caption of the table. While the model as no closed-form solution, the approximations for both coefficient delivers a good fit, with R^2 statistics of 0.969 and 0.999 respectively. The regressions produce estimates of -35.424, 7.489, 4.142, and -17.623 for a_0, a_1, the intercept c_0, and the slope c_1. Given the limited number of observations on an annual frequency, to avoid overfitting and noisy estimates in the GMM tests of the model, I only estimate a_0 and c_0, while I set a_1 and c_1 to the values reported above.

[Insert Table 3 Here]

Table 4 reports the results for the estimation. The results are consistent with those of the unconditional tests in Table 2. The estimates of a_0 and c_0 have the expected sign when all test assets are considered in Column 4. The estimates in Columns 1-3 confirm that, as in unconditional tests, the Fama-French 30 industry portfolios play an important role in the inference. Finally, the Corporate CAPM appears to have a good pricing performance, with mean absolute pricing errors below 0.8% per year, and R^2 statistics well above 0.8.

[Insert Table 4 Here]

6.3. Comparison Among Models

Table 5 compares the pricing performance of the Corporate CAPM and that of the most popular existing asset pricing models. I consider three other models: the CAPM (Column 1), the Fama and French three-factor model (Column 2), and the Consumption CAPM (Column 3). Columns 4 and 5 report the results for both unconditional and conditional tests of the Corporate CAPM. In terms of test assets, Panel A refers to the Fama-French 25 portfolios, Panel B to the 25 portfolios sorted by HML and market beta, Panel C to the 30 Fama-French industry portfolios, and Panel D to all portfolios together.

[Insert Table 5 Here]

As in previous studies, the CAPM and the Consumption CAPM are not successful in pricing the tests assets. The MAE is high, ranging from 1.362% per annum to 1.911% per annum, and the R^2 is consistently low across all test assets. The Fama-French model instead performs rather well, with mean absolute pricing errors ranging from 0.673% to 1.095% per year, and R^2 between 0.630 for the 30 Fama-French portfolios and 0.915 for the portfolios sorted by size and book-to-market. With respect to these two indicators, the Corporate CAPM outperforms all models on all test assets, both in its unconditional and conditional specification. Not surprisingly, and consistent with the findings in Ahn and Gadarowski (2004), Burnside (2010), Lewellen, Nagel, and Shanken (2010), and Daniel and Titman (2012), the formal tests based on HJ and J statistics are uninformative, and are unable to reject any model. Although these findings should be interpreted with caution due to the well-known issues with the testing framework, the Corporate CAPM seems to have a satisfactory pricing performance.

[Insert Figure 6 Here]

Figure 6 summarizes the previous comparison among models, in line with Figure 4. Panels A through D depict predicted versus realized average excess returns for the CAPM, the Fama-French model, the Consumption CAPM, and the Corporate CAPM. The figure refers to all the test assets together. Panels A and C show that the points are far from the 45-degree line for the CAPM and the Consumption CAPM, while they line up fairly well for the Fama and French's model (Panel B), and especially for the Corporate CAPM (Panel D).

6.4. Industry Breakdowns

As I discuss in Section 4, Proposition 3 provides an irrelevance result that I dub as the relativity property. In the model, as long as the number of firms used in the aggregation process is large, any choice of the set of benchmark firms for the computation of the factors allows to back out the same approximate stochastic discount factor.

Table 6 reports unconditional (Panel A) and conditional (Panel B) tests of the Corporate CAPM with respect to five large reference industry, namely the Fama-French industries (consumer goods, manufacturing, hi-tech, healthcare, other). The test assets are all the previous portfolios together. The results appear to be consistent with the relativity property. Regardless of the reference industry, mean absolute pricing errors are rather low, with R^2 statistics between 0.720 to 0.854. In addition, the estimates for the coefficients on the net worth and profitability factors are respectively negative and positive as predicted by the model.

These results represent a starting point to understand common procedures that focus on "comparable" firms, and that practitioners ordinarily use for company valuation, such as relative valuation based on multiples or bottom-up betas (Damodaran (2008)). In fact, unlike classical macro-based asset pricing models, the present framework allows to formally introduce the concept of benchmark set of firms. Future research may extend the present model to analyze the conditions under which the irrelevance result breaks, and attempt to rationalize such commonly used practices.

[Insert Table 6 Here]

7. Conclusions

Recent corporate finance studies document that hedging motives represent a key determinant of corporate decisions. In a dynamic contracting model, I recover a stochastic discount factor from firm's investment and financing policies. This leads to a novel asset pricing model, the Corporate CAPM. In the model, firms hedge by transferring resources to future states where they are more valuable. Firms have limited funds because of collateral constraints that endogenously arise from agency conflicts between firms and lenders. The amount of resources firms can devote to hedging is therefore limited. In this context, the shareholders' stochastic discount factor measures the importance of each state for firm's value. Value maximization provides a motive for firms to hedge most important states, in a tradeoff with their funding needs for current investment and distributions. On the corporate finance side, a calibrated version of the model is quantitatively consistent with investment, financing, and payout policies of US listed firms. On the asset pricing side, the Corporate CAPM finds support in the data. The model performs well in pricing different test assets, also in comparison to popular asset pricing models, namely the CAPM, the Consumption CAPM, and the Fama and French three-factor model.

This work has implications for future research not only for production-based asset pricing, but also for consumption-based models, and for empirical work on the cross-section of expected returns. The present framework may represent a complementary tool to advance the understanding of the consumption side of the economy. As Cochrane (2011) points out, the ultimate goal of asset pricing theory should be to provide a general equilibrium explanation of how asset returns and consumption are jointly determined. In general equilibrium, the stochastic discount factor obtained from both the production and consumption side of the economy must have consistent properties. These additional restrictions may provide guidance in modeling the household side on the economy. Another implication of this paper is that the state variables of the firm's optimization problem, in other words the determinants of firms' decisions, enter the stochastic discount factor directly. For empirical work, this observation may provide insights for the development of new testable hypotheses for cross-sectional differences in returns.

References

Ábrahám, Á., and K. White, 2006, The dynamics of plant-level productivity in us manufacturing, *Center for Economic Studies Working Paper*, 6, 20.

Abreu, D., D. Pearce, and E. Stacchetti, 1990, Toward a theory of discounted repeated games with imperfect monitoring, *Econometrica*, 58(5), 1041–1063.

Ahn, S. C., and C. Gadarowski, 2004, Small sample properties of the GMM specification test based on the Hansen–Jagannathan distance, *Journal of Empirical Finance*, 11(1), 109–132.

Albuquerque, R., and H. A. Hopenhayn, 2004, Optimal lending contracts and firm dynamics, *The Review of Economic Studies*, 71(2), 285–315.

Aliprantis, C. D., and K. C. Border, 2006, *Infinite dimensional analysis: a hitchhiker's guide*. Springer.

Bazdrech, S., F. Belo, and X. Lin, 2013, Labor hiring, investment and stock return predictability in the cross section, *Journal of Political Economy*, forthcoming.

Belo, F., 2010, Production-based measures of risk for asset pricing, *Journal of Monetary Economics*, 57(2), 146–163.

Berk, J. B., R. C. Green, and V. Naik, 1999, Does the stock market overreact?, *The Journal of Finance*, 54(5), 1513–1607.

Bolton, P., H. Chen, and N. Wang, 2011, A unified theory of Tobin's Q, corporate investment, financing, and risk management, *The Journal of Finance*, 66(5), 1545–1578.

Brandt, M. W., and D. A. Chapman, 2006, Linear approximations and tests of conditional pricing models, Working Paper.

Burnside, C., 2010, Identification and inference in linear stochastic discount factor models, Working Paper.

Burnside, C. A., M. S. Eichenbaum, and S. T. Rebelo, 1996, Sectoral Solow residuals, *European Economic Review*, 40(3), 861–869.

Campbell, J. Y., and T. Vuolteenaho, 2004, Bad beta, good beta, *American Economic Review*, pp. 1249–1275.

Cochrane, J. H., 1993, Rethinking production under uncertainty, *Manuscript, University of Chicago.*

——— , 1996, A cross-sectional test of an investment-based asset pricing model, *Journal of Political Economy*, 104(3), 572–621.

——— , 2001, *Asset Pricing*. Princeton University Press, Princeton, NJ.

——— , 2011, Presidential address: Discount rates, *The Journal of Finance*, 66(4), 1047–1108.

Colla, P., F. Ippolito, and K. Li, 2013, Debt specialization, *The Journal of Finance*, forthcoming.

Cook, W. J., W. H. Cunningham, W. R. Pulleyblank, and A. Schrijver, 2011, *Combinatorial optimization*, vol. 33. Wiley. com.

Cooley, T. F., and E. C. Prescott, 1995, Economic growth and business cycles, *Frontiers of business cycle research*, pp. 1–38.

Covas, F., and W. J. Den Haan, 2012, The role of debt and equity finance over the business cycle, *The Economic Journal*, 122(565), 1262–1286.

Damodaran, A., 2008, *Damodaran on valuation*. Wiley. com.

Daniel, K., and S. Titman, 2012, Testing factor-model explanations of market anomalies, *Critical Finance Review*, 1(1), 103–139.

DeAngelo, H., L. DeAngelo, and T. M. Whited, 2011, Capital structure dynamics and transitory debt, *Journal of Financial Economics*, 99(2), 235–261.

Denardo, E. V., 1970, On linear programming in a Markov decision problem, *Management Science*, pp. 281–288.

Fernald, J., 2009, A quarterly, utilization-adjusted series on total factor productivity, Working Paper.

Ferson, W. E., and S. R. Foerster, 1994, Finite sample properties of the Generalized Method of Moments in tests of conditional asset pricing models, *Journal of Financial Economics*, 36(1), 29–55.

Garlappi, L., and H. Yan, 2011, Financial distress and the cross-section of equity returns, *The Journal of Finance*, 66(3), 789–822.

Ghysels, E., P. Santa-Clara, and R. Valkanov, 2004, The MIDAS touch: mixed data sampling regression models, Working Paper.

Gomes, J. F., and L. Schmid, 2010, Levered returns, *The Journal of Finance*, 65(2), 467–494.

Hansen, L. P., and R. Jagannathan, 1997, Assessing specification errors in stochastic discount factor models, *The Journal of Finance*, 52(2), 557–590.

Hansen, L. P., and S. F. Richard, 1987, The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models, *Econometrica*, pp. 587–613.

Hansen, L. P., and K. J. Singleton, 1982, Generalized instrumental variables estimation of nonlinear rational expectations models, *Econometrica: Journal of the Econometric Society*, pp. 1269–1286.

Hennessy, C. A., and T. M. Whited, 2005, Debt dynamics, *The Journal of Finance*, 60(3), 1129–1165.

——— , 2007, How costly is external financing? Evidence from a structural estimation, *The Journal of Finance*, 62(4), 1705–1745.

Howard, R. A., 1960, *Dynamic Programming and Markov Processes*. John Wiley.

Jagannathan, R., and Z. Wang, 1996, The conditional CAPM and the cross-section of expected returns, *The Journal of Finance*, 51(1), 3–53.

Jermann, U. J., 2010, The equity premium implied by production, *Journal of Financial Economics*, 98(2), 279–296.

Kamihigashi, T., 2012, Elementary results on solutions to the Bellman equation of dynamic programming: Existence, uniqueness, and convergence, working paper.

Levenberg, K., 1944, A method for the solution of certain problems in least squares, *Quarterly of Applied Mathematics*, 2, 164–168.

Lewellen, J., S. Nagel, and J. Shanken, 2010, A skeptical appraisal of asset pricing tests, *Journal of Financial Economics*, 96(2), 175–194.

Li, S., and T. Whited, 2013, Endogenous financial constraints, taxes, and leverage, Working Paper.

Livdan, D., H. Sapriza, and L. Zhang, 2009, Financially constrained stock returns, *The Journal of Finance*, 64(4), 1827–1862.

Marquardt, D. W., 1963, An algorithm for least-squares estimation of nonlinear parameters, *Journal of the Society for Industrial and Applied Mathematics*, 11(2), 431–441.

McGrattan, E. R., 1997, *Application of weighted residual methods to dynamic economic models*. Federal Reserve Bank of Minneapolis, Research Department.

Mendoza, E. G., 2000, Sudden stops, financial crises, and leverage, *American Economic Review*, 100(5), 1941–1966.

Merton, R. C., 1973, An intertemporal capital asset pricing model, *Econometrica: Journal of the Econometric Society*, pp. 867–887.

Nemhauser, G. L., and L. A. Wolsey, 1988, *Integer and combinatorial optimization*, vol. 18. Wiley New York.

Nikolov, B., L. M. Schmid, and R. Steri, 2013, Dynamic corporate liquidity, Working Paper.

Obreja, I., 2013, Book-to-market, financial leverage, and the cross-section of stock returns, *Review of Financial Studies*, forthcoming.

Pruitt, W. E., 1966, Summability of independent random variables, *Journal of Mathematical and Mechanics*, 15, 769–776.

Pucci de Farias, D., and B. Van Roy, 2003, The linear programming approach to approximate dynamic programming, *Operations Research*, 51(6), 850–865.

Rabin, M., 2000, Risk aversion and expected-utility theory: A calibration theorem, *Econometrica*, 68(5), 1281–1292.

Rampini, A. A., A. Sufi, and S. Viswanathan, 2013, Dynamic risk management, *Journal of Financial Economics*, forthcoming.

Rampini, A. A., and S. Viswanathan, 2010, Collateral, risk management, and the distribution of debt capacity, *The Journal of Finance*, 65(6), 2293–2322.

——— , 2013, Collateral and capital structure, *Journal of Financial Economics*, 109(2), 466–492.

Rohatgi, V., 1971, Convergence of weighted sums of independent random variables, in *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 69, pp. 305–307. Cambridge Univ Press.

Ross, S. M., 1983, *Recursive Methods in Economic Dynamics*. Academic Press, New York, NY.

Rust, J., 1996, Numerical dynamic programming in economics, *Handbook of computational economics*, 1, 619–729.

Schmid, L. M., and R. Steri, 2013, A quantitative dynamic agency model of financing constraints, Working Paper.

Schrijver, A., 1998, *Theory of linear and integer programming*. Wiley. com.

Stokey, N. L., and R. E. Lucas, 1989, *Recursive Methods in Economic Dynamics*. Harvard University Press, Cambridge, MA.

Sufi, A., 2009, Bank lines of credit in corporate finance: An empirical analysis, *Review of Financial Studies*, 22(3), 1057–1088.

Tarski, A., 1955, A lattice-theoretical fixpoint theorem and its applications, *Pacific journal of Mathematics*, 5(2), 285–309.

Thomas, J., and T. Worrall, 1994, Foreign direct investment and the risk of expropriation, *The Review of Economic Studies*, 61(1), 81–108.

Trick, M. A., and S. E. Zin, 1993, A linear programming approach to solving stochastic dynamic programs, Working Paper.

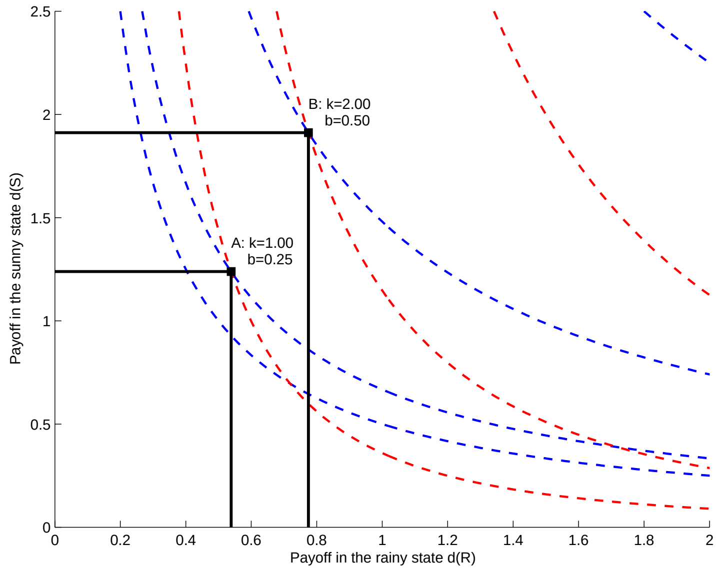
———, 1997, Spline approximations to value functions: A linear programming approach, *Macroeconomic Dynamics*, 1(1), 255–277.

Yogo, M., 2006, A consumption-based explanation of expected stock returns, *The Journal of Finance*, 61(2), 539–580.

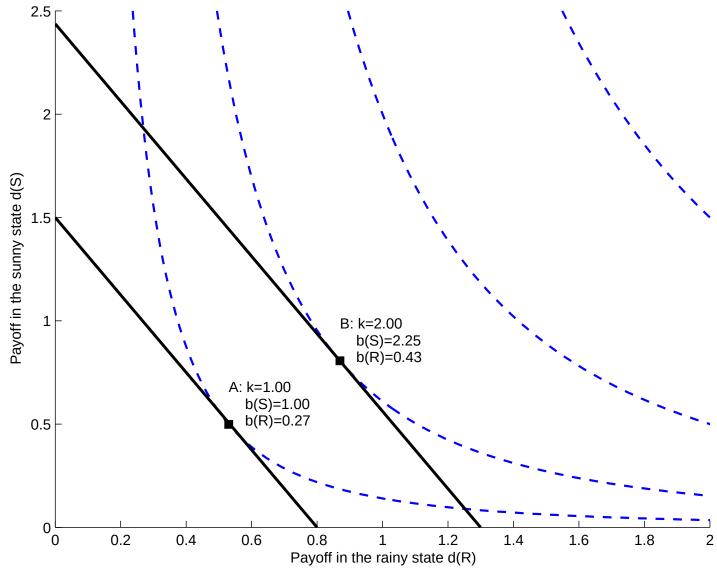
Zhang, L., 2005, The value premium, *The Journal of Finance*, 60(1), 67–103.

Figure 1. Collateral-Based Asset Pricing: Illustration

Panel A: No Hedging

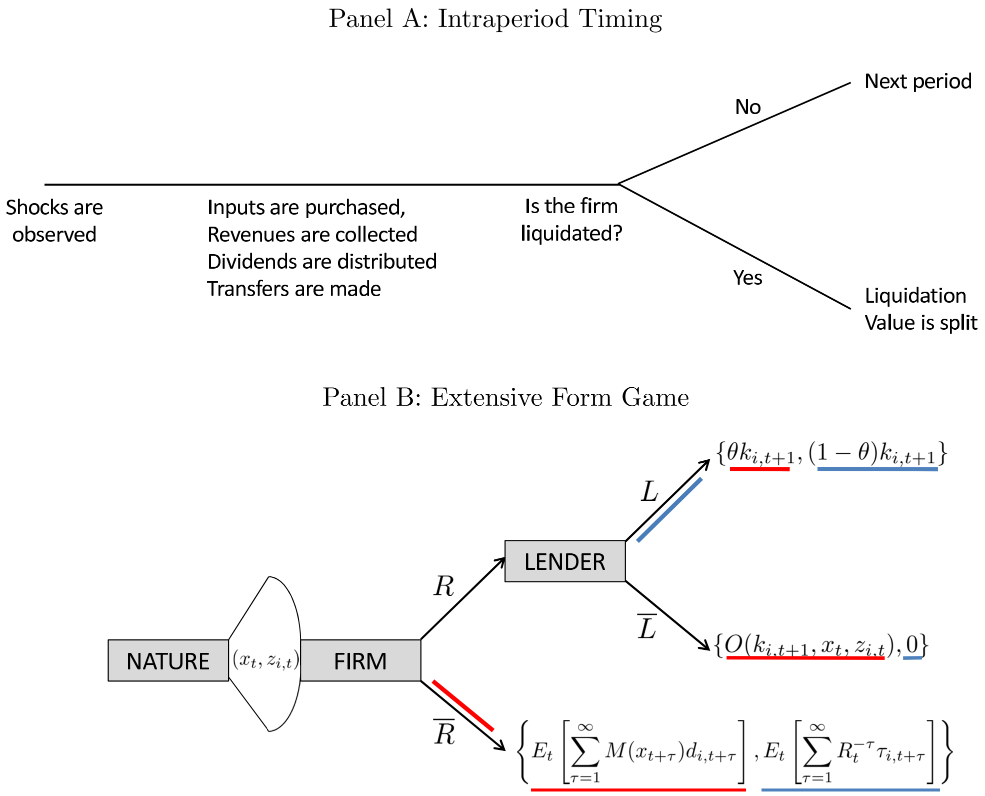


Panel B: Hedging



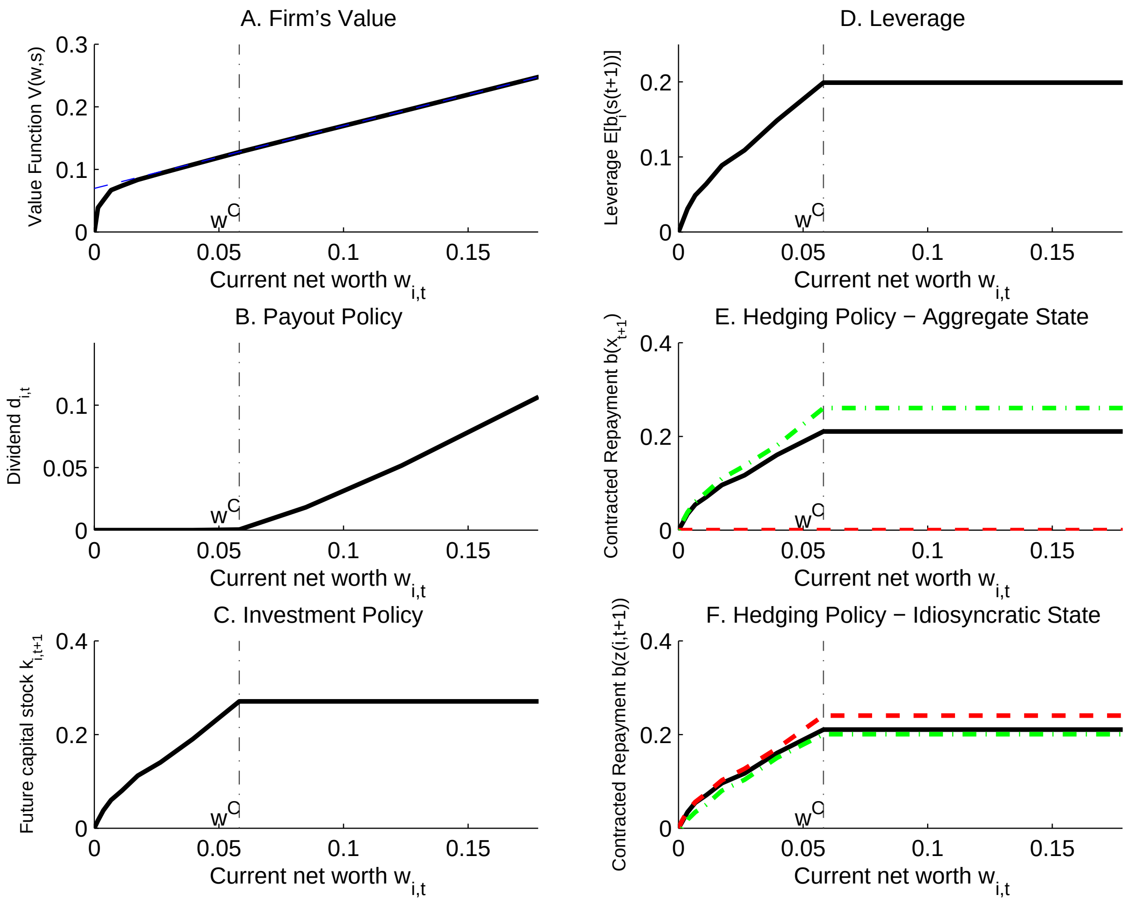
The figure illustrates the set of possible payoffs of a firm with and without hedging in the context of the example in Section 2. Panel A depicts the case of no hedging, while Panel B introduces hedging. In Panel A, the thick solid lines represent the firm's payoff in the sunny ($d(S)$) and rainy ($d(R)$) states for a given payout $d(F)$ in the foggy state. k in capital investment, and b is the debt stock. Blue and red dashed lines represent two possible sets of indifference curves for the representative investor. The equilibrium marginal rate of substitution, and hence the stochastic discount factor, cannot be backed out because the kinks at any decision point are consistent with more than one indifference curve. In Panel B, the firm can transfer resources across states by arranging state-contingent debt repayments $b(S)$, $b(F)$, and $b(R)$ in the sunny, foggy, and rainy states, in the presence of collateral constraints. The payout set is linear, and in equilibrium its slope must be equal to the slope of indifference curves.

Figure 2. The Dynamic Limited Enforcement Problem



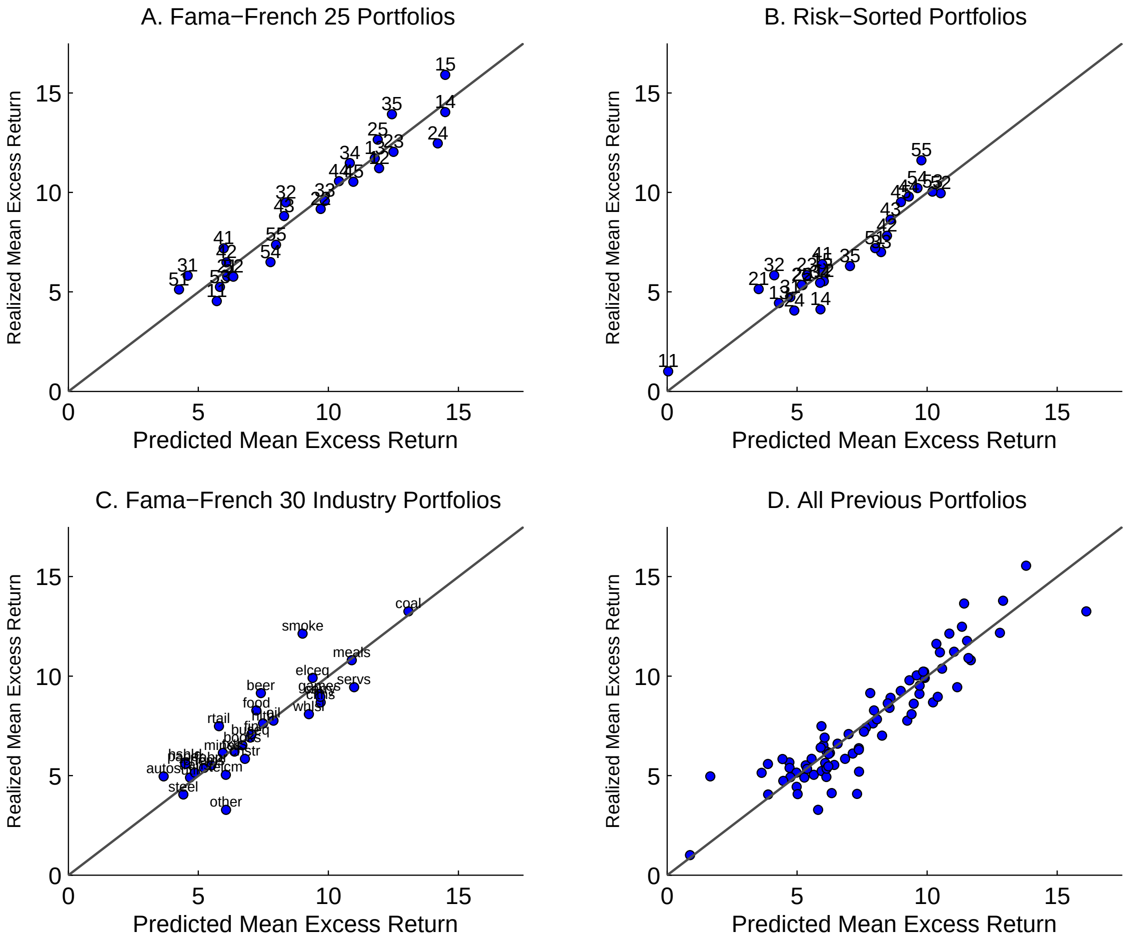
The figure depicts the timing of events in the dynamic limited enforcement problem, as described in the text. Panel A represents the sequence of events that occur each period after the long-term contract between the lender and the borrower is signed. Panel B shows the extensive form of the game from which enforcement constraints arise as an equilibrium outcome. In Panel B, red lines and blue lines represent optimal strategies and payoffs for the firm and the lender respectively. The possible strategies for the borrower are either to renege the contract ($\overline{R}$), or to continue running the firm (R). If the borrower decides to renege the contract, The possible strategies for the lender are either to liquidate the firm (L), or to not liquidate the firm ($\overline{L}$). At time t and for firm i, $M(x_t)$ denotes the stochastic discount factor, R_t is the risk-neutral lender's discount rate, $d_{i,t}$ the dividend payment, $\tau_{i,t}$ the repayment to the lender, $k_{i,t}$ the firm's capital stock, $O(k_{i,t+1}, s_{i,t})$ the value of the outside opportunity for the entrepreneur, and $1-\theta$ the fraction of capital the lender can expropriate upon liquidation. $s_{i,t}$ in the state of the economy, and consists of an aggregate shock x_t, and of a firm-specific shock $z_{i,t}$.

Figure 3. Firm's policy: Illustration



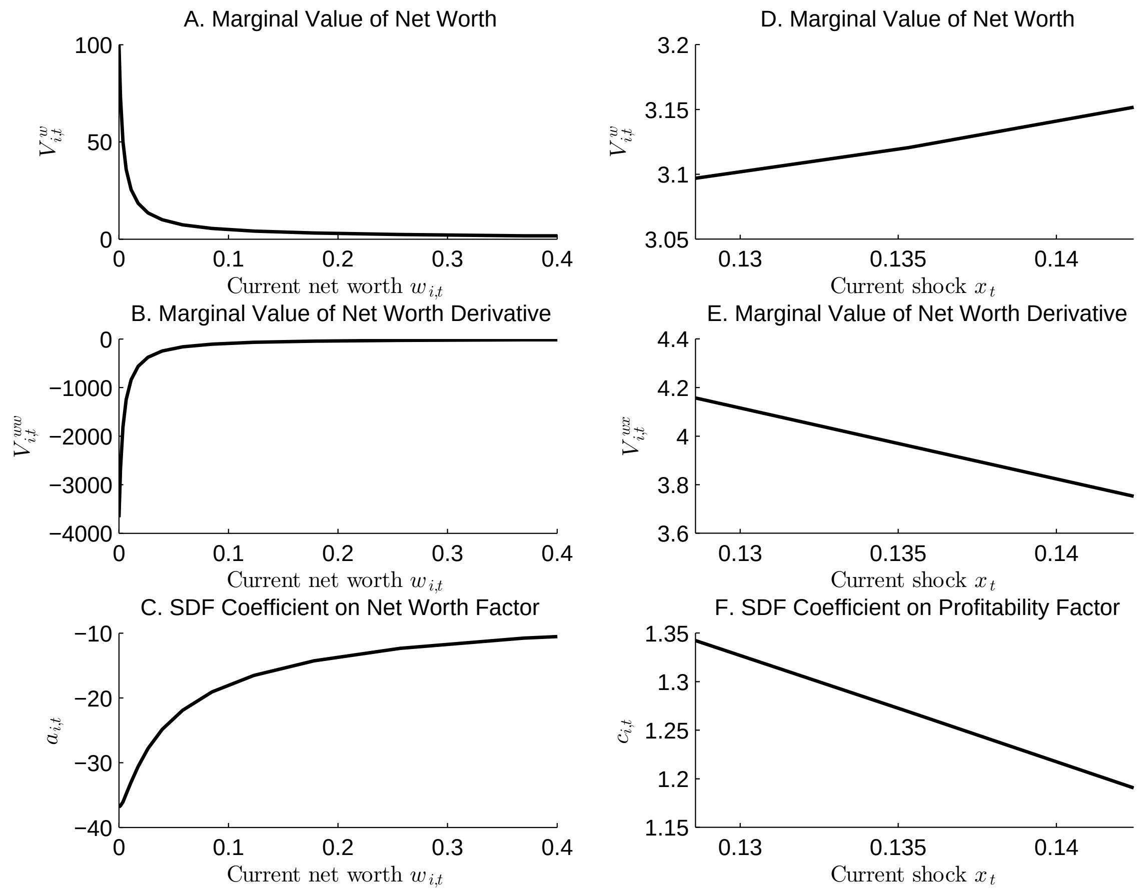
The figure illustrates the investment, payout, financing, and hedging policy of the firm as a function of current net worth $w_{i,t}$. The model is solved under the baseline calibration in Table 1. Panels A through F show: firm's equity value $V(w(s_{i,t}), s_{i,t})$, dividend payouts $d_{i,t}$ the new capital stock $k_{i,t+1}$, the observed debt stock $E[b(s_{i,t+1})]$, the debt repayment in three different aggregate states $b(x_{t+1})$, and the debt repayment in three different idiosyncratic states $b(z_{i,t+1})$. In all Panels, w^C denotes the net worth cutoff that delimits the region in which the firm is paying dividends. In Panel A, the dashed blue line represents the 45-degree slope of the value function in the region where dividends are paid. In Panels E and F, the solid line refer to the repayment in the middle state, the dashed red line to the one-state-down repayment, and the dash-dotted green line to the one-state-up repayment.

Figure 4. Predicted vs Realized Excess Returns: Corporate CAPM.



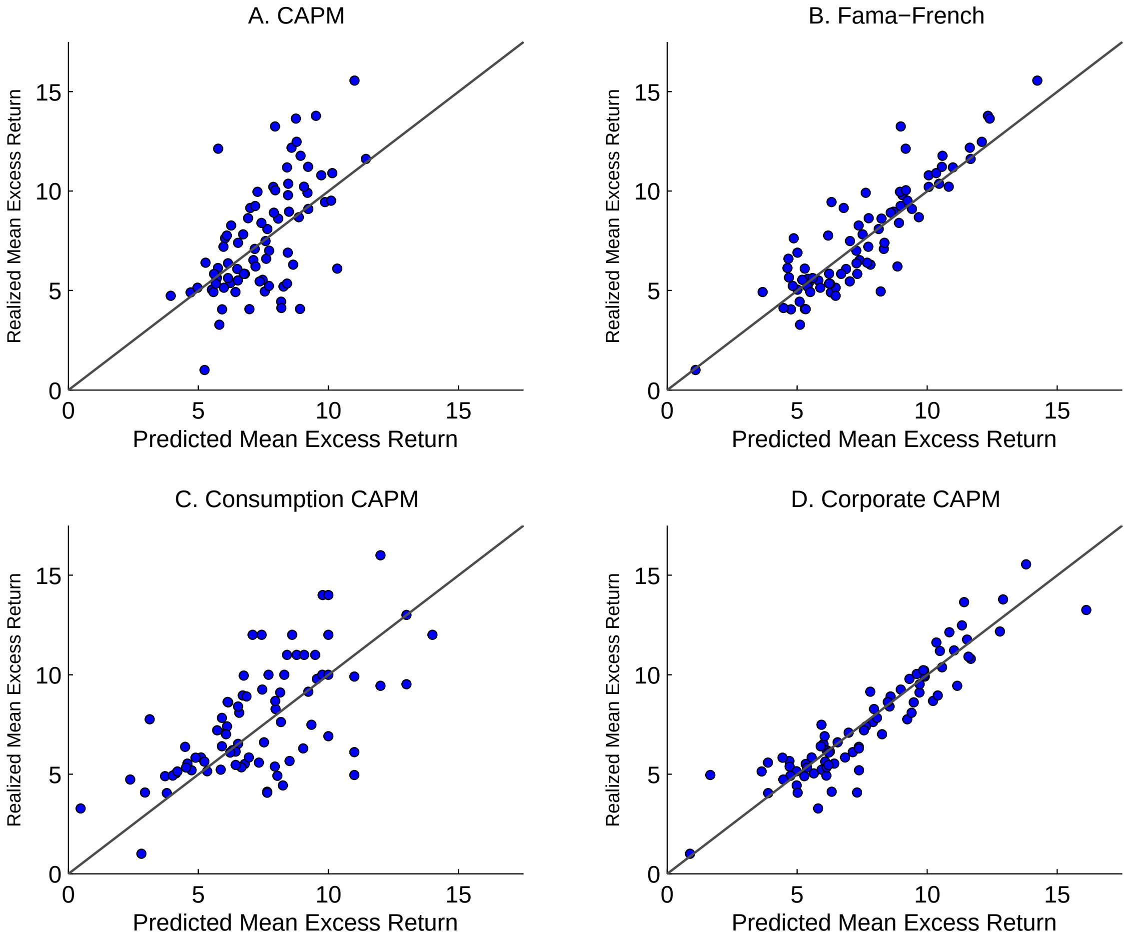
The figure illustrates annual predicted and realized excess returns for the first-stage GMM estimation of the Corporate CAPM as in Table 2. Panels A through D refer to the following test assets: the 25 Fama and French's portfolios sorted on size and book-to-market equity, 25 risk-sorted portfolios on pre-ranking market and HML betas as in Yogo (2006), the 30 Fama-French industry portfolios, and all the previous portfolios. In Panel A, the first digit of the label corresponds to the size quintile, and the second digit to the book-to-market equity quintile. In Panel B, the first digit of the label corresponds to the pre-ranking HML beta quintile, and the second digit to the market beta within each HML beta group. In Panel C, the labels are mnemonics for Fama and French 30-Industry classification as on Kenneth French's website. Accounting data for the construction of the Corporate CAPM factors are from Compustat Annual. The sample period is from 1965 to 2010.

Figure 5. Coefficients on the Hedging and Profitability Factors



Panels A through C depict the marginal value of net worth $V_w(w_{i,t}, s_{i,t})$, its derivative with respect to current net worth $V_{ww}(w_{i,t}, s_{i,t})$, and the coefficient $\overline{a}_{i,t} \equiv \frac{V_{ww}(w_{i,t},s_{i,t})}{V_w(w_{i,t},s_{i,t})}$, all as a function of current net worth $w_{i,t}$. The pictures refer to the steady state for both aggregate and idiosyncratic shocks. Panels D through F depict the marginal value of net worth $V_w(w_{i,t}, s_{i,t})$, its derivative with respect to current net worth $V_{ww}(w_{i,t}, s_{i,t})$, and the coefficient $\overline{c}_{i,t} \equiv \frac{V_{wx}(w_{i,t},s_{i,t})}{V_w(w_{i,t},s_{i,t})}$, all as a function of the current aggregate shock x_t. The pictures refer to the steady state for both net worth and aggregate shocks. The coefficients $\overline{a}_{i,t}$ and $\overline{c}_{i,t}$ on the net worth and aggregate profitability factors are aggregated for conditional tests and lead to the specification of the Corporate CAPM in Equation (43).

Figure 6. Predicted vs Realized Excess Returns: Comparison Among Models.



The figure illustrates predicted and realized excess returns for the first-stage GMM estimation of different asset pricing models. All returns are annual and in excess of the riskfree rate. The test assets are the 25 Fama and French's portfolios sorted on size and book-to-market equity, 25 risk-sorted portfolios on pre-ranking market and HML betas as in Yogo (2006), and the 30 Fama-French industry portfolios, all together. Panels A through D refer to the asset pricing models estimated in Table 5: the CAPM, the three factor model of Fama and French, the Consumption CAPM, and the Corporate CAPM (unconditional estimation). Accounting data for the construction of the Corporate CAPM factors are from Compustat Annual. The sample period is from 1965 to 2010.

Table 1. Model Calibration.

The table reports actual and simulated moments, together with the corresponding choice of structural parameters. Panel A reports a set of moments that refers to corporate policies, and the corresponding data values. Calculations of data moments in Panel A are based on a sample of nonfinancial, unregulated firms from the annual 2012 Compustat Industrial database. The sample period is from 1988 to 2001. Operating income is defined as $(x_{t+1}z_{t+1}k_t^{\alpha})/k_t$, investment as $i_t = k_{t+1} - (1-\delta)k_t$, leverage as $E[b(s_{t+1})]/(E[b(s_{t+1})] + V(w_t, s_t))$, distributions as d_t/k_t and Tobin's Q as $(V(w_t, s_t) + E[b(s_{t+1})])/k_t$. Panel B reports a set of simulated aggregate asset pricing moments, whose data counterparts are from previous studies. Panel C reports the chosen values for structural parameters. Parameters in Group I are those whose value can be restricted from previous works or maps directly into data moments. Parameters in Group II pertain to the pricing kernel and are set to match the average real riskfree rate, the real riskfree rate volatility, and the average Sharpe ratio. Parameters in Group III are set to match simulated moments to data moments. α is the curvature of the production function, θ is the fraction of diverted capital in case of liquidation, δ is the depreciation rate, β, γ_0, and γ_1 are the parameters in the stochastic discount factor, μ_x, ρ_x, σ_x are the parameters driving the dynamics of the aggregate shock, ρ_z, and σ_z are the parameters driving the dynamics of the idiosyncratic shock, and ϕ is the fraction of incumbents per period.

Panel A: Corporate Policy Moments		
	Simulated Moments	Data Moments
Mean of operating income	0.2115	0.1387
Variance of operating income	0.0077	0.0068
Serial correlation of operating income	0.6706	0.7920
Mean of investment	0.1609	0.2018
Variance of investment	0.0568	0.0516
Mean of leverage	0.3931	0.2820
Variance of leverage	0.0427	0.0546
Serial correlation of leverage	0.6493	0.7723
Average distributions	0.0486	0.0310
Mean Tobin's Q	1.6522	1.5594

Panel B: Aggregate Moments		
	Simulated Moments	Data Moments
Mean of riskfree rate	0.0219	0.0290
Volatility of riskfree rate	0.0375	0.0300
Mean of Sharpe Ratio	0.3499	0.4100
Average excess returns	0.0627	0.0790
Variance of aggregate returns	0.0228	0.0317
Mean of firm-level return variances	0.0804	0.1149

Panel C: Calibrated Parameters											
Group I				Group II			Group III				
δ	ρ_x	σ_x	ϕ	β	γ_0	γ_1	α	θ	μ_x	ρ_z	σ_z
0.1500	0.8145	0.0280	0.0200	0.9400	12.5	-120	0.7600	0.3000	-2.0	0.8700	0.0750

Table 2. Unconditional Tests of the Corporate CAPM.

The table reports the estimated factor loading on the net worth, and profitability factors for the Corporate CAPM. The test assets are the 25 Fama and French's portfolios sorted on size and book-to-market equity, 25 risk-sorted portfolios sorted on pre-ranking market and HML betas as in Yogo (2006), the 30 Fama-French industry portfolios, and all the previous portfolios. All returns are annual and in excess of the riskfree rate. In this specification, the model is estimated unconditionally, and the curvature parameter α is set to the calibrated value of 0.76. Estimation is by two-step GMM. HAC standard errors are in parentheses. The kernel is Newey-West with a lag length of 1 year. MAE denotes the percent mean absolute pricing error, and the R^2 is computed as in Campbell and Vuolteenaho (2004). The latter two statistics are based on first-stage estimates. HJ denotes the Hansen-Jagannathan distance, computed as in Jagannathan and Wang (1996). $p(HJ)$ is the p-value for the HJ test corrected for degrees of freedom as in Ferson and Foerster (1994). J and $p(J)$ denote the test statistic and the p-value for a test of overidentifying restrictions. Accounting data for the construction of the Corporate CAPM factors are from Compustat Annual. The sample period is from 1965 to 2010.

	Test Assets			
Estimate	25 S/BM	FF 30 Ind	Risk-Sorted	All
Net Worth	2.853	-8.392	0.879	-6.344
	(1.623)	(0.978)	(1.574)	(1.367)
Profitability	23.145	26.927	50.785	27.621
	(1.285)	(1.572)	(5.826)	(5.831)
MAE (%)	0.764	0.790	0.676	0.838
R^2	0.923	0.771	0.872	0.846
HJ Distance	0.773	0.828	0.669	-
$p(HJ)$	(0.768)	(0.982)	(0.913)	-
J	22.333	22.405	17.730	22.487
$p(J)$	(0.500)	(0.762)	(0.772)	(1.000)

Table 3. Conditional Tests: Nonlinear Regression for $\overline{a}_{i,t}$ and $\overline{c}_t$.

The table reports estimated coefficients and the R^2 for a nonlinear regression of the time-varying coefficient $\overline{a}_{i,t}$, and of a linear regression of $\overline{c}_t$, for the conditional specification of empirical tests of the Corporate CAPM. The values of $\overline{a}_{i,t}$ are regressed from the numerical solution of the model on the endogenous state variable $w_{i,t}$, with the functional form:

$$a_0 \frac{1}{1 + a_1 w_{i,t}}$$

Estimation is based the algorithm in Levenberg (1944) and Marquardt (1963). The values of $\overline{c}_t \equiv \frac{1}{N} \sum_{j=1}^{N} \overline{c}_{j,t}$ are regressed from the numerical solution of the model on the state variable x_t, with the functional form:

$$c_0 + c_1 \rho_t^A$$

Standard errors are in parentheses. The R^2 is from a cross-sectional regression of fitted on actual values.

Dependent Variable: $\overline{a}_{i,t}$			
Functional Form	a_0	a_1	R^2
$a_0 \cdot \frac{1}{1+a_1 w_{i,t}}$	-35.424	7.489	0.969
	(0.295)	(0.271)	

Dependent Variable: $\overline{c}_t$			
Functional Form	c_0	c_1	R^2
$c_0 + c_1 \rho_t^A$	4.142	-17.623	0.999
	(0.072)	(0.529)	

Table 4. Conditional Tests of the Corporate CAPM

The table reports the estimated factor loading on the net worth, and profitability factors for the Corporate CAPM. The test assets are the 25 Fama and French's portfolios sorted on size and book-to-market equity, 25 risk-sorted portfolios sorted on pre-ranking market and HML betas as in Yogo (2006), the 30 Fama-French industry portfolios, and all the previous portfolios. All returns are annual and in excess of the riskfree rate. In this specification, the model is estimated conditionally with the stochastic discount factor in Equation (43), in which the coefficient $\overline{a}_{i,t}$ for "net worth" factor is time varying, and as in Table 3, is parametrized as:

$$a_0 \frac{1}{1 + a_1 w_{i,t}}$$

and the estimated coefficient for the "profitability factor" is parametrized as:

$$c_0 + c_1 \rho_t^A$$

The table reports the estimates for a_0 and c_0, while a_1 is set to 7.489, and c_1 is set to -17.623 as estimated in Table 3. The curvature parameter α is set to the calibrated value of 0.76. Estimation is by two-step GMM. Standard errors are in parentheses, and are computed with HAC standard error. The kernel is Newey West with a lag length of 1 year. MAE denotes the percent mean absolute pricing error, and the R^2 is computed as in Campbell and Vuolteenaho (2004). The latter two statistics are based on first-stage estimates. HJ denotes the Hansen-Jagannathan distance, computed as in Jagannathan and Wang (1996). $p(HJ)$ is the p-value for the HJ test corrected for degrees of freedom as in Ferson and Foerster (1994). J and $p(J)$ denote the test statistic and the p-value for a test of overidentifying restrictions. Accounting data for the construction of the Corporate CAPM factors are from Compustat Annual. The sample period is from 1965 to 2010.

	Test Assets			
Estimate	25 S/BM	FF 30 Ind	Risk-Sorted	All
Net Worth	19.536	-97.200	7.272	-21.700
	(38.975)	(14.240)	(40.198)	(5.009)
Profitability	34.361	27.007	40.231	28.097
	(2.788)	(0.934)	(3.613)	(4.160)
MAE (%)	0.634	0.784	0.557	0.722
R^2	0.944	0.820	0.911	0.888
HJ Distance	0.876	0.810	0.783	-
$p(HJ)$	(0.711)	(0.981)	(0.901)	-
J	22.714	22.863	18.997	22.254
$p(J)$	(0.478)	(0.740)	(0.701)	(1.000)

Table 5. Comparison Among Models.

Columns 1 through 5 report performance measures for the CAPM, the three factor model of Fama and French, the consumption CAPM, and the Corporate CAPM. For the Corporate CAPM, the results for unconditional estimates are in Column 5, and those for conditional estimates are in Column 6. Panels A through D refer to the following test assets: the 25 Fama and French's portfolios sorted on size and book-to-market equity, 25 risk-sorted portfolios on pre-ranking market and HML betas as in Yogo (2006), the 30 Fama-French industry portfolios, and all the previous portfolios. All returns are annual and in excess of the riskfree rate. Estimation is by two-step GMM. HAC standard error are in parentheses. The kernel is Newey-West with a lag length of 1 year. MAE denotes the percent mean absolute pricing error, and the R^2 is computed as in Campbell and Vuolteenaho (2004). The latter two statistics are based on first-stage estimates. HJ denotes the Hansen-Jagannathan distance, computed as in Jagannathan and Wang (1996). p(HJ) is the p-value for the HJ test corrected for degrees of freedom as in Ferson and Foerster (1994). J and $p(J)$ denote the test statistic and the p-value for a test of overidentifying restrictions. Accounting data for the construction of the Corporate CAPM factors are from Compustat Annual. The sample period is from 1965 to 2010.

	CAPM	Fama-French	CCAPM	Corporate CAPM (Unconditional)	Corporate CAPM (Conditional)
Panel A. 25 Fama-French Portfolios					
MAE (%)	1.764	0.673	1.414	0.752	0.634
R^2	0.510	0.915	0.586	0.923	0.944
HJ	0.871	0.863	0.869	0.804	0.876
$p(HJ)$	(0.736)	(0.293)	(0.885)	(0.735)	(0.711)
J	19.508	21.459	20.913	21.968	22.714
$p(J)$	(0.724)	(0.493)	(0.644)	(0.522)	(0.478)
Panel B. 25 Risk-Sorted Portfolios					
MAE (%)	1.857	0.815	1.911	0.758	0.557
R^2	0.217	0.837	0.196	0.852	0.911
HJ	0.761	0.761	0.773	0.693	0.713
$p(HJ)$	(0.912)	(0.683)	(0.942)	(0.894)	(0.901)
J	19.897	19.886	21.809	20.354	18.997
$p(J)$	(0.703)	(0.590)	(0.591)	(0.620)	(0.701)
Panel C. 30 Fama-French Industry Portfolios					
MAE (%)	1.362	1.095	1.629	0.935	1.015
R^2	0.264	0.630	0.159	0.743	0.784
HJ	0.846	0.848	0.877	0.822	0.820
$p(HJ)$	(0.988)	(0.906)	(0.993)	(0.982)	(0.981)
J	20.232	20.838	22.306	22.112	22.863
$p(J)$	(0.886)	(0.794)	(0.807)	(0.776)	(0.740)
Panel D. All 80 Portfolios					
MAE (%)	1.703	0.990	1.829	0.838	0.722
R^2	0.378	0.791	0.349	0.846	0.888
J	22.349	22.358	22.475	22.487	22.254
$p(J)$	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)

Table 6. The Corporate CAPM: Industry Breakdowns.

The table reports the estimated factor loading on the net worth, and profitability factors for the Corporate CAPM. The test assets are the 25 Fama and French's portfolios sorted on size and book-to-market equity, 25 risk-sorted portfolios on pre-ranking market and HML betas as in Yogo (2006), and the 30 Fama-French industry portfolios, all together. All returns are annual and in excess of the riskfree rate. The first row reports the reference set of firms with respect to the Corporate CAPM factors are computed, and corresponds to Fama and French's five-industry classification. Panel A refers to unconditional tests, implemented as in Table 2. Panel B refers to conditional tests, implemented as in Table 4. Estimation is by two-step GMM. HAC standard errors are in parentheses. The kernel is Newey-West with a lag length of 1 year. MAE denotes the percent mean absolute pricing error, and the R^2 is computed as in Campbell and Vuolteenaho (2004). The latter two statistics are based on first-stage estimates. J and $p(J)$ denote the test statistic and the p-value for a test of overidentifying restrictions. Accounting data for the construction of the Corporate CAPM factors are from Compustat Annual. The sample period is from 1965 to 2010.

Panel A: Unconditional Tests

	Reference Industry				
Estimate	Cnsmr	Manuf	HiTec	Hlth	Other
Net Worth	-8.567	-2.148	-2.359	-3.795	-5.224
	(1.976)	(0.480)	(0.508)	(0.821)	(1.105)
Profitability	34.680	15.858	11.321	19.065	20.690
	(7.337)	(3.346)	(2.392)	(4.025)	(4.379)
MAE (%)	0.930	0.959	0.898	1.214	0.895
R^2	0.822	0.854	0.841	0.830	0.823
J	22.487	22.481	22.465	22.470	22.432
$p(J)$	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)

Panel B: Conditional Tests

	Reference Industry				
Estimate	Cnsmr	Manuf	HiTec	Hlth	Other
Net Worth	-17.700	-27.000	-20.300	-18.900	-47.500
	(7.239)	(5.861)	(4.772)	(4.462)	(10.343)
Profitability	12.975	4.902	6.526	10.107	7.189
	(2.751)	(1.034)	(1.384)	(2.132)	(1.529)
MAE (%)	1.212	1.297	0.856	0.951	0.864
R^2	0.720	0.728	0.847	0.835	0.822
J	22.485	22.480	22.487	22.475	22.457
$p(J)$	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)

Appendix A. Proofs of Propositions

Proof of Lemma 1. Denote by Y the set of the possible values for the state variables $w_{i,t}$ and $s_{i,t}$, by $\Gamma(y)$ the set of possible actions $k_{i,t+1}$ and $b(s_{i,t+1})$ for each $y \in Y$. Let V be the set of functions from Y to $(-\infty, \infty)$. In the remainder of the proof, I use the shorthands V^{LB} for $V^{LB}(w_{i,t}, s_{i,t})$, V^{UB} for $V^{UB}(w_{i,t}, s_{i,t})$, and V^* for $V(w_{i,t}, s_{i,t})$. Denote by $\leq$ be partial order operator for the functions on V, and by T the Bellman operator defined by

$$(Tv)(y) = \sup_{a \in \Gamma(y)} \left(d(y,a) + E_t \left[\beta M_0(x_{t+1}) v(y') \right] \right), \qquad y, y' \in Y, v \in V \tag{A.1}$$

In this setting, the number of states is assumed to be finite, and by no arbitrage we have $M_0(\cdot) > 0$. Therefore, from the definition of T, it follows that T is monotone. Furthermore, $T(V^{UB}) \leq V^{UB}$, and $T(V^{LB}) \geq V^{LB}$. Under these conditions, the Knaster-Tarski fixed-point theorem (Aliprantis and Border (2006), Theorem 1.10) guarantees that the Bellman operator has at least one fixed point V^{FP} in $[V^{LB}, V^{UB}]$. Define the sequence V_n^{LB}, with $n = 0, 1, 2, ...$ such that $V_0^{LB} = V^{LB}$, and $V_{n+1}^{LB} = TV_n^{LB}$. Since any fixed point of T in $[V^{LB}, V^{UB}]$ is bounded above by V^{UB}, the increasing sequence V_n^{LB} must converge to a fixed point $\hat{V}^{LB}$ in $[V^{LB}, V^{UB}]$. By definition of fixed point, $V^{FP} = TV^{FP}$, and, by construction, $V_n^{LB} \leq V^{FP}$, for all n. Thus, $\hat{V}^{LB} \leq V^{FP}$. By (23), and since the number of states is finite, the conclusion of Theorem 4.3 in Stokey and Lucas (1989) goes through. Therefore $V^* = V^{FP}$. Finally, the assumptions for Lemma 4.3 in Kamihigashi (2012) are satisfied, and this guarantees that $V^* \leq \hat{V}^{LB}$. As a consequence, the following chain of inequalities holds:

$$V^* \leq \hat{V}^{LB} \leq V^{FP} = V^* \tag{A.2}$$

This establishes that the uniqueness result in part (i), and the convergence results in part (ii). □

Proof of Proposition 1. Part (i). As Equation (26) states, the first-order conditions of problem (17)-(22) with respect to $b(s_{i,t+1})$ are:

$$R_t \nu(s_{i,t+1}) M(x_{t+1}) = \frac{\nu_{i,t}}{1 + \lambda_{i,t}} \tag{A.3}$$

Solving the previous equation for $M(x_{t+1})$, the stochastic discount factor can be obtained as:

$$M(x_{t+1}) = \frac{\nu_{i,t}}{R_t V_w(w(s_{i,t+1}), x_{t+1}, z_{i,t+1})(1 + \lambda_{i,t})} \tag{A.4}$$

The envelope condition (17)-(22) with respect to the state variable $w_{i,t}$ is:

$$\nu_{i,t} = V_w(w_{i,t}, x_t, z_{i,t}) \tag{A.5}$$

Plugging the expression of the multiplier $\nu_{i,t}$ from Equation (A.5) into (A.4) yields:

$$\begin{aligned} M(x_{t+1}) &= \frac{V_w(w_{i,t}, x_t, z_{i,t})}{R_t V_w(w(s_{i,t+1}), x_{t+1}, z_{i,t+1})(1+\lambda_{i,t})} = \\ &= \mu^M_{i,t} \frac{V_w(w_{i,t}, x_t, z_{i,t})}{V_w(w(s_{i,t+1}), x_{t+1}, z_{i,t+1})} \end{aligned} \tag{A.6}$$

Part (ii). Taking the log of both sides of (27) yields

$$\log M(x_{t+1}) = \mu^M_{i,t} + \log \frac{V_w(w_{i,t}, x_t, z_{i,t})}{V_w(w(s_{i,t+1}), x_{t+1}, z_{i,t+1})} \tag{A.7}$$

Define $f(w(s_{i,t+1}, s_{i,t+1}) \equiv \log \frac{V_w(w(s_{i,t+1}), x_{t+1}, z_{i,t+1})}{V_w(w_{i,t}, x_t, z_{i,t})}$. A first-order Taylor expansion of $f(w(s_{i,t+1}, s_{i,t+1})$ around the previous period realization $(w_{i,t}, s_{i,t})$ leads to:

$$\begin{aligned} f(w(s_{i,t+1}, s_{i,t+1}) &\simeq f(w_{i,t}, s_{i,t}) + f_w(w_{i,t}, s_{i,t})(w(s_{i,t+1}) - w_{i,t}) + f_z(w_{i,t}, s_{i,t})(z_{i,t+1} - z_{i,t}) \\ &+ f_x(w_{i,t}, s_{i,t})(x_{i,t+1} - x_{i,t}) \end{aligned} \tag{A.8}$$

Since

$$f(w_{i,t}, s_{i,t}) = 1 \tag{A.9}$$

$$f_w(w_{i,t}, s_{i,t}) = \frac{V_{ww}(w_{i,t}, s_{i,t})}{V_w(w_{i,t}, s_{i,t})} \tag{A.10}$$

$$f_z(w_{i,t}, s_{i,t}) = \frac{V_{wz}(w_{i,t}, s_{i,t})}{V_w(w_{i,t}, s_{i,t})} \tag{A.11}$$

$$f_x(w_{i,t}, s_{i,t}) = \frac{V_{wx}(w_{i,t}, s_{i,t})}{V_w(w_{i,t}, s_{i,t})} \tag{A.12}$$

and because, expressing x_t as a Solow residual and recovering $z_{i,t}$ as a function of it, I obtain:

$$x_t = \rho^A_t \tag{A.13}$$

$$z_{i,t} = \frac{\rho_{i,t}}{\rho^A_t} \tag{A.14}$$

Then Equation (A.8) simplifies as

$$\log \frac{V_w(w(s_{i,t+1}), x_{t+1}, z_{i,t+1})}{V_w(w_{i,t}, x_t, z_{i,t})} = \bar{a}_{i,t}(w(s_{i,t+1}) - w_{i,t}) + \bar{b}_{i,t}\left(\frac{\rho_{i,t+1}}{\rho^A_{t+1}} - \frac{\rho_{i,t}}{\rho^A_t}\right) + \bar{c}_{i,t}\left(\rho^A_{t+1} - \rho^A_t\right) \tag{A.15}$$

Plugging (A.15) into (A.7) yields the result. □

Proof of Proposition 2. The stochastic discount factor can be log-linearized at the first-order as

$$\frac{M(x_{t+1})}{E_t[M(x_{t+1})]} \approx 1 + \log M(x_{t+1}) - E_t[\log M(x_{t+1})] \tag{A.16}$$

that, using equation (28), can be written as:

$$\frac{M(x_{t+1})}{E_t[M(x_{t+1})]} \approx \mu^M_{j,t} - \overline{a}_{j,t}(w(s_{j,t+1}) - w_{j,t}) - \overline{b}_{j,t}\left(\frac{\rho_{j,t+1}}{\rho^A_{t+1}} - \frac{\rho_{j,t}}{\rho^A_t}\right) - \overline{c}_{j,t}\left(\rho^A_{t+1} - \rho^A_t\right) \tag{A.17}$$

The SDF can therefore be approximated with a two-factor linear representation, that is

$$\frac{M(x_{t+1})}{E_t[M(x_{t+1})]} \approx \mu^M_{j,t} - \overline{a}_{j,t} f^1_{j,t+1} - \overline{b}_{j,t} f^2_{j,t+1} - \overline{c}_{j,t} f^3_{j,t+1} \tag{A.18}$$

with

$$f^1_{j,t+1} \equiv w(s_{i,t+1}) - w_{i,t} \tag{A.19}$$

$$f^2_{j,t+1} \equiv \frac{\rho_{j,t+1}}{\rho^A_{t+1}} - \frac{\rho_{j,t}}{\rho^A_t} \tag{A.20}$$

$$f^3_{j,t+1} \equiv \rho^A_{t+1} - \rho^A_t \tag{A.21}$$

$M(x_{t+1})$ is a valid stochastic discount factor for equity returns $R_{i,t+1}$, and for the riskfree return R^f_t. Therefore:

$$E_t[M(x_{t+1})R_{i,t+1}] = E_t[M(x_{t+1})R^f_t] = 1 \tag{A.22}$$

The previous equation can be rewritten as

$$E_t\left[M(x_{t+1})(R_{i,t+1} - R^f_t)\right] = 0 \tag{A.23}$$

The constant in the SDF is measurable with respect to the time-t information set. Thus, I obtain

$$E_t\left[\frac{M(x_{t+1})}{E_t[M(x_{t+1})]}(R_{i,t+1} - R^f_t)\right] = 0 \tag{A.24}$$

that is

$$Cov_t\left[\frac{M(x_{t+1})}{E_t[M(x_{t+1})]}, R_{i,t+1} - R^f_t\right] + E_t\left[\frac{M(x_{t+1})}{E_t[M(x_{t+1})]}\right] E_t[R_{i,t+1} - R^f_t] = 0 \tag{A.25}$$

Substituting the approximated expression for the SDF in equation (A.18):

$$\begin{aligned} E_t[R_{i,t+1} - R^f_t] &\approx -Cov_t\left[\mu^M_{j,t} - \overline{a}_{j,t} f^1_{j,t+1} - \overline{b}_{j,t} f^2_{j,t+1} - \overline{c}_{j,t} f^3_{j,t+1}, R_{i,t+1} - R^f_t\right] = \\ &= -Cov_t\left[-\overline{a}_{j,t} f^1_{j,t+1} - \overline{b}_{j,t} f^2_{j,t+1} - \overline{c}_{j,t} f^3_{j,t+1}, R_{i,t+1} - R^f_t\right] \end{aligned} \tag{A.26}$$

Consider the column vector $f_{j,t+1}$ obtained by stacking $f^1_{j,t+1}$, $f^2_{j,t+1}$, and $f^3_{j,t+1}$. The variance-covariance matrix of the factors is

$$Var_t\left[f_{j,t+1}\right] \equiv Var_t\begin{bmatrix} f^1_{j,t+1} \\ f^2_{j,t+1} \\ f^3_{j,t+1}\end{bmatrix} = \begin{bmatrix} Var_t(f^1_{j,t+1}) & Cov_t(f^1_{j,t+1}, f^2_{j,t+1}) & Cov_t(f^1_{j,t+1}, f^3_{j,t+1}) \\ Cov_t(f^2_{j,t+1}, f^1_{j,t+1}) & Var_t(f^2_{j.t+1}) & Cov_t(f^2_{j,t+1}, f^3_{j,t+1}) \\ Cov_t(f^3_{j,t+1}, f^1_{j,t+1}) & Cov_t(f^3_{j,t+1}, f^2_{j,t+1}) & Var_t(f^3_{j.t+1}) \end{bmatrix} \tag{A.27}$$

and the vector $\widetilde{b}_{j,t}$ as

$$\widetilde{b}_{j,t} \equiv \begin{bmatrix} -\overline{a}_{j,t} \\ -\overline{b}_{j,t} \\ -\overline{c}_{j,t} \end{bmatrix}$$

Then, it follows that:

$$\begin{aligned} E_t[R_{i,t+1} - R^f_t] &\approx -\widetilde{b}^T_{j,t} Cov_t\left[f_{j,t+1}, R_{i,t+1} - R^f_t\right] = \qquad (A.28) \\ &= -\widetilde{b}^T_{j,t} Var_t\left[f_{j,t+1}\right] Var_t\left[f_{j,t+1}\right]^{-1} Cov_t\left[f_{j,t+1}, R_{i,t+1} - R^f_t\right] = \\ &= \widetilde{\lambda}^T_{j,t}\beta_{i,t} \end{aligned}$$

where

$$\widetilde{\lambda}^T_{j,t} \equiv -\widetilde{b}^T_{j,t} Var_t\left[f_{j,t+1}\right] \tag{A.29}$$

$$\beta_{j,t} \equiv Var_t\left[f_{j,t+1}\right]^{-1} Cov_t\left[f_{j,t+1}, R_{i,t+1} - R^f_t\right] \tag{A.30}$$

Substituting back the explicit expressions for $f^1_{j,t+1}$, $f^2_{j,t+1}$, and $f^3_{j,t+1}$ completes the proof. □

Proof of Proposition 3. Part (i). The current aggregate state imposes a restriction of firms' investment, and financing policy such that the left-hand side of equation (28) is equalized across firms. Therefore:

$$\begin{aligned} \frac{1}{N}\sum_{j\in\Omega} \log M(x_{t+1}) &= \log M(x_{t+1}) \approx \qquad (A.31) \\ &\approx \frac{1}{N}\sum_{j\in\Omega}\left[\mu^M_{j,t} - \overline{a}_{j,t}(w(s_{j,t+1}) - w_{j,t}) - \overline{b}_{j,t}\left(\frac{\rho_{j,t+1}}{\rho^A_{t+1}} - \frac{\rho_{j,t}}{\rho^A_t}\right) - \overline{c}_{j,t}\left(\rho^A_{t+1} - \rho^A_t\right)\right] \end{aligned}$$

The proof of the covariance representation in Equation (35) follows as in the previous proof by replacing

$$\begin{bmatrix} f^1_{j,t+1} \\ f^2_{j,t+1} \\ f^3_{j,t+1} \end{bmatrix} \text{ with } \begin{bmatrix} \frac{1}{N}\sum\limits_{j\in\Omega} \overline{a}_{j,t} f^1_{j,t+1} \\ \frac{1}{N}\sum\limits_{j\in\Omega} \overline{b}_{j,t} f^2_{j,t+1} \\ \frac{1}{N}\sum\limits_{j\in\Omega} \overline{c}_{j,t} f^3_{j,t+1} \end{bmatrix}.$$

Part (ii). Because $\frac{\rho_{j,t+1}}{\rho^A_{t+1}} - \frac{\rho_{j,t}}{\rho^A_t} = z_{j,t+1} - z_{j,t}$ has zero mean, the process for $z_{j,t}$ has a finite support, and $z_{j,t}$ and $z_{i,t}$ are independent for each $i \neq j$, the assumptions in Pruitt (1966) and Rohatgi (1971) hold and, for $N \to \infty$:

$$\frac{1}{N}\sum_{j\in\Omega}\bar{b}_{j,t}(z_{j,t+1}-z_{j,t}) \to 0 \tag{A.32}$$

by the strong law of large numbers. □

Appendix B. Solution by Mixed-Integer Programming

In this section, I discuss the numerical solution method of the model. I introduce the main results on which the solution algorithm is based, and I provide details on its implementation. Proofs are relegated to Supplementary Appendix F. I start considering the perfect enforcement problem without the borrowing constraint (21), and I show the equivalence between the dynamic program and the linear program, along the lines of Ross (1983).

Lemma 2 (Perfect Enforcement Problem as a Linear Program) *The solution of problem (17) subject to (18), (19), (20), and (22) on a discrete grid is equivalent to the solution of the following linear programming problem:*

$$\min_{v_{w,s}} \sum_{w=1}^{nw}\sum_{s=1}^{nx\cdot nz} v_{w,s} \tag{B.1}$$

$$s.t.$$

$$v_{w,s} \geq d_{w,s,a} + \sum_{s'=1}^{nx\cdot nz} \pi(s'|s)M(s')v_{a,s'} \qquad \forall w,s,a \tag{B.2}$$

where nw, nx, and nz are the number of grid points on the grids for $w_{i,t}$, x_t, and $z_{i,t}$ respectively, $v_{w,s}$ is the value function on the grid point indexed by w and s, a is an index for an action on the grid for both capital and state-contingent debt repayments, and $d_{w,s,a}$ denotes the payout corresponding to the action a starting from the state indexed by w and s.

The previous lemma shows that the linear programming solution method does not require the Bellman operator be a contraction mapping. I now incorporate the borrowing constraint (21) into the linear programming representation above. Because the dynamic programming problem with perfect enforcement has a unique solution, there is only one binding constraint (i.e. one optimal action a on the grid) for each state (w,s) in the equivalent linear programming representation. The enforcement constraint (21) dictates that the optimal action $a^*_{w,s}$ for each state (w,s) satisfies

$$\theta k(a^*_{w,s}) \leq \sum_{s'=1}^{nx \cdot nz} \pi(s'|s) M(s') v_{a^*_{w,s},s'} \qquad \forall w,s \quad \text{(B.3)}$$

where $k(a^*_{w,s})$ denotes the point on the capital grid corresponding to the action $a^*_{w,s}$. In the following lemma, I show that the linear programming representation augmented with constraints (B.3) can be solved as a mixed-integer programming problem.

Lemma 3 (Equivalent Mixed-Integer Programming Representation) *The problem in (B.1)-(B.2) with the borrowing constraints in (B.3) is equivalent to:*

$$\min_{v_{w,s}} \sum_{w=1}^{nw} \sum_{s=1}^{nx \cdot nz} v_{w,s} + \sum_{w=1}^{nw} \sum_{s=1}^{nx \cdot nz} \sum_{a \in \Gamma(w,s)} \epsilon \cdot D_{w,s,a} \quad \text{(B.4)}$$

$$s.t.$$

$$d_{w,s,a} + \sum_{s'=1}^{nx \cdot nz} \pi(s'|s) M(s') v_{a,s'} \leq v_{w,s} \qquad \forall w,s,a \quad \text{(B.5)}$$

$$-v_{w,s} + d_{w,s,a} + \sum_{s'=1}^{nx \cdot nz} \pi(s'|s) M(s') v_{a,s'} + N D_{w,s,a} \geq 0 \qquad \forall w,s,a \quad \text{(B.6)}$$

$$\sum_{s'=1}^{nx \cdot nz} \pi(s'|s) M(s') v_{a,s'} + N D_{w,s,a} \geq \theta k(a) \qquad \forall w,s,a \quad \text{(B.7)}$$

where $D_{w,s,a}$ are binary variables, $\epsilon \to 0$ is a positive small number, $N \to \infty$ is a positive large number, and $\Gamma(w,s)$ is the set of feasible actions if the current state is (w,s).

It is important to remark that the mixed-integer problem in the previous lemma is in general *less* constrained than the "first-best" problem with perfect enforcement. In fact, some actions that are feasible in the "first-best" problem do not satisfy the borrowing constraints, and are excluded from $\Gamma(w,s)$. Consistent with this observation, the minimized objective in the problem with limited enforcement is better and, as I show below, results in a lower optimal equity value for each state.

As Trick and Zin (1993) discuss, solving the full mixed-integer program (as well as the full linear problem) would require to store a huge matrix, because the number of constraints in the problem is very large. This would be impractical, in that hardware, memory, and computational requirements would be enormous. For this reason, I resort to constraint generation, which is a standard technique in operation research to solve problems with a large number of constraints. Specifically, constraint generation begins with the solution a relaxed problem with the same objective and only a subset of the constraints. Then, the procedure identifies the remaining constraints in the full problem that are violated. A subset of the violated constraints is then added to the relaxed problem according to a selection rule. The procedure is iterated until all constraints are satisfied. The next lemma proposes a constraint generation algorithm, and shows it converges to the unique fixed point in Lemma 1.

Lemma 4 (Constraint Generation) *The sequence of functions* $\{v^n(w,s)\}_{n=1}^{\infty}$ *generated by the following algorithm converges to the fixed point* $V(w,s)$ *specified in Lemma 1:*
1. solve the problem in Lemma 3 with only the constraints corresponding to zero capital and zero debt for each state (w,s)*;*
2. if all constraints $a \in \Gamma^n(w,s)$*, for all* (w,s)*, are satisfied, terminate the algorithm (where* $\Gamma^n(w,s)$ *is the set of feasible actions at iteration* n*);*
3. for each state (w,s) *add the constraint* $a \in \Gamma^n(w,s)$ *that generates the highest violation in (B.5) with respect to the current solution* $v^n(w,s)$*;*
4. solve the problem with the current set of constraints;
5. go back to step 2.

The constraint generation algorithm above extends the procedure in Trick and Zin (1993), and Trick and Zin (1997). The procedure starts from a solution which is feasible in that it does not violate the enforcement constraint. Then, at iteration n and for each state (w,s), constraints are added using the same rule which is used in value function iteration, namely maximizing the sum of distributions and the expected continuation value given the current maximized value $v^n(w,s)$. In the mixed-integer programming representation, this rule corresponds to selecting the most violated constraint for each state in the feasible set $\Gamma^n(w,s)$. As Trick and Zin (1993) document and the results in Pucci de Farias and Van Roy (2003) suggest, constraint generation allow to achieve significance speed gains. Most important, it avoids to solve the full problem, which would be computationally too demanding.

However, to make the method implementable, one last critical issue must be addressed. The selection of the most violated constraint in the third step of the constraint generation procedure requires searching over a huge vector of grid points for all the choice variables. The computational and memory requirement would still be excessive for a problem with many controls variables. In this setting, this issue is exacerbated by the presence of state-contingent actions. To make the constraint generation operational, I use a separation oracle, that is an auxiliary linear programming problem that identifies the most violated constraint. Separation oracles are standard tools in operation research (Nemhauser and Wolsey (1988), Schrijver (1998), Cook, Cunningham, Pulleyblank, and Schrijver (2011)), and have been recently used in corporate finance by Nikolov, Schmid, and Steri (2013). I detail and describe the separation oracle for this problem at the end of this appendix.

Operatively, the problem is solved using the algorithm in Lemma 4, and the separation oracle. Codes are implemented with Matlab®, and the solver for the mixed-integer programming problems is CPLEX®. Matlab® and CPLEX® are interfaced through the CPLEX Class API®. The workstation has with a CPU with 8 cores and 32GB of RAM. The model is solved with three grid points for the aggregate shock, seven grid points for the idiosyncratic shock, 500 grid points for capital and each state-contingent debt variable, and 27 grid points for net worth. Following McGrattan (1997), the grid for net worth is not evenly spaced, but more points are collocated in the low net worth region, where the curvature of value function is more relevant. Simulated data from the model are based on panels of 5000 firms and 2000 time periods.

Separation Oracle

$$\max_{a=\{k',b(s')\}} d_{w,s,a} + \sum_{s'=1}^{nx \cdot nz} \pi(s'|s)M(s')v_{a,s'} - v_{w,s} \tag{B.8}$$

$s.t.$

$$\underline{k} \leq k' \leq \overline{k} \tag{B.9}$$

$$\underline{b} \leq b(s') \leq \overline{b} \qquad \forall s' \tag{B.10}$$

$$\sum_{s'=1}^{nx \cdot nz} \pi(s'|s)M(s')v_{w(s'),s'} \geq \theta k' \tag{B.11}$$

$$0 \leq p(i_k) \leq 1 \qquad \forall i_k = 1, ..., n_k \tag{B.12}$$

$$\sum_{i_k=1}^{n_k} p(i_k) = 1 \tag{B.13}$$

$$k' = \sum_{i_k=1}^{n_k} p(i_k)k^G(i_k) \tag{B.14}$$

$$d_{w,s,a} = w - k' + \sum_{s'=1}^{nx \cdot nz} \pi(s'|s)M(s')b(s') \tag{B.15}$$

$$d_{w,s,a} \geq 0 \tag{B.16}$$

$$f(k') = \sum_{i_k=1}^{n_k} p(i_k)(k^G(i_k))^\alpha \tag{B.17}$$

$$w(s') = A(s')f(k') + k'(1-\delta) - R_t b(s') \qquad \forall s' = 1...nx \cdot nz \tag{B.18}$$

Equations (B.9) and (B.10) define the bounds for capital and debt, Equation (B.11) is the enforcement constraint and allows to select feasible actions from $\Gamma^n(w, s)$, Equations (B.12) and (B.13) define the variables $p(i_k)$ that have the role to select a grid point for capital on the grid $k^G(i_k)$ and linearize the term k'^α in the production function, Equation (B.14) picks the grid point for the chosen capital stock from $k^G(i_k)$, Equations (B.15) and (B.16) define dividends and impose their positivity, Equation (B.17) computes the nonlinear term in capital in the production function, and Equation (B.18) defines future net worth in each state s'. The solution of the separation oracle for state-contingent debt is a continuous variable and is interpolated to the nearest point on the corresponding grid.