Firm Investment, Costly Reversibility, and Stock Returns

February 2022

Preliminary Version

ABSTRACT

I study the asset pricing implications of investment and disinvestment options with a production-based model featuring costly reversibility. Investment options are contingent claims on assets in place so that they are riskier and earn higher expected returns. Disinvestment options with costly reversibility reduce exposure to aggregate risks amid deteriorating business conditions and lower expected returns on a firm. The inextricable link between investment options and disinvestment options explains the coexistence of the profitability premium and the value premium while retains a positive relation between profitability and market valuation ratios. My model also generates a procyclical profitability premium and a countercyclical value premium.

1 Introduction

Firms with high gross profitability (GPA) on average earn higher returns than do those with low gross profitability (the gross profitability premium, see Novy-Marx (2013)). More profitable firms have low book-to-market ratio[1] (BM) and resemble growth firms, but growth firms tend to be outperformed by value firms, which is the so-called value premium. Both the gross profitability and the book-to-market ratio positively predict future stock returns but the two characteristics are negatively correlated. The concurrence of the profitability premium and the value premium poses a serious challenge to existing asset pricing models. Most structural models assume a perfectly negative correlation between profitability and book-to-market ratio, and they can explain one anomaly but go in the opposite direction in explaining the other [2]. On the empirical side, the Fama-French three-factor model can absorb the value premium but it magnifies the profitability premium. This perplexing phenomenon motivates Fama and French (2015) to include a profitability factor to better capture the cross section of stock returns.

[Insert Figure 1 Here]

In data, the profitability premium and the value premium are not independent of each other. Figure 1 illustrates that the profitability premium is strongest within growth firms (11.6% per year in growth firms versus 3.5% per year in value firms), and that the value premium is strongest within unprofitable firms (10.8% per year in the lowest GPA quintile versus 2.7% per year in the highest GPA quintile). Conventional wisdom claims high profitability, synonymous with a low book-to-market ratio, signifies ample growth opportunities

[1] I define book-to-market ratio (BM) in Section 2, and market valuation ratio (MB) is the multiplicative inverse of BM. As a common practice in the empirical asset pricing literature, firms with high book-to-market ratio are classified as value firms, and those with low book-to-market ratio are classified as growth firms. As such, I use BM to sort stocks and also report BM as one of the portfolio characteristics in the empirical analyses of the paper. In Sections 3,4, and 6, I primarily use MB for the simplicity of describing the relationships between variables.

[2] For example, Carlson, Fisher, and Giammarino (2004) and Zhang (2005) explain the value premium via the operating leverage so their models imply a counterfactual negative profitability premium. Hackbarth and Johnson (2015) propose a model to study the profitability premium, but their approach "makes the value puzzle worse".

for a firm. A low book-to-market ratio intensifies the profitability premium (top two panels in Figure 1), whereas the profitability premium is not enhanced by a high profitability per se (bottom two panels in Figure 1). Equating the abundance of growth opportunities with high profitability or a low book-to-market ratio cannot account for the manner in which they predict future stock returns interactively. The correspondence of the two asset pricing puzzles calls for an exhaustive inspection of the fundamental determinants of stock returns.

A buoyant economy replete with lucrative opportunities encourages companies to expand, while plunging demand and plummeting profits induce companies to contract. Assuming that firms are made of assets in place and growth opportunities potentially misrepresents the fact that firms are able to scale up or down in response to the uncertain operating environment. In this paper, I employ a neoclassical model of a firm that derives its profits from existing assets and has the flexibility to increase or decrease its capital stock. Thus, the value of the firm has three components: assets in place, investment options, and disinvestment options. The present value of all future profits generated by a firm's current capital makes up the value of assets in place. I follow the convention that investment options are contingent claims on assets in place so that they are riskier and earn higher expected returns. Disinvestment options, on the other hand, serve as a hedge for existing assets amid deteriorating business conditions and deserve a lower expected return. A firm's profits and its real options are inextricably linked. High profitability drives firms' further expansion, fostering more opportunities to invest. Hence, the positive profitability premium reflects the return difference between firms owning more investment options and those with fewer investment options.

A notable feature in data is that gross profitability (GPA) and market valuation ratio (MB) are positively correlated. The introduction of disinvestment options with two dimensions of firm-level heterogeneity helps tackle the challenge. In my model, firms face stochastic productivity and stochastic costs of investment. More productive firms are more profitable, endowed with more investment options and valued high by the market. This channel preserves a positive relation between GPA and MB. When firms have the flexibility to scale down, rising costs of investment increase the valuation ratio of growth firms by increasing the value of disinvestment options. Growth firms face opportunities aplenty outside their own venture and are more likely to scale down and allocate the resource to external productive businesses. This makes the value of disinvestment options a significant portion for growth firms, reducing the expected return on growth firms and addressing the value premium. Therefore, my model fundamentally dismantles the one-to-one injection between GPA and MB in a cross section while still preserves a general positive relation between them, and reconciles the coexistence of a positive profitability premium and a positive value premium.

My model quantitatively reproduces the profitability premium, the value premium, and the negative relation between gross profitability and the book-to-market ratio. Employing the simulated method of moments (SMM), I estimate three parameters, including the discount of capital resale price, the price of aggregate demand risk, and the price of investment price risk, targeting seven empirical moments, including market excess return, average profitability premium, average value premium, GPA spread of GPA sorted decile portfolios, logBM spread over GPA sorted decile portfolios, GPA spread over BM sorted decile portfolios, and logBM spread over BM sorted decile portfolios. My model generates an annualized profitability premium of 5.2% with t-statistic of 14.5 and an annualized value premium of 4.6% with t-statistic of 11.9, close to the empirical values of 5.3% for the profitability premium and 4.8% for the value premium over the period of 1964 to 2020. Furthermore, my model retains the negative relation between GPA and logBM and replicates the failure of CAPM.

In my model, the implicit value of marginal productivity of capital ($\frac{F}{C}$), defined as the ratio of marginal product of capital (F) to marginal cost of investment (C), is a sufficient statistic of the expected return (ER). In my benchmark parametrization, the expected return is a concave function increasing in the implicit capital productivity. Profitable firms and value firms have higher implicit capital productivity in the cross section and so high expected returns. Among less profitable firms and growth firms (low $\frac{F}{C}$ firms), the expected return is more sensitive to firm characteristics. This is because the slope of $ER - \frac{F}{C}$ curve is steeper in the low $\frac{F}{C}$ region, where the vertical difference, representing the return spread, corresponding to a same horizontal difference, representing the $\frac{F}{C}$ spread, is magnified. This explains why the profitability premium is stronger within growth firms and why the value premium is stronger within less profitable firms.

My model also provides an explanation for the cyclicality of the two premiums. The profitability premium is procyclical (see Kogan, Li, and Zhang (2021)) and the value premium is countercyclical (see Petkova and Zhang (2005) and Zhang (2005)). The dispersion of profitability is larger in good economic states, creating stronger return spread across GPA sorted portfolios. The non-linearity of return prediction in my model indicates that expected returns are more sensitive to market valuation ratios during bad times. Hence the predictability of market valuation ratios on stock returns is mainly concentrated in economic downturns.

Recent works trying to explain the coexistence of the profitability premium and the value premium include Kogan and Papanikolaou (2013), Ma and Yan (2015), and Kogan et al. (2021). The main difference between my paper and previous studies is that I explicitly explore the role of disinvestment options. Existing studies generally model the arrivals and departures of growth perspectives to be exogenous so firm value is the sum of the value of assets in place and the present value of growth opportunities. When one component, say assets in place, yields a higher return than the other, say growth opportunities, high-return companies have more assets in place than growth opportunities, and this implies a counterfactual positive relation between book-to-market ratio and gross profitability. Ai and Kiku (2013), Cooper (2006), and Hackbarth and Johnson (2015) are prior studies focusing on the link between real options and either the value premium or the profitability premium. Nevertheless, the lack of firm-level heterogeneity could go against the other premium.

More broadly, my paper is related to the literature on firms' investment decisions under uncertainty. Early works include Abel (1983), Abel and Eberly (1993), Abel and Eberly

(1996), McDonald and Siegel (1986), and Dixit and Pindyck (2012). I follow Bertola (1988) and Bertola (1998) in setting up and solving the model. I extend investment theories to analyzing their asset pricing implications.

The rest of this paper proceeds as follows. Section 2 describes data sources, variable construction, and motivating empirical evidence. Section 3 introduces the continuous-time production-based model with its analytical solution. Section 4 discusses the asset pricing implications of the model. Calibration, estimation, and time-series simulation are presented in Section 5. I inspect the mechanism in Section 6. I conclude in Section 7.

2 Data and Empirical Motivation

In this section, I first describe data sources and variable construction. Then I present main empirical results of univariate sorts, double sorts, and Fama-MacBeth regressions to motivate the structural model.

2.1 Data and variable definitions

Monthly stock return data are from the Center for Research in Security Prices (CRSP) database, and firm-level accounting data are from Compustat annual database. I only include common stocks (CRSP item SHRCD = 10 or 11) traded in NYSE, AMEX, and NASDAQ (CRSP item EXCHCD = 1, 2, or 3). I remove financial firms (SIC between 6000 and 6999) and utility firms (SIC between 4950 and 4999). The sample period is from July 1964 to June 2020.

Following Novy-Marx (2013), I define profitability (GPA) as gross profit (Compustat item GP) divided by total assets (Compustat item AT). Following Fama and French (1993), I define book-to-market ratio (BM) as the book value of common equity (Compustat item CEQ) plus deferred taxes (Compustat item TXDB) and investment tax credit (Compustat item ITCB), minus the book value of preferred stock, calculated as the coalescence in the order of redemption value (Compustat item PSTKRV), liquidation value (Compustat item PSTKL), and par value (Compustat item PSTK), divided by market equity (CRSP December market capitalization), divided by the market capitalization at the end of the previous fiscal year.

2.2 Decile portfolios from univariate sorts

Novy-Marx (2013) documents a gross profitability premium that firms with higher profitability have higher average returns. This seems to be at odds with the well-known value premium that value stocks (firms with high BM) tend to outperform growth stocks (firms with low BM), as more profitable firms are seemingly analogous with growth firms. As such, the profitability premium is dubbed "the other side of value" in Novy-Marx (2013). I first replicate the two anomalies in an updated sample. At the end of June of each year from 1964 to 2020, firms are allocated into deciles based on their previous fiscal year end gross profitability (GPA) and book-to-market ratio (BM), respectively. The portfolio characteristics [3], value-weighted returns, and results of asset pricing tests are reported in Tables 1 and 2.

[Insert Table 1 Here]

Panel A of Table 1 reports characteristics of GPA decile portfolios. I find that firms with high gross profitability are similar to growth firms in that they have low BM, high past return, high Tobin's Q, high investment ratio, and high cash holdings. Panel B of Table 1 reports the value-weighted returns, results of CAPM and Fama-French three-factor model tests for the 10 GPA portfolios and the long-short strategy. More profitable firms have higher

[3] Portfolio characteristics include GPA, logBM, logME (logarithm of market capitalization at the end of June), Mom (cumulative return over past twelve to two months), Q (Tobin's Q, defined as the sum of market capitalization, long-term debt (Compustat item DLC), preferred stock redemption value (Compustat item PSTKRV) minus total inventories (Compustat item INVT) and deferred tax in balance sheet (Compustat item TXDB), divided by gross property, plant and equipment (Compustat item PPEGT)), IK (investment-to-capital ratio, defined as capital expenditure (Compustat item CAPX) divided by PPEGT), and CHK (cash holdings, defined as cash and short-term investments (Compustat item CHE) divided by PPEGT).

excess returns than do less profitable firms, and the average annual return spread is 5.32% (t-statistic = 2.40). Moreover, the profitability premium cannot be explained by the CAPM or the Fama-French three-factor model, as the CAPM α spread is 7.31% (t-statistic = 3.29) and the FF3F α spread is 8.49% (t-statistic = 4.04). The spread is magnified when the SMB factor and the HML factor are included in the time-series regression due to the negative loadings on the two factors (the exposure to SMB is -0.39 with t-statistic = -5.04, and the exposure to HML is -0.20 with t-statistic = -2.04).

[Insert Table 2 Here]

Table 2 presents characteristics and returns of decile portfolios formed on BM. As can be seen from Panel A, from decile 1 to decile 10, profitability, size, past return, Tobin's Q, investment ratio, and cash holdings generally decrease, which suggests resemblance between growth (value) firms and more profitable (less profitable) firms. Panel B shows the annualized return spread between value stocks and growth stocks is 4.78% (t-statistic = 2.04). Though not captured by the CAPM, the value premium can be completely absorbed when Fama-French three factors (particularly the HML factor) are controlled.

One observation from Table 1 is that both logBM and the loadings on the HML factor exhibit a hump-shaped pattern across 10 portfolios formed on GPA. In other words, the relation between GPA and BM is not necessarily linear. High GPA firms are alike growth firms, and so are those with the lowest GPA. Low BM firms should comprise those with high GPA as well as some small firms with low profitability, but the latter are tucked away in the lowest BM decile in Table 2 for they are dominated by other larger and more profitable firms. It can be deceptively drawn from univariate sorts that there is a close correspondence between more profitable (less profitable) firms and growth (value) firms.

2.3 The interplay of profitability and book-to-market

To explore the different information contained in gross profitability and book-to-market and how they predict stock returns interactively, I first conduct a double sort on BM and GPA. Table 3 reports the average excess returns and characteristics of the 25 portfolios. The relation between GPA and excess return is positive in all five BM groups. While the GPA spread is comparable across the five BM groups, the return spread is larger among firms with lower BM. Similarly, the relation between BM and excess return is also positive when GPA is controlled, and it is stronger among firms with lower GPA though the dispersion in logBM is not larger there.

[Insert Table 3 Here]

I then construct 25 portfolios sorted on gross profitability only and report the results in Table 4. Once again, I observe a positive relation between GPA and excess return; the finer partition, in addition, reveals a concave relation between GPA and excess return, namely that the return does not increase as fast as does the profitability. Conventional wisdom claims a high GPA, synonymous with a low BM, signifies ample growth opportunities for a firm. If both GPA and BM contain similar information, the profitability premium would be stronger among low BM firms as well as high GPA firms, so a convex relation between GPA and excess return should be anticipated.

[Insert Table 4 Here]

I perform a set of Fama and MacBeth (1973) cross-sectional regressions to test whether book-to-market ratio or gross profitability magnifies the positive relation between GPA and future excess returns. This regression analysis allows me to control for other determinants of the cross section of stock returns, including market capitalization (ME) and cumulative return over past twelve to two months (Mom). The equation to estimate each month t is

$$R_{i,t} = b_{0,t} + b_{1,t} X_{i,t} \times GPA_{i,t} + b_{2,t} GPA_{i,t} + b_{3,t} \log(B/M)_{i,t} + b_{4,t} \log(ME)_{i,t} + b_{5,t} Mom_{i,t} + \epsilon_{i,t} \tag{1}$$

$R_{i,t}$ is the month t excess return on stock i. The subscript t of the explanatory variables represents the timing of accounting variables consistent with portfolio sorts conducted before. $X_{i,t}$ stands for $\log(B/M)_{i,t}$ in specification (1) and $GPA_{i,t}$ in specification (2). The results are presented in Table 5.

[Insert Table 5 Here]

In specification (1), I find the coefficient of the interaction term $\log(B/M)_{i,t} \times GPA_{i,t}$ is significantly negative. This indicates that the predictability of profitability on returns decreases as book-to-market increases. The coefficient of the term $GPA_{i,t} \times GPA_{i,t}$ is significantly negative in specification (2). This implies that the positive relation between profitability and return diminishes as profitability increases. The results of the Fama-MacBeth regressions reinforce the findings of double sorts that the profitability premium is prominent among firms with low BM and low GPA. To sum up, profitability and book-to-market are interconnected but they reveal different aspects of firms. In the next section, I introduce a structural model of investment which is able to reconcile the stylized facts described above.

3 The Model Set-up

In this section, I develop a continuous-time model of a representative firm making investment decisions under uncertainty. I make reasonable assumptions to obtain a closed-form solution for the value of the firm.

3.1 Production technology

Assume that a representative firm i's maximum gross profit, defined as total revenue minus variable costs evaluated at the optimal quantity of perfectly flexible inputs, at instant t is given by

$$\Pi_{i,t} = \frac{1}{1+\psi} X_t^{\phi} Z_{i,t}^{\phi-1} K_{i,t}^{1+\psi}, \quad -1 < \psi < 0 \text{ and } \phi > 1 \tag{2}$$

where K denotes the book value of physical capital, X is the aggregate demand, Z is the firm-specific productivity, and ψ is the capital curvature index and ϕ is profit's elasticity to the aggregate demand. ψ and ϕ are two constants depicting decreasing returns to scale. [4] Unless the firm adjusts physical capital, K changes only due to depreciation

$$\mathrm{d}K_t = -\delta K_t \, \mathrm{d}t \tag{3}$$

where δ denotes the rate of depreciation. Both X and Z are random variables and evolve according to a geometric Brownian motion

$$\frac{\mathrm{d}X}{X} = \mu_X \, \mathrm{d}t + \sigma_X \, \mathrm{d}w_X \tag{4}$$

$$\frac{\mathrm{d}Z}{Z} = \mu_Z \, \mathrm{d}t + \sigma_Z \, \mathrm{d}w_Z \tag{5}$$

where $\mathrm{d}w_X$ and $\mathrm{d}w_Z$ are increments of two independent standard Wiener processes. Let $A_{i,t} = X_t^{\phi} Z_{i,t}^{\phi-1}$ encapsulate the systematic and idiosyncratic operation conditions for firm i at time t. A is higher when the aggregate demand X is higher and when the representative firm's productivity Z is higher. It follows that A can be described by the stochastic process

$$\frac{\mathrm{d}A}{A} = \mu_A \, \mathrm{d}t + \sigma_A \, \mathrm{d}w_A \tag{6}$$

[4] In Appendix A, I show that this is the case where a firm produces with physical capital and flexible inputs according to a Cobb-Douglas production function with decreasing returns to scale and both the price and the productivity of its flexible inputs are stochastic.

where

$$
\begin{aligned}
\mu_A &\equiv \phi\mu_X + (\phi-1)\mu_Z + \frac{\phi(\phi-1)}{2}\sigma_X^2 + \frac{(\phi-1)(\phi-2)}{2}\sigma_Z^2 \\
\sigma_A &\equiv \sqrt{\phi^2\sigma_X^2 + (\phi-1)^2\sigma_Z^2} \\
\mathrm{d}w_A &\equiv \frac{\phi\sigma_X\,\mathrm{d}w_X + (1-\phi)\sigma_Z\,\mathrm{d}w_Z}{\sigma_A}
\end{aligned}
$$

3.2 Valuation

The firm incurs a random cost whenever it adjusts capital, and investment is characterized by costly reversibility, where capital is worth less when being sold. Let C denote the exogenous cost of adjusting a unit of capital. C is the product of market price of investment Y and firm-specific investment price multiplier U. When the firm is setting up new capital, U is high if the firm is inefficient and uses more resources; when the firm is winding down and selling existing capital, U is high if the firm can find a buyer willing to pay a high price. Then for firm i at instant t,

$$C_{i,t} = Y_t U_{i,t} \tag{7}$$

Firm i pays an effective price of C_i to put a new unit of capital stock in place, and installed capital can be resold at unit price of θC_i ($0 < \theta < 1$). Costly reversibility is captured by the economy-wide resale discount θ. When $\theta = 1$, there is no resale discount. When $\theta = 0$, all investment costs are sunk costs and irreversible. There is no option to disinvest, as is the case in Cooper (2006) and Kogan and Papanikolaou (2013).

I assume that Y and U follow a geometric Brownian motion

$$\frac{\mathrm{d}Y}{Y} = \mu_Y\,\mathrm{d}t + \sigma_Y\,\mathrm{d}w_Y \tag{8}$$

$$\frac{\mathrm{d}U}{U} = \mu_U\,\mathrm{d}t + \sigma_U\,\mathrm{d}w_U \tag{9}$$

where $\mathrm{d}w_Y$ and $\mathrm{d}w_U$ are increments of two independent standard Wiener processes. Then the process of C can be expressed as follows

$$\frac{\mathrm{d}C}{C} = \mu_C \,\mathrm{d}t + \sigma_C \,\mathrm{d}w_C \tag{10}$$

where

$$\begin{aligned} \mu_C &\equiv \mu_Y + \mu_U \\ \sigma_C &\equiv \sqrt{\sigma_Y^2 + \sigma_U^2} \\ \mathrm{d}w_C &\equiv \frac{\sigma_Y \,\mathrm{d}w_Y + \sigma_U \,\mathrm{d}w_U}{\sigma_C} \end{aligned}$$

Further assume that the increments of X and Y are independent and that ρ_{ZU} indexes the correlation between $\mathrm{d}w_Z$ and $\mathrm{d}w_U$. This allows me to replace the four primitive state variables X, Y, Z, and U with two summarized variables A and C: μ_A and σ_A are functions of μ_X, σ_X, μ_Z, and σ_Z; μ_C and σ_C are functions of μ_Y, σ_Y, μ_U, and σ_U; $\rho_{AC} \equiv (\phi - 1)\frac{\rho_{ZU}\sigma_Z\sigma_U}{\sigma_A\sigma_P}$ reflects the dependence between Z and U.

I assume all firms are risk-neutral when solving the model[5], so the value of a representative firm at instant t can then be conveniently written as the maximized expected present value of cash flows discounted at a constant risk-free rate r

$$V(A_t, C_t, K_t) = \max_{I_t} \left\{ \mathop{\mathrm{E}}_t \int_t^\infty e^{-r(\tau - t)} \left(\frac{1}{1+\psi} A_\tau K_\tau^{1+\psi} \,\mathrm{d}\tau - C_\tau \,\mathrm{d}I_\tau \right) \right\} \tag{11}$$

$$\text{s.t.} \quad \mathrm{d}K_t = -\delta K_t \,\mathrm{d}t + C_t \,\mathrm{d}I_t \tag{12}$$

Define $F(A_t, K_t)$ as the marginal product of capital and $Q(F_t, C_t)$ as the marginal value

[5] Constantinides (1978) develops the risk-neutral valuation rule. Cooper (2006) proves that risk premiums are positive under the risk-neutral assumption because firm values comove with aggregate risks.

of capital. That is,

$$F(A_t, K_t) \equiv \frac{\partial \Pi_t}{\partial K_t} = A_t K_t^{\psi} \tag{13}$$

$$Q(F_t, C_t) \equiv \frac{\partial V_t}{\partial K_t} = \mathop{\mathrm{E}}_{t} \int_t^{\infty} e^{-(r+\delta)(\tau - t)} F_\tau \, \mathrm{d}\tau \tag{14}$$

The necessary conditions to solve for the optimal investment policy I_t are

$$\mathrm{d}I_t > 0 \quad \text{if } Q_t > C_t \tag{15}$$

$$\mathrm{d}I_t = 0 \quad \text{if } \theta C_t \leq Q_t \leq C_t \tag{16}$$

$$\mathrm{d}I_t < 0 \quad \text{if } Q_t < \theta C_t \tag{17}$$

In words, the firm purchases new capital only when the marginal value of capital Q exceeds the marginal cost C, and the firm sells existing capital only when the marginal value of capital Q falls below the marginal cost θC. When Q is in the range of $[\theta C, C]$, the firm neither invests nor disinvests, and physical capital decreases only via depreciation. Since the model does not incorporate other adjustment frictions, whenever exogenous variations in A or C are about to bring Q outside the inaction range, the firm is able to make small adjustments instantaneously to keep Q within the inaction region $[\theta C, C]$.

The existence of the inaction region hinges on the assumption that $\theta < 1$. When $\theta = 1$, conditions (15) and (17) are equivalent, so the firm is continuously investing or disinvesting to maintain the equality between marginal value of capital Q and marginal cost of capital C. Since both investment (a negative dividend payment) and disinvestment (a positive dividend payment) do not affect firm value, it can be inferred that all firms in this economy have a same static expected return. When $\theta = 0$, $\mathrm{d}I \geq 0$ always holds. This is the case of irreversibly. Firms invest when marginal value of capital Q passes marginal cost of capital C, and the diminution of capital stock is only due to depreciation.

Following McDonald and Siegel (1986), I conjecture that Q is homogeneous of degree one in F and C so that the inaction region defined in (16) associated with Q and C is now prescribed by F and C:

$$b\theta C \leq F \leq hC \tag{18}$$

where b and h (with $0 < b < h$)[6] are two unknown constants solved in Appendix B that regulate the ratio of marginal product of capital F to effective unit cost of purchasing new capital C. As a result, doubling a given pair of values of F and C shall not alter the firm's investment decision, since the marginal value of capital Q is also doubled and the relation between Q and C is preserved.

I derive in Appendix B the total value of the firm at instant t as

$$V(A_t, C_t, K_t) = D_0 A_t K_t^{1+\psi} + D_1 A_t^{\eta_1} C_t^{1-\eta_1} K_t^{1+\psi\eta_1} + D_2 A_t^{\eta_2} C_t^{1-\eta_2} K_t^{1+\psi\eta_2} \tag{19}$$

where parameters $D_0 > 0$, $D_1 > 0$, $D_2 > 0$, $\eta_1 > 1$, and $\eta_2 < 0$ are described in the appendix. In equation (19), the first term $D_0 A_t K_t^{1+\psi}$ is the value of assets in place ($V_{AP,t}$), the second term $D_1 A_t^{\eta_1} C_t^{1-\eta_1} K_t^{1+\psi\eta_1}$ is the value options to invest ($V_{IO,t}$), and the third term $D_2 A_t^{\eta_2} C_t^{1-\eta_2} K_t^{1+\psi\eta_2}$ is the value of options to disinvest ($V_{DO,t}$). Their properties are studied in the following section.

4 Asset Pricing Implications

In this section, I use the firm value (equation (19)) solved in Section 3 to study the relation between firms' asset composition and their expected returns.

[6] $0 < b < h$ because $0 < \theta < 1$. When $\theta = 1$, $b = h$; when $\theta = 0$, $b = 0$.

4.1 Asset composition

The three components of total firm value respond differently to two economy-wide shocks - aggregate demand X and investment price Y, which in turn determine firm's overall risk exposure. Assets in place load only on X_t with the exposure

$$\beta_X^{AP} = \frac{\partial \log V_{AP,t}}{\partial \log X_t} = \phi > 1 \tag{20}$$

This is equal to the exposure of optimized gross profit to X_t, as is discussed in Appendix A. The value of existing assets is the present value of all expected future profits generated by the current level of capital stock allowing for depreciation but ruling out contingent actions by the firm to scale it up or down in the future. Investment options and disinvestment options load on both X_t and Y_t:

$$\beta_X^{IO} = \frac{\partial \log V_{IO,t}}{\partial \log X_t} = \eta_1 \phi > \phi, \quad \beta_Y^{IO} = \frac{\partial \log V_{IO,t}}{\partial \log Y_t} = 1 - \eta_1 < 0 \tag{21}$$

$$\beta_X^{DO} = \frac{\partial \log V_{DO,t}}{\partial \log X_t} = \eta_2 \phi < 0, \quad \beta_X^{DO} = \frac{\partial \log V_{DO,t}}{\partial \log Y_t} = 1 - \eta_2 > 0 \tag{22}$$

An investment option can be interpreted as a call option on firm's assets in place - in particular, a long position in the present value of future profits and a short position in investment costs. Exercising an investment option involves exchanging more assets in place with total investment costs as the strike price. As investment costs depend on Y and are independent of X and assets in place load only on X, paying the strike price levers up the exposure of investment options to X and creates a hedge to Y. Hence, $\beta_X^{IO} > \beta_X^{AP}$ and $\beta_Y^{IO} < 0$. This is consistent with the intuition that the investment option is worth more and more likely to exercise when the firm is in better business condition (high X) and faces lower exercising costs (low Y). Similarly, the disinvestment option, a put option on firm's assets in place, comes into play amid deteriorating business conditions (low X) or exorbitant costs of investment (high Y).

The weights of assets in place, investment options, and disinvestment options are given by

$$\omega_{AP,t} \equiv \frac{D_0 F_t/C_t}{D_0 F_t/C_t + D_1 (F_t/C_t)^{\eta_1} + D_2 (F_t/C_t)^{\eta_2}} \tag{23}$$

$$\omega_{IO,t} \equiv \frac{D_1 (F_t/C_t)^{\eta_1}}{D_0 F_t/C_t + D_1 (F_t/C_t)^{\eta_1} + D_2 (F_t/C_t)^{\eta_2}} \tag{24}$$

$$\omega_{DO,t} \equiv \frac{D_2 (F_t/C_t)^{\eta_2}}{D_0 F_t/C_t + D_1 (F_t/C_t)^{\eta_1} + D_2 (F_t/C_t)^{\eta_2}} \tag{25}$$

respectively. All three weights are a function of the ratio of F_t to C_t and are time-varying. $\frac{F_{i,t}}{C_{i,t}}$ is the intrinsic state variable for firm i in my model. It is the ratio of marginal product of capital F to marginal cost of investment C, so it measures the *implicit* per dollar value of capital productivity. Plugging the expression of $F_{i,t}$ as in equation (13) and the expression of $C_{i,t}$ as in equation (7), I can expand $\frac{F_{i,t}}{C_{i,t}}$ as

$$\frac{F_{i,t}}{C_{i,t}} = \frac{X_t^{\phi} Z_{i,t}^{\phi-1} K_{i,t}^{\psi}}{Y_t U_{i,t}} \tag{26}$$

It is the product of an aggregate condition indicator $\frac{X_t^{\phi}}{Y_t}$ and a firm status indicator $\frac{Z_{i,t}^{\phi-1} K_{i,t}^{\psi}}{U_{i,t}}$. In Appendix C, I show that $\omega_{IO,t}$ is a monotonically increasing function of $\frac{F_t}{C_t}$ and that $\omega_{DO,t}$ is a monotonically decreasing function of $\frac{F_t}{C_t}$. This mirrors the previous result of their aggregate risk exposures and embodies firm-specific ingredients. In a given state of economy, more productive firms (high Z) operating with lower idiosyncratic investment costs (low U) have more investment options and fewer disinvestment options. The monotonicity of $\omega_{AP,t}$ depends on $\frac{F_t}{C_t}$. When $\frac{F_t}{C_t}$ is low, disinvestment options dominate. An increase in $\frac{F_t}{C_t}$ reduces disinvestment options and gives rise to more assets in place. The numerators of $\omega_{AP,t}$ and $\omega_{IO,t}$ suggest that investment options increase with $\frac{F_t}{C_t}$ at a rate higher than that of assets in place. When $\frac{F_t}{C_t}$ is sufficiently high, investment options take up the relative volume of assets in place and dominate the latter.

4.2 Risk premium

The two independent aggregate sources of risk (volatile total demand X and fluctuating investment price Y) with a constant risk-free rate r in the model economy motivate the following pricing kernel

$$\frac{\mathrm{d}m}{m} = -r\,\mathrm{d}t - \gamma_X\,\mathrm{d}w_X - \gamma_Y\,\mathrm{d}w_Y \tag{27}$$

Both the demand shock and the price shock are priced and have constant prices of risk γ_X and γ_Y, respectively.

The firm's risk premium is determined by its covariance with the stochastic discount factor. In the model, the exposure of assets in place, investment options, and disinvestment options to systematic risks are not time-varying, so the firm's expected excess return (ER_t) can be written as a weighted average of the expected excess returns on assets in place (ER_{AP}), investment options (ER_{IO}), and disinvestment options (ER_{DO}), i.e.,

$$ER_t = \omega_{AP,t}ER_{AP} + \omega_{IO,t}ER_{IO} + \omega_{DO,t}ER_{DO} \tag{28}$$

where

$$ER_{AP} = \phi\gamma_X\sigma_X \tag{29}$$

$$ER_{IO} = \eta_1\phi\gamma_X\sigma_X + (1-\eta_1)\gamma_Y\sigma_Y \tag{30}$$

$$ER_{DO} = \eta_2\phi\gamma_X\sigma_X + (1-\eta_2)\gamma_Y\sigma_Y \tag{31}$$

In Section 5.1, I estimate $\gamma_X > 0$ and $\gamma_Y < 0$[7]. Then

$$ER_{IO} - ER_{AP} = (\eta_1 - 1)(\phi\gamma_X\sigma_X - \gamma_Y\sigma_Y) > 0 \tag{32}$$

$$ER_{DO} - ER_{AP} = (\eta_2 - 1)(\phi\gamma_X\sigma_X - \gamma_Y\sigma_Y) < 0 \tag{33}$$

Investment options are contingent claims on future profits, so on average they are riskier than assets in place and require a higher return. Disinvestment options, on the other hand, pay out in recession and serve as a hedge against economic bat times; for this reason they deserve a lower expected return. Hence, the expected excess return on a firm is related to its asset composition. The more investment options and the fewer disinvestment options a firm commands, the higher its risk premium.

Plugging weights from equations (23), (24), and (25) and risk premiums from equations (29), (30), and (31) into equation equation (28), I derive in Appendix C that ER_t is an increasing function of $\frac{F_t}{C_t}$. If $\frac{F_t}{C_t}$ belongs to the set of all positive real numbers, the expected excess return must be bounded below by ER_{DO} when disinvestment options dominate the firm value and bounded up by ER_{IO} when investment options dominate the firm value. As ER_t is a smooth function, there exists a convex region when $\frac{F_t}{C_t}$ is low and a concave region when $\frac{F_t}{C_t}$ is high. As such, the sensitivity of ER_t to $\frac{F_t}{C_t}$ depends on $\frac{F_t}{C_t}$ per se: in the convex region, the change in ER_t induced by a marginal change in $\frac{F_t}{C_t}$ grows as $\frac{F_t}{C_t}$ increases; as $\frac{F_t}{C_t}$ continues to increase to the concave region, the change in ER_t caused by a unit change in $\frac{F_t}{C_t}$ diminishes. Hence, the magnitude of expected return spread derived from the spread of any firm characteristics is subject to the level of the elemental state variable $\frac{F_t}{C_t}$, which consists of both systematic and idiosyncratic contents. In my model, $\frac{F_t}{C_t}$ is restricted in the range of $[b\theta, h]$, so the practical convexity of ER_t depends on the model parameterization.

[7]Appendix C shows that $\gamma_Y < 0$ is a sufficient condition to generate the return patterns discussed here. The estimate of γ_Y from the simulated method of moments is -0.598 with a standard error of 0.010. This is also consistent with the observation in data that the relative price of investment goods covaries negatively with the business cycle (see Li (2018)).

4.3 The profitability premium and the value premium

Each component in the expected return (see equation (28)) contains a time-invariant expected return and a time-varying weight, suggesting that expected returns of long-short portfolios constructed from sorting firm characteristics reflect differences in three constituent weights. The expected profitability premium (PMU_t) is the difference between returns on profitable companies (those with high profitability, measured empirically by GPA, whose aggregate return at time t is ER_t^P) and unprofitable companies (those with low GPA whose aggregate return at time t is ER_t^U)

$$\begin{aligned} PMU_t &= ER_t^P - ER_t^U \\ &= [(\eta_1 - 1)(\omega_{IO,t}^P - \omega_{IO,t}^U) + (\eta_2 - 1)(\omega_{DO,t}^P - \omega_{DO,t}^U)](\phi\gamma_X\sigma_X - \gamma_Y\sigma_Y) \end{aligned} \tag{34}$$

The expected value premium (VMG_t) is the difference between returns on value stocks (those with low valuation ratio, measured empirically by market-to-book ratio or MB, whose aggregate return at time t is ER_t^V) and growth stocks (those with high MB whose aggregate return at time t is ER_t^G)

$$\begin{aligned} VMG_t &= ER_t^V - ER_t^G \\ &= [(\eta_1 - 1)(\omega_{IO,t}^V - \omega_{IO,t}^G) + (\eta_2 - 1)(\omega_{DO,t}^V - \omega_{GO,t}^U)](\phi\gamma_X\sigma_X - \gamma_Y\sigma_Y) \end{aligned} \tag{35}$$

Given $\eta_1 > 1$, $\eta_2 < 0$, and $\phi\gamma_X\sigma_X - \gamma_Y\sigma_Y > 0$, a sufficient condition to generate an ex-ante positive profitability premium and an ex-ante positive value premium states that

$$\omega_{IO,t}^P > \omega_{IO,t}^U \quad \text{and} \quad \omega_{DO,t}^P < \omega_{DO,t}^U, \quad \forall t \tag{36}$$

$$\omega_{IO,t}^V > \omega_{IO,t}^G \quad \text{and} \quad \omega_{DO,t}^V < \omega_{DO,t}^G, \quad \forall t \tag{37}$$

Since investment options earn the highest expected return in my model, firms with higher returns (profitable firms and value firms) are those with more investment options. Berk, Green, and Naik (1999) and Gomes, Kogan, and Zhang (2003) claim that value firms have more *growth* options than assets in place to explain the positive value premium, for their models ignore disinvestment options and suggest that *growth* options demand a higher return than do assets in place. However, as firms are endowed with the opportunity to dispose of capital in my model, the wedge between a high market value and a low book value among growth firms comes from disinvestment options, which is indicated by condition (37). Separating the two types of real options, my model does not imply that growth firms are not option-intensive, since combining investment options and disinvestment options still allows growth firms to own more options than assets in place.

The closed-form expression of the firm value in equation (19) can be implemented to dissect the two firm characteristics, defined in the model as

$$\mathrm{GPA} = \frac{\Pi}{K} = \frac{AK^{\psi}}{1+\psi} = \frac{F}{1+\psi} \tag{38}$$

$$\begin{aligned} \mathrm{MB} &= \frac{V}{K} = D_0 A K^{\psi} + D_1 A^{\eta_1} C^{1-\eta_1} K^{\psi\eta_1} + D_2 A^{\eta_2} C^{1-\eta_2} K^{\psi\eta_2} \\ &= D_0 F + D_1 F^{\eta_1} C^{1-\eta_1} + D_2 F^{\eta_2} C^{1-\eta_2} \end{aligned} \tag{39}$$

In the cross section, aggregate variables X and Y are fixed. Equation (38) suggests that GPA is an increasing function of Z. As $\frac{F}{C}$ also increases with Z, more profitable (high Z) firms have higher $\frac{F}{C}$ and earn higher returns, which explains the positive profitability premium. Simple derivation from equation (39) shows that MB is an increasing function of both Z and U. MB and GPA are therefore positively correlated due to their common reliance on Z. However, if disinvestment options were ignored, increasing Z increases $\frac{F}{C}$ and MB as $\frac{\partial \mathrm{MB}}{\partial F} = D_0 + \eta_1 D_1 \left(\frac{F}{C}\right)^{\eta_1 - 1} > 0$, implying a negative value premium. The inclusion of disinvestment options helps generate the positive value premium, for increasing U decreases $\frac{F}{C}$ but increases MB given $\frac{\partial \mathrm{MB}}{\partial C} = (1-\eta_1) D_1 \left(\frac{F}{C}\right)^{\eta_1} + (1-\eta_2) D_2 \left(\frac{F}{C}\right)^{\eta_2} > 0$. The intuition is that rising resale price increases the valuation ratio by increasing the value of their disinvestment options. Section 6.1 confirms that the model generates a positive value premium in simulation. To sum up, my model fundamentally dismantles the one-to-one injection between GPA and MB in a cross section with the second source of heterogeneity U while still preserves a general positive relation between them via Z, and reconciles the coexistence of a positive profitability premium and a positive value premium.

The relation between GPA, MB and Z, U in the cross section can also be used to deduce why some characteristic-based premiums are more prominent within certain firms following analysis in Section 4.2. As long as the vertical spread, representing the ER spread, corresponding to a same horizontal spread, representing the $\frac{F}{C}$ spread or equivalently the characteristic spread, can be magnified in some regions, where there are growth firms and unprofitable firms, the characteristic-based premiums would be larger.

Another empirical property of the profitable premium and the value premium is that they covary differently with the business cycle. Kogan et al. (2021) record a procyclical profitability premium, and Petkova and Zhang (2005) and Zhang (2005) document a countercyclical value premium. In my model, the unconditional covariance between the two premiums is given by

$$\begin{aligned}
\mathrm{Cov}(PMU_t, VMG_t) = & \mathrm{Cov}\left((\eta_1 - 1)(\omega^P_{IO,t} - \omega^U_{IO,t}) + (\eta_2 - 1)(\omega^P_{DO,t} - \omega^U_{DO,t}),\right. \\
& \left.(\eta_1 - 1)(\omega^V_{IO,t} - \omega^G_{IO,t}) + (\eta_2 - 1)(\omega^V_{DO,t} - \omega^G_{DO,t})\right) \\
= & (\eta_1 - 1)^2 \,\mathrm{Cov}\left(\omega^P_{IO,t} - \omega^U_{IO,t}, \omega^V_{IO,t} - \omega^G_{IO,t}\right) \\
& + (\eta_2 - 1)^2 \,\mathrm{Cov}\left(\omega^P_{DO,t} - \omega^U_{DO,t}, \omega^V_{DO,t} - \omega^G_{DO,t}\right) \\
& + (\eta_1 - 1)(\eta_2 - 1) \,\mathrm{Cov}\left(\omega^P_{IO,t} - \omega^U_{IO,t}, \omega^V_{DO,t} - \omega^G_{DO,t}\right) \\
& + (\eta_1 - 1)(\eta_2 - 1) \,\mathrm{Cov}\left(\omega^P_{DO,t} - \omega^U_{DO,t}, \omega^V_{IO,t} - \omega^G_{IO,t}\right)
\end{aligned} \tag{40}$$

Their different patterns of moving with the business cycle can be produced from the following two conditions

$$\mathrm{Cov}\left(\omega^{P}_{IO,t}-\omega^{U}_{IO,t},\omega^{V}_{IO,t}-\omega^{G}_{IO,t}\right)<0,\quad \forall t \tag{41}$$

$$\mathrm{Cov}\left(\omega^{P}_{DO,t}-\omega^{U}_{DO,t},\omega^{V}_{DO,t}-\omega^{G}_{DO,t}\right)<0,\quad \forall t \tag{42}$$

Conditions (41) and (42) suggest that the contrast in investment (disinvestment) options owned by more profitable firms versus less profitable firms moves in the opposite direction with the same contrast from value firms versus growth firms. In other words, during periods when more profitable firms have significantly more/fewer investment (disinvestment) options than do less profitable firms, the surplus/deficit in investment options owned by value firms versus growth firms is negligible, and when the difference of investment (disinvestment) options between more profitable firms and less profitable firms is tiny, that between value firms and growth firms is large. Section 6.3 confirms these conditions in simulation.

5 Quantitative Evaluation

I first calibrate and estimate the model in this section. Then I present the time-series simulation results under benchmark parameterization.

5.1 Model calibration and structural estimation

Table 6 summarizes the annual value of sixteen parameters. Panel A in Table 6 reports thirteen parameters guided by the existing literature. I first set the risk-free rate r_f to 0.07, the capital curvature index ψ to -0.07, the elasticity to aggregate demand ϕ to 1.67, and the annual depreciation rate δ to 0.1, the drift of aggregate demand μ_X to 0.012, the volatility of aggregate demand σ_X to 0.15, the drift of market price of investment μ_Y to 0.096, the volatility of market price of investment σ_Y to 0.06 following previous studies on investment under uncertainty. I then restrict the drift of idiosyncratic productivity μ_Z and the correlation between idiosyncratic productivity and idiosyncratic investment cost to 0 and set the volatility of idiosyncratic productivity σ_Z to 0.001, the drift of idiosyncratic investment cost μ_U to -0.03, and the volatility of idiosyncratic investment cost σ_U to 0.01 to match the aggregate level of firm characteristics.

[Insert Table 6 Here]

I estimate the remaining three parameters using the simulated method of moments. Denote Ψ as the vector of moments in the data and $\hat{\Psi}(\mathbf{p})$ as the vector of moments from the simulated sample generated with a set of parameters $\mathbf{p}$. The parameter vector $\mathbf{p}$ is estimated from the following minimization problem

$$\hat{\mathbf{p}} = \underset{\mathbf{p}}{\operatorname{argmin}}[\hat{\Psi}(\mathbf{p}) - \Psi]' W [\hat{\Psi}(\mathbf{p}) - \Psi] \tag{43}$$

where W is the weighting matrix. I simulate the model 100 times with 800 firms over 100 years at a weekly frequency and then accrue or re-sample all variables to an annual frequency to match the empirical moments. Following Bloom (2009), I use the annealing algorithm to solve the minimization problem and obtain the global minimum.

Panel B in Table 6 reports the estimation results. I estimate the discount of capital resale price θ to 0.414 (with a standard error of 0.006). θ has to be large enough to generate a positive profitability premium and a positive value premium while not close to one in order to maintain a sizable inaction region to create firm-level dispersion. I set the two prices of risk γ_X and γ_Y to 0.632 (with a standard error of 0.009) and -0.598 (with a standard error of 0.010), respectively, to match market excess return, average profitability premium, and average value premium.

[Insert Table 7 Here]

Table 7 compares simulation results with empirical data. My model reproduces the co-existence of the profitability premium and the value premium with a larger profitability premium among growth stocks and a larger value premium among less profitable firms. My model is also able to generate a negative relation between GPA and BM.

5.2 Time-series simulation

Table 8 reports the average annualized excess returns and results of CAPM test for ten gross profitability (GPA) deciles in Panel A and ten book-to-market (BM) deciles in Panel B from simulated data. My model is able to reproduce a positive profitability premium of 5.15% (compared to the 5.32% return spread in data) with t-statistic of 14.46 and a positive value premium of 4.64% (compared to the 4.78% return spread in data) with t-statistic of 11.94.

[Insert Table 8 Here]

Table 8 also demonstrates the failure of CAPM. The annualized CAPM alphas of the GPA long-short strategy and the BM long-short strategy are significant (3.61% with t-stat of 5.51 and 3.07% with t-stat of 4.41, respectively), while the beta across decile portfolios shows no monotonicity.

6 Results and Discussions

I first inspect the mechanism of the model at a cross section in subsection 6.1. Then I use the model to analyze the failure of CAPM in subsection 6.2 and the cyclicality of the profitability premium and the value premium in subsection 6.3.

6.1 Cross section of firms and returns

I inspect the mechanism of the model in this subsection, with the model solution visualized using a fixed set of X, Y, Z, U, and K whose ranges are from the time-series simulation in section 5.1. Section 4 argues that the ratio of marginal product of capital F to effective investment costs C is the intrinsic variable that indicates asset composition and risk premium. In this subsection, I first illustrate their relations, based on which I explore how stock returns are related with profitability (GPA) and market-to-book ratio (MB) in a cross section.

[Insert Figure 2 Here]

The top panel of Figure 2 plots the relation between weights of assets in place (ω_{AP}), investment options (ω_{IO}), and disinvestment options (ω_{DO}) and $\frac{F}{C}$. It shows that ω_{IO} monotonically increases with $\frac{F}{C}$ and that ω_{DO} monotonically decreases with $\frac{F}{C}$. When $\frac{F}{C}$ is close to zero, disinvestment options dominate all firm value. As $\frac{F}{C}$ increases, assets in place and investment options come out, and the former gives way to the latter at a high level of $\frac{F}{C}$, creating a hump-shaped relation between ω_{AP} and $\frac{F}{C}$. The bottom panel of Figure 2 portrays that expected return increases with $\frac{F}{C}$ and is bounded below on the left by ER_{DO} and up on the right by ER_{IO}. Based on my model calibration, the ER-$\frac{F}{C}$ relation is mainly concave over the feasible domain of $\frac{F}{C}$. Figure 2 authenticates the analysis in Section 4 with my parameters. Given the concavity between ER and $\frac{F}{C}$, the slope of ER-$\frac{F}{C}$ curve is steeper, which magnifies the vertical spread corresponding to a same horizontal difference. This explains why the profitability premium is stronger among growth firms and why the value premium is stronger among unprofitable firms.

[Insert Figure 3 Here]

Demonstrating how risk premium varies across different firms, Figure 3 compiles the time-series simulation results when aggregate state variables are fixed at the average. The top panel of Figure 3 plots firm's annual risk premium against two idiosyncratic variables - productivity Z and investment cost U, and shows that expected return increases with Z and decreases with U. The bottom panel of Figure 3 plots firm's risk premium against GPA and

MB jointly. Two patterns become legible. First, GPA is positively correlated with expected return and MB is negatively correlated with expected return. Second, the return spread over GPA is larger when MB is higher. This unveils my following analysis on how a firm's GPA and MB are differently attached to its asset composition and, more fundamentally, how underlying state variables affect firm's characteristics distinctively.

[Insert Figure 4 Here]

The top, middle, and bottom panels of Figure 4 plot ω_{AP}, ω_{IO}, and ω_{DO}, respectively, against GPA and MB. The middle and bottom panels exhibit that both more profitable (high GPA) firms and value (low MB) firms have more investment options and fewer disinvestment options, in accord with the conditions derived in Section 4 under which a positive profitability premium and a positive value premium coexist. It can also be observed that a low level of GPA marks deficiency in investment options (middle panel) and overload of disinvestment options (bottom panel) when MB is high. Therefore, the highest profitability premium among high MB firms as well as the lowest value premium among low GPA firms in essence reflects the return difference between investment options and disinvestment options.

[Insert Figure 5 Here]

The relation between firm characteristics, asset composition, and risk premium is the incorporation of separate basic state variables. The top and bottom panels of Figure 5 plot how a firm's GPA and MB, respectively, vary with $\frac{F}{C}$, the implicit productivity, and two primitive firm-specific variables, productivity Z and investment cost U when capital stock K is fixed. The top left panel suggests that GPA increases in Z and is solely driven by Z. Since Z is positively linked with $\frac{F}{C}$, the correlation between GPA and $\frac{F}{C}$ is also positive, displaying a positive profitability premium in the model. The bottom left panel shows that MB increases in both Z and U. The common reliance of MB and GPA on Z favors a positive relation between MB and GPA. Moreover, the bottom right panel of Figure 5 demonstrates that MB increases with $\frac{F}{C}$ for each level of U, as is the case in Hackbarth and Johnson (2015); but the overall relation between MB and $\frac{F}{C}$ is negative, when U, negatively related to $\frac{F}{C}$, is included in the model. This gives rise to the value premium, where firms with low MB earn a higher expected return than do those with high MB.

6.2 Failure of the CAPM

Tables 1 and 2 show the failure of CAPM in explaining the profitability strategy and the value strategy in data. This is a significant feature that can be generated by my model as demonstrated in Section 5.2. The CAPM assumes one single risk factor, and the test involves regressing a time-series of security excess returns on the market excess return.

Let $f_{X,t}$ and $f_{Y,t}$ be the factor realizations of X and Y at instant t. With the risk exposures and constituent weights derived in Section 4.1, I can write the realized excess return of stock i as

$$R_{i,t} = \phi f_{X,t} + [(\eta_1 - 1)\omega_{IO,t} + (\eta_2 - 1)\omega_{DO,t}](\phi f_{X,t} - f_{Y,t}) \tag{44}$$

This is a two-factor model, in which all firms have a common exposure to $f_{X,t}$ and a time-varying exposure to a combined factor, $(\phi f_{X,t} - f_{Y,t})$. This has two immediate implications. First, the CAPM fails in my model, for a single market factor does not trace all movements in stock returns. Second, any long-short strategy reflects the contrasts in the time-varying exposures to the combined factor, or, more fundamentally, $\omega_{IO,t}$ and $\omega_{DO,t}$.

6.3 Cyclicality of profitability premium and value premium

One salient feature about the profitability premium and the value premium is that they comove differently with the business cycle. The profitability premium is procyclical (see Kogan et al. (2021)) while the value premium is countercyclical (see Petkova and Zhang (2005) and Zhang (2005)). I first replicate the cyclicality of the two premiums and report the results in Table 9.

[Insert Table 9 Here]

I classify different states of economy based on the real GDP growth rate in the past year. Table 9 shows that the average profitability premium is larger when the state of economy is good (7.76% in the good state compared to 1.47% in the bad state) and that the average value premium is larger when the state of economy is bad (4.72% in the bad state compared to 1.44% in the good state). Moreover, the 95th percentile-5th percentile spreads of GPA and logBM are larger in the good state than in the bad state. My model provides an explanation for the cyclicality of two premiums with their corresponding sorting variables.

In the model, the state of economy is measured by $\frac{X^\phi}{Y}$. A large value of $\frac{X^\phi}{Y}$ indicates a good state, which features a high demand and a low investment price.

[Insert Figure 6 Here]

Figure 6 plots the spread of ω_{IO}, ω_{DO}, and ER against GPA and MB. In the upper panels of Figure 6, the solid lines represent the average ω_{IO}-GPA, ω_{DO}-GPA, and ER-GPA relationships in the good state (high $\frac{X^\phi}{Y}$) and the dashed lines represent those relationships in the bad state (low $\frac{X^\phi}{Y}$). The lower panels of Figure 6 repeat the same test on the average ω_{IO}-MB, ω_{DO}-MB, and ER-MB relationships in two states of economy. In the good state, not only is GPA higher on average, but the spread of GPA is also wider because the elasticity of GPA with respect to X is $\phi > 1$, contributing to a larger return spread between high GPA firms and low GPA firms in the good state relative to that in the bad state. This creates a procyclical profitability premium with the contrast between investment options and that between disinvestment options also procyclical. On the other hand, the horizontal difference between average MB does not change drastically across two states, but the vertical difference is larger in the bad state. Since $\frac{F}{C}$ tends to be low when $\frac{X^\phi}{Y}$ is low, the ER-$\frac{F}{C}$ curve becomes steeper given the concavity implied in Figure 2. Thus, the value premium is countercyclical.

GPA and MB reveal different information in different states of economy. The left four panels in Figure 6 show that during periods when more profitable firms have significantly more/fewer investment (disinvestment) options than do less profitable firms, the surplus/deficit in investment options owned by value firms versus growth firms is negligible, and that when the difference of investment (disinvestment) options between more profitable firms and less profitable firms is tiny, that between value firms and growth firms is large.

7 Conclusion

I study the effects of costly reversibility on firms' asset composition and risk premiums and find that both investment options and disinvestment options are indispensable for the concurrence of the profitability premium and the value premium. As firms derive profits from existing assets, the option to invest is a contingent claim on future profits so that it earns a higher expected return, and the option to disinvest serves as a hedge for existing assets amid deteriorating business conditions and deserves a lower expected return.

Within the model, profitable firms have more investment options and yield higher expected returns, and growth firms have more disinvestment options and yield lower expected returns. While the market-to-book ratio is always increasing in profitability, costly reversibility increases the valuation ratio of less profitable firms via raising the value of disinvestment options. Therefore, my model still retains a positive relation between GPA and MB. In the model simulation, I find that out of growth firms, more profitable ones have more investment options while less profitable firms have more disinvestment options. The difference between high profitability and low profitability reflects the contrast between investment options and disinvestment options. This explains why the profitability premium is stronger among growth firms. My model also generates a procyclical profitability premium and a countercyclical value premium, consistent with the pattern observed in data.

A Firm's profit maximization problem

Consider the following static profit maximization problem for a representative firm

$$\Pi = \max_{L} \left\{ X \left[K^{\alpha} (SL)^{1-\alpha} \right]^{\varphi} - WL \right\} \tag{A.45}$$

where L is a perfectly flexible production factor that is augmented by technology S and can be purchased at price W, X is the aggregate demand, K represents the capital stock, $0 < \alpha < 1$ indexes the output elasticity of capital, and $0 < \varphi < 1$ indexes the coefficient of returns to scale. X, S, and W are exogenous shocks, L is firm's intratemporal control variable, and K is constant in the static problem.

Setting the first order derivative of Π with respect to L to zero solves the optimal quantity of flexible inputs

$$L^* = \left[\frac{(1-\alpha)\varphi X K^{\alpha\varphi} S^{(1-\alpha)\varphi}}{W} \right]^{\frac{1}{1-(1-\alpha)\varphi}} \tag{A.46}$$

Plugging L^* into the profit function yields

$$\Pi^* = \left\{ [(1-\alpha)\varphi]^{\frac{(1-\alpha)\varphi}{1-(1-\alpha)\varphi}} - [(1-\alpha)\varphi]^{\frac{1}{1-(1-\alpha)\varphi}} \right\} X^{\frac{1}{1-(1-\alpha)\varphi}} \left(\frac{S}{W} \right)^{\frac{(1-\alpha)\varphi}{1-(1-\alpha)\varphi}} K^{\frac{\alpha\varphi}{1-(1-\alpha)\varphi}} \tag{A.47}$$

Let $\psi = \frac{-(1-\varphi)}{1-(1-\alpha)\varphi}$ and $\phi = \frac{1}{1-(1-\alpha)\varphi}$. Simple algebra shows that $-1 < \psi < 0$ and $\phi > 1$ when $0 < \alpha < 1$ and $0 < \varphi < 1$. Without loss of generality, I replace $\frac{S}{W}$ with A, which can be interpreted as per dollar value of flexible input productivity. Then Π^* can be expressed as

$$\Pi^* = (1+\psi) \left\{ [(1-\alpha)\varphi]^{\frac{(1-\alpha)\varphi}{1-(1-\alpha)\varphi}} - [(1-\alpha)\varphi]^{\frac{1}{1-(1-\alpha)\varphi}} \right\} \frac{1}{1+\psi} X^{\phi} A^{1-\phi} K^{1+\psi} \tag{A.48}$$

I omit the constant $(1+\psi) \left\{ [(1-\alpha)\varphi]^{\frac{(1-\alpha)\varphi}{1-(1-\alpha)\varphi}} - [(1-\alpha)\varphi]^{\frac{1}{1-(1-\alpha)\varphi}} \right\}$ in the main part for simplicity, which does not impair model implications.

This static model reveals that (1) the technology shock and the price shock to flexible inputs can be condensed into one shock; (2) the exposure of Π^* to X is larger than 1 as a result of profit maximization.

B Firm's value maximization problem

The value of the representative firm at instant t is

$$V(A_t, C_t, K_t) = \max_{I_t} \left\{ \mathop{\mathrm{E}}_t \int_t^\infty e^{-r(\tau - t)} \left(\frac{1}{1+\psi} A_\tau K_\tau^{1+\psi} \,\mathrm{d}\tau - C_\tau \,\mathrm{d}I_\tau \right) \right\} \tag{B.49}$$

$$\text{s.t.} \quad \mathrm{d}K_t = -\delta K_t \,\mathrm{d}t + \mathrm{d}I_t \tag{B.50}$$

Define $F(A_t, K_t)$ as the marginal product of capital and $Q(F_t, C_t)$ as the marginal value of capital. That is,

$$F(A_t, K_t) \equiv \frac{\partial \Pi_t}{\partial K_t} = A_t K_t^\psi \tag{B.51}$$

$$Q(F_t, C_t) \equiv \frac{\partial V_t}{\partial K_t} = \mathop{\mathrm{E}}_t \int_t^\infty e^{-(r+\delta)(\tau - t)} F_\tau \,\mathrm{d}\tau \tag{B.52}$$

The necessary conditions to solve for the optimal investment policy I_t are

$$Q_t > C_t \quad \text{if } \mathrm{d}I_t > 0 \tag{B.53}$$

$$\theta C_t \le Q_t \le C_t \quad \text{if } \mathrm{d}I_t = 0 \tag{B.54}$$

$$Q_t < \theta C_t \quad \text{if } \mathrm{d}I_t < 0 \tag{B.55}$$

Following McDonald and Siegel (1986), I conjecture that Q is homogeneous of degree one in F and C so that the inaction region defined in equation (B.54) associated with Q and C is now prescribed by F and C:

$$b\theta C \le F \le hC \tag{B.56}$$

where b and h (with $0 < b < h$)[8] are two unknown constants I seek.

Let $\mathring{F}(A_t, K_t)$ denote function $F(A_t, K_t)$ in the inaction region ($\mathrm{d}I_t = 0$ or $\mathrm{d}K_t = -\delta K_t\,\mathrm{d}t$). Applying Ito's lemma to $\mathring{F}(A_t, K_t)$, I obtain [9]

$$\begin{aligned}\mathrm{d}\mathring{F} &= \mathring{F}_A\,\mathrm{d}A + \mathring{F}_K\,\mathrm{d}K + \frac{1}{2}\left[\mathring{F}_{AA}(\mathrm{d}A)^2 + \mathring{F}_{KK}(\mathrm{d}K)^2 + 2\mathring{F}_{AK}(\mathrm{d}A\,\mathrm{d}K)\right]\\ &= \underbrace{(\mu_A - \psi\delta)}_{\equiv\mu_F}\underbrace{AK^{\psi}}_{=\mathring{F}}\mathrm{d}t + \sigma_A\underbrace{AK^{\psi}}_{=\mathring{F}}\mathrm{d}w_A\end{aligned}$$

so $\mathring{F}$ follows a geometric Brownian motion

$$\frac{\mathrm{d}\mathring{F}}{\mathring{F}} = \mu_F\,\mathrm{d}t + \sigma_A\,\mathrm{d}w_A \tag{B.57}$$

Let H_t be a non-decreasing function of time representing the cumulative increase in F_t resulting from capital acquisition when $F_t = hC_t$, and let B_t be a non-decreasing function of time representing the cumulative increase in F_t resulting from capital sale when $F_t = b\theta C_t$. The relation between F_t and $\mathring{F}_t$ is

$$F_t = \frac{\mathring{F}_t B_t}{H_t}$$

and so

$$\mathrm{d}F = \mu_F F\,\mathrm{d}t + \sigma_A F\,\mathrm{d}w_A + \frac{b\theta C\,\mathrm{d}B}{B} - \frac{hC\,\mathrm{d}H}{H} \tag{B.58}$$

[8] $0 < b < h$ because $0 < \theta < 1$. When $\theta = 1$, $b = h$; when $\theta = 0$, $b = 0$.

[9] To be concise, I omit the subscript t in state variables and omit the independent variables in functions F, $\mathring{F}$, Q, and their derivatives whenever no ambiguity is caused.

Applying Ito's lemma to $Q(F_t, K_t)$, with $\mathrm{d}F$ plugged in as in (B.58), yields

$$\begin{aligned}
\mathrm{d}Q = {} & Q_F(F,C)(\mu_F F\,\mathrm{d}t + \sigma_A F\,\mathrm{d}w_A) + Q_F(b\theta C, C)\frac{b\theta C\,\mathrm{d}B}{B} - Q_F(hC, C)\frac{hC\,\mathrm{d}H}{H} \\
& + Q_C(F,C)(\mu_C C\,\mathrm{d}t + \sigma_C C\,\mathrm{d}w_C) \\
& + \frac{1}{2}(Q_{FF}(F,C)\sigma_A^2 F^2\,\mathrm{d}t + Q_{CC}(F,C)\sigma_C^2 C^2\,\mathrm{d}t + 2Q_{FC}(F,C)\rho_{AC}\sigma_A\sigma_C FC\,\mathrm{d}t)
\end{aligned} \tag{B.59}$$

$Q(F_t, C_t)$ must satisfy the following transversality condition

$$\lim_{T\to\infty} \mathop{\mathrm{E}}_t \left[e^{-(r+\delta)(T-t)} Q(F_T, C_T)\right] = 0 \tag{B.60}$$

where $e^{-(r+\delta)(T-t)}Q(F_T, C_T)$ can be expanded as

$$e^{-(r+\delta)(T-t)}Q(F_T, C_T) = Q(F_t, C_t) + \int_t^T Q(F_\tau, C_\tau)\,\mathrm{d}e^{-(r+\delta)(\tau-t)} + \int_t^T e^{-(r+\delta)(\tau-t)}\,\mathrm{d}Q(F_\tau, C_\tau)$$

with $\mathrm{d}Q$ described in (B.59). Then I obtain the following partial differential pricing equation

$$\mu_F F Q_F + \mu_C C Q_C + \frac{\sigma_A^2}{2}F^2 Q_{FF} + \frac{\sigma_C^2}{2}C^2 Q_{CC} + \rho_{AC}\sigma_A\sigma_C FC Q_{FC} - (r+\delta)Q + F = 0 \tag{B.61}$$

with two smooth-pasting conditions

$$Q(hC, C) = C \tag{B.62}$$

$$Q(b\theta C, C) = \theta C \tag{B.63}$$

and two high-contact conditions

$$Q_F(hC, C) = 0 \quad \forall C > 0 \tag{B.64}$$

$$Q_F(b\theta C, C) = 0 \quad \forall C > 0 \tag{B.65}$$

I first consider the following functional form for the firm value

$$V(A,C,K) = D_0 A K^{1+\psi} + D_1 A^{\eta_1} C^{1-\eta_1} K^{1+\psi\eta_1} + D_2 A^{\eta_2} C^{1-\eta_2} K^{1+\psi\eta_2} \tag{B.66}$$

with parameters $D_0, D_1, D_2 > 0$ so that

$$Q(F,C) \equiv \frac{\partial V}{\partial K} = E_0 F + E_1 F^{\eta_1} C^{1-\eta_1} + E_2 F^{\eta_2} C^{1-\eta_2} \tag{B.67}$$

is homogeneous of degree one in F and C, and that the two high-contact conditions hold for all $C > 0$. Accordingly, the smoothing-pasting conditions (B.62) and (B.63) and high-contact conditions (B.64) and (B.65) become

$$E_0 + \eta_1 E_1 h^{\eta_1 - 1} + \eta_2 E_2 h^{\eta_2 - 1} = 0 \tag{B.68}$$

$$E_0 + \eta_1 E_1 (b\theta)^{\eta_1 - 1} + \eta_2 E_2 (b\theta)^{\eta_2 - 1} = 0 \tag{B.69}$$

$$E_0 h + E_1 h^{\eta_1} + E_2 h^{\eta_2} = 1 \tag{B.70}$$

$$E_0 b\theta + E_1 (b\theta)^{\eta_1} + E_2 (b\theta)^{\eta_2} = \theta \tag{B.71}$$

Plugging equation (B.67) and its derivatives into equation (B.61) yields

$$E_0 = \frac{1}{r + \delta - \mu_F} \tag{B.72}$$

and η_1 and η_2 are positive and negative roots, respectively, of the following quadratic equation in η

$$\frac{\sigma_A^2 + \sigma_C^2 - 2\rho_{AC}\sigma_A\sigma_C}{2}\eta^2 + \left(\mu_F - \mu_C - \frac{\sigma_A^2 + \sigma_C^2 - 2\rho_{AC}\sigma_A\sigma_C}{2}\right)\eta - (r + \delta - \mu_C) = 0 \tag{B.73}$$

Inserting the values of E_0, η_1, and η_2 into conditions (B.68) - (B.71) generates a system of four nonlinear equations with four unknowns E_1, E_2, h, and b that can be solved numerically.

D_0, D_1, and D_2 can finally be worked out.

C Asset composition and the expected return

The first order derivatives of the weights of assets in place ($\omega_{AP,t}$ as in equation (23)), investment options ($\omega_{IO,t}$ as in equation (24)), and disinvestment options ($\omega_{DO,t}$ as in equation (25))with respect to $\frac{F_t}{C_t}$ are

$$\frac{\partial \omega_{AP,t}}{\partial\left(\frac{F_t}{C_t}\right)} = \frac{D_0\left[D_2(1-\eta_2)\left(\frac{F_t}{C_t}\right)^{\eta_2} - D_1(\eta_1-1)\left(\frac{F_t}{C_t}\right)^{\eta_1}\right]}{\left[D_0\left(\frac{F_t}{C_t}\right) + D_1\left(\frac{F_t}{C_t}\right)^{\eta_1} + D_2\left(\frac{F_t}{C_t}\right)^{\eta_2}\right]^2} \tag{C.74}$$

$$\frac{\partial \omega_{IO,t}}{\partial\left(\frac{F_t}{C_t}\right)} = \frac{D_1\left(\frac{F_t}{C_t}\right)^{\eta_1-1}\left[D_0(\eta_1-1)\left(\frac{F_t}{C_t}\right) + D_2(\eta_1-\eta_2)\left(\frac{F_t}{C_t}\right)^{\eta_2}\right]]}{\left[D_0\left(\frac{F_t}{C_t}\right)_t + D_1\left(\frac{F_t}{C_t}\right)^{\eta_1} + D_2\left(\frac{F_t}{C_t}\right)^{\eta_2}\right]^2} \tag{C.75}$$

$$\frac{\partial \omega_{DO,t}}{\partial\left(\frac{F_t}{C_t}\right)} = -\frac{D_2\left(\frac{F_t}{C_t}\right)^{\eta_2-1}\left[D_0(1-\eta_2)\left(\frac{F_t}{C_t}\right) + D_1(\eta_1-\eta_2)\left(\frac{F_t}{C_t}\right)^{\eta_1}\right]]}{\left[D_0\left(\frac{F_t}{C_t}\right)_t + D_1\left(\frac{F_t}{C_t}\right)^{\eta_1} + D_2\left(\frac{F_t}{C_t}\right)^{\eta_2}\right]^2} \tag{C.76}$$

Since $D_0, D_1, D_2 > 0$, $\eta_1 > 1$ and $\eta_2 < 0$, it is obvious that $\frac{\partial \omega_{IO,t}}{\partial(F_t/C_t)} > 0$ and $\frac{\partial \omega_{DO,t}}{\partial(F_t/C_t)} < 0$. Observing equation (C.74), I find that $D_2(1-\eta_2)\left(\frac{F_t}{C_t}\right)^{\eta_2}$ decreases with $\frac{F_t}{C_t}$ and that $D_1(\eta_1-1)\left(\frac{F_t}{C_t}\right)^{\eta_1}$ increases with $\frac{F_t}{C_t}$. Equating the two expressions solves the critical point of $\left(\frac{F_t}{C_t}\right)^* = \left[\frac{D_2(1-\eta_2)}{D_1(\eta_1-1)}\right]^{\frac{1}{\eta_1-\eta_2}}$, below which $\frac{\partial \omega_{AP,t}}{\partial(F_t/C_t)} > 0$ and above which $\frac{\partial \omega_{AP,t}}{\partial(F_t/C_t)} < 0$.

The first order derivatives of the expected return with respect to $\frac{F_t}{C_t}$ is

$$\begin{aligned}
\frac{\partial ER_t}{\partial\left(\frac{F_t}{C_t}\right)} &= ER_{AP}\frac{\partial \omega_{AP,t}}{\partial\left(\frac{F_t}{C_t}\right)} + ER_{IO}\frac{\partial \omega_{IO,t}}{\partial\left(\frac{F_t}{C_t}\right)} + ER_{DO}\frac{\partial \omega_{DO,t}}{\partial\left(\frac{F_t}{C_t}\right)} \\
&= \frac{\phi\gamma_X\sigma_X - \gamma_Y\sigma_Y}{\left[D_0\left(\frac{F_t}{C_t}\right) + D_1\left(\frac{F_t}{C_t}\right)^{\eta_1} + D_2\left(\frac{F_t}{C_t}\right)^{\eta_2}\right]^2} \\
&\quad \cdot \left[D_0D_1(\eta_1-1)^2\left(\frac{F_t}{C_t}\right)^{\eta_1} + D_0D_2(1-\eta_2)^2\left(\frac{F_t}{C_t}\right)^{\eta_2} + D_1D_2(\eta_1-\eta_2)^2\left(\frac{F_t}{C_t}\right)^{\eta_1+\eta_2-1}\right]
\end{aligned}$$

This is positive if and only if $\phi\gamma_X\sigma_X - \gamma_Y\sigma_Y > 0$.

References

Abel, Andrew B, 1983, Optimal investment under uncertainty, *The American Economic Review* 73, 228–233.

Abel, Andrew B, and Janice C Eberly, 1993, A unified model of investment under uncertainty, Technical report, National Bureau of Economic Research.

Abel, Andrew B, and Janice C Eberly, 1996, Optimal investment with costly reversibility, *The Review of Economic Studies* 63, 581–593.

Ai, Hengjie, and Dana Kiku, 2013, Growth to value: Option exercise and the cross section of equity returns, *Journal of Financial Economics* 107, 325–349.

Berk, Jonathan B., Richard C. Green, and Vasant Naik, 1999, Optimal investment, growth options, and security returns, *Journal of Finance* 54, 1553–1607.

Bertola, Giuseppe, 1988, *Adjustment costs and dynamic factor demands: investment and employment under uncertainty*, Ph.D. thesis, Massachusetts Institute of Technology.

Bertola, Giuseppe, 1998, Irreversible investment, *Research in Economics* 52, 3–37.

Bloom, Nicholas, 2009, The impact of uncertainty shocks, *Econometrica* 77, 623–685.

Carlson, Murray, Adlai Fisher, and Ron Giammarino, 2004, Corporate investment and asset price dynamics: Implications for the cross-section of returns, *Journal of Finance* 59, 2577–2603.

Constantinides, George M, 1978, Market risk adjustment in project valuation, *The Journal of Finance* 33, 603–616.

Cooper, Ilan, 2006, Asset pricing implications of nonconvex adjustment costs and irreversibility of investment, *Journal of Finance* 61, 139–170.

Dixit, Robert K, and Robert S Pindyck, 2012, *Investment under uncertainty* (Princeton university press).

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3 – 56.

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1 – 22.

Fama, Eugene F, and James D MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of political economy* 81, 607–636.

Gomes, Joao, Leonid Kogan, and Lu Zhang, 2003, Equilibrium cross section of returns, *Journal of Political Economy* 111, 693–732.

Gu, Lifeng, Dirk Hackbarth, and Tim Johnson, 2018, Inflexibility and stock returns, *The Review of Financial Studies* 31, 278–321.

Hackbarth, Dirk, and Timothy Johnson, 2015, Real options and risk dynamics, *The Review of Economic Studies* 82, 1449–1482.

Kogan, Leonid, Jun Li, and Harold H Zhang, 2021, Operating hedge and gross profitability premium, *Available at SSRN 3947165* .

Kogan, Leonid, and Dimitris Papanikolaou, 2013, Firm characteristics and stock returns: The role of investment-specific shocks, *Review of Financial Studies* .

Li, Jun, 2018, Explaining momentum and value simultaneously, *Management Science* 64, 4239–4260.

Ma, Liang, and Hong Yan, 2015, The value and profitability premiums, working paper.

McDonald, Robert, and Daniel Siegel, 1986, The value of waiting to invest, *The quarterly journal of economics* 101, 707–727.

Novy-Marx, Robert, 2013, The other side of value: The gross profitability premium, *Journal of Financial Economics* 108, 1–28.

Petkova, Ralitsa, and Lu Zhang, 2005, Is value riskier than growth?, *Journal of Financial Economics* 78, 187–202.

Zhang, Lu, 2005, The value premium, *Journal of Finance* 60, 67–103.

Table 1. Gross profitability decile portfolios: characteristics and returns

At the end of June of each year t, firms are sorted into decile portfolios based on their fiscal year $t-1$ end gross profitability (GPA). This table reports the characteristics, value-weighted returns, and results of asset pricing tests of 10 portfolios formed on GPA. Panel A reports characteristics of GPA portfolios, including gross profitability (GPA), log book-to-market ratio (logBM), log June-end market capitalization (logME), cumulative return over past twelve to two months (Mom), Tobin's Q (Q), investment-to-capital ratio (IK), and cash holdings (CHK). Panel B reports the average excess returns, results of CAPM test, and results of Fama-French three-factor model test. The sample period is from July 1964 to June 2020. Newey-West t-statistics adjusted for heteroscedasticity and autocorrelation are provided in parentheses.

Panel A: Characteristics of GPA deciles

Decile	GPA	logBM	logME	Mom	Q	IK	CHK
Low	0.05	-0.54	4.11	-4.63	1.37	0.11	0.29
2	0.15	-0.14	4.85	3.00	0.82	0.10	0.11
3	0.21	-0.19	5.06	4.28	0.78	0.10	0.08
4	0.27	-0.28	4.98	5.63	0.95	0.10	0.09
5	0.32	-0.35	4.88	6.63	1.36	0.10	0.14
6	0.38	-0.43	4.79	6.46	2.14	0.11	0.21
7	0.44	-0.51	4.80	7.12	2.70	0.12	0.25
8	0.52	-0.59	4.80	7.75	3.29	0.12	0.23
9	0.64	-0.69	4.77	8.83	3.64	0.14	0.21
High	0.88	-0.77	4.55	8.43	3.32	0.13	0.20

Panel B: Returns of GPA deciles

	Ret-Rf	α(CAPM)	MKT	α(FF3F)	MKT	SMB	HML
Low	4.05	-3.99	1.28	-4.25	1.18	0.43	-0.05
	(1.23)	(-2.26)	(28.65)	(-2.50)	(25.28)	(5.44)	(-0.78)
2	4.98	-1.40	1.01	-2.09	1.06	-0.09	0.21
	(2.14)	(-1.36)	(37.03)	(-2.22)	(41.84)	(-2.41)	(5.30)
3	4.35	-2.01	1.01	-2.74	1.05	-0.03	0.20
	(1.89)	(-1.86)	(36.48)	(-2.70)	(43.36)	(-0.70)	(4.79)
4	7.50	1.24	0.99	0.78	1.03	-0.06	0.14
	(3.41)	(1.36)	(43.99)	(0.88)	(46.47)	(-1.41)	(3.26)
5	6.70	0.30	1.02	0.17	0.99	0.11	0.00
	(3.00)	(0.39)	(52.60)	(0.21)	(49.21)	(3.27)	(0.10)
6	6.64	-0.11	1.07	0.59	1.03	0.06	-0.20
	(2.76)	(-0.11)	(34.26)	(0.66)	(37.78)	(1.51)	(-4.10)
7	5.62	-0.87	1.03	-0.04	1.00	0.00	-0.22
	(2.40)	(-0.87)	(47.58)	(-0.04)	(46.02)	(0.07)	(-5.36)
8	6.59	0.24	1.01	1.26	0.97	-0.01	-0.26
	(2.71)	(0.23)	(43.03)	(1.36)	(43.99)	(-0.42)	(-7.61)
9	8.93	3.13	0.92	4.14	0.91	-0.12	-0.23
	(4.06)	(2.76)	(31.09)	(3.97)	(30.97)	(-2.91)	(-3.92)
High	9.37	3.32	0.96	4.24	0.92	0.03	-0.25
	(3.97)	(2.91)	(37.24)	(4.09)	(36.36)	(1.00)	(-5.05)
High - Low	5.32	7.31	-0.32	8.49	-0.26	-0.39	-0.20
	(2.40)	(3.29)	(-5.18)	(4.04)	(-4.51)	(-5.04)	(-2.04)

Table 2. Book-to-market decile portfolios: characteristics and returns

At the end of June of each year t, firms are sorted into decile portfolios based on their fiscal year $t-1$ end book-to-market ratio (BM). This table reports the characteristics, value-weighted returns, and results of asset pricing tests of 10 portfolios formed on BM. Panel A reports characteristics of BM portfolios, including gross profitability (GPA), log book-to-market ratio (logBM), log June-end market capitalization (logME), cumulative return over past twelve to two months (Mom), Tobin's Q (Q), investment-to-capital ratio (IK), and cash holdings (CHK). Panel B reports the average excess returns, results of CAPM test, and results of Fama-French three-factor model test. The sample period is from July 1964 to June 2020. Newey-West t-statistics adjusted for heteroscedasticity and autocorrelation are provided in parentheses.

Panel A: Characteristics of BM deciles

Decile	GPA	logBM	logME	Mom	Q	IK	CHK
Low	0.49	-1.82	5.54	4.56	4.14	0.14	0.22
2	0.40	-1.18	5.69	6.59	2.71	0.12	0.19
3	0.37	-0.89	5.48	6.04	1.88	0.12	0.15
4	0.34	-0.67	5.31	6.29	1.44	0.11	0.13
5	0.31	-0.49	5.09	6.09	1.15	0.11	0.13
6	0.29	-0.31	4.84	6.29	0.91	0.10	0.11
7	0.26	-0.14	4.62	5.62	0.83	0.10	0.13
8	0.24	0.04	4.28	5.36	0.70	0.10	0.12
9	0.23	0.27	3.84	5.22	0.49	0.11	0.10
High	0.22	0.71	3.16	4.90	0.26	0.10	0.11

Panel B: Returns of BM deciles

	Ret-Rf	α(CAPM)	MKT	α(FF3F)	MKT	SMB	HML
Low	6.32	-0.32	1.06	1.35	1.02	-0.10	-0.41
	(2.51)	(-0.30)	(46.63)	(1.70)	(53.60)	(-3.04)	(-13.71)
2	5.92	-0.37	1.00	0.32	0.99	-0.05	-0.17
	(2.65)	(-0.46)	(49.06)	(0.44)	(57.86)	(-1.90)	(-3.82)
3	6.64	0.38	0.99	0.75	0.98	0.00	-0.10
	(3.04)	(0.51)	(53.28)	(1.05)	(50.46)	(-0.07)	(-3.08)
4	6.19	0.20	0.95	0.05	0.93	0.10	0.01
	(2.90)	(0.25)	(51.16)	(0.06)	(51.59)	(3.47)	(0.36)
5	8.00	1.75	0.99	1.49	0.98	0.08	0.05
	(3.64)	(2.10)	(49.96)	(1.79)	(51.71)	(2.06)	(1.33)
6	7.33	0.92	1.02	0.40	1.00	0.15	0.09
	(3.14)	(0.91)	(44.08)	(0.42)	(46.28)	(4.88)	(2.63)
7	9.10	2.52	1.05	1.67	1.05	0.11	0.19
	(3.88)	(2.84)	(41.25)	(2.02)	(44.65)	(2.95)	(4.78)
8	5.79	-0.74	1.04	-1.81	1.04	0.15	0.24
	(2.38)	(-0.60)	(31.66)	(-1.61)	(33.47)	(2.73)	(5.33)
9	9.06	2.61	1.02	0.79	1.05	0.18	0.43
	(3.80)	(1.80)	(32.11)	(0.61)	(41.39)	(4.53)	(7.80)
High	11.11	4.04	1.12	1.41	1.11	0.46	0.56
	(3.84)	(2.34)	(23.19)	(1.05)	(35.00)	(7.78)	(8.15)
High - Low	4.78	4.35	0.07	0.06	0.09	0.55	0.97
	(2.04)	(1.79)	(1.06)	(0.03)	(2.37)	(7.77)	(12.59)

Table 3. Portfolios double sorted on book-to-market and gross profitability

At the end of June of each year t, firms are sorted into quintile portfolios based on their fiscal year $t-1$ end book-to-market ratio (BM); within each quintile, firms are further sorted into five portfolios based on their fiscal year $t-1$ end gross profitability (GPA). This table reports the value-weighted returns and characteristics of the 25 BM and GPA double-sorted portfolios. Portfolio characteristics include gross profitability (GPA), log book-to-market ratio (logBM), log June-end market capitalization (logME), cumulative return over past twelve to two months (Mom), Tobin's Q (Q), investment-to-capital ratio (IK), and cash holdings (CHK). The sample period is from July 1964 to June 2020.

BMGroup	GPAGroup	Ret-Rf	GPA	logBM	logME	Mom	Q	IK	CHK
	1	-2.66	0.08	-1.69	4.50	-4.92	3.74	0.17	0.37
	2	4.64	0.27	-1.50	5.54	3.33	3.04	0.14	0.22
1	3	3.44	0.43	-1.48	5.84	5.43	4.89	0.13	0.31
	4	6.84	0.59	-1.54	5.96	9.57	5.42	0.14	0.29
	5	8.90	0.82	-1.65	5.76	10.53	5.90	0.15	0.27
	1	2.42	0.15	-0.85	5.02	-0.28	1.75	0.13	0.17
	2	4.74	0.28	-0.83	5.86	5.76	1.80	0.11	0.15
2	3	6.35	0.39	-0.84	5.65	7.04	2.33	0.11	0.23
	4	9.57	0.51	-0.86	5.35	8.26	2.63	0.12	0.22
	5	9.59	0.75	-0.87	5.06	8.92	2.02	0.13	0.18
	1	4.91	0.15	-0.42	5.03	0.94	0.93	0.11	0.12
	2	7.88	0.26	-0.42	5.55	6.21	0.94	0.10	0.10
3	3	7.80	0.35	-0.43	5.20	7.80	1.18	0.11	0.15
	4	10.93	0.46	-0.43	4.85	8.41	1.40	0.12	0.20
	5	9.62	0.70	-0.44	4.51	8.73	1.13	0.12	0.15
	1	7.52	0.13	-0.04	4.75	2.94	0.65	0.10	0.11
	2	7.88	0.23	-0.05	5.05	5.95	0.57	0.10	0.08
4	3	11.89	0.32	-0.06	4.54	6.37	0.65	0.10	0.10
	4	12.37	0.43	-0.06	4.23	7.53	0.97	0.11	0.19
	5	10.25	0.70	-0.07	3.84	7.44	0.61	0.11	0.15
	1	8.10	0.09	0.57	3.72	0.92	0.36	0.08	0.15
	2	8.35	0.19	0.49	4.07	4.94	0.30	0.10	0.09
5	3	12.12	0.28	0.44	3.73	6.15	0.32	0.11	0.10
	4	11.48	0.39	0.44	3.35	6.38	0.30	0.10	0.13
	5	11.62	0.68	0.46	3.06	7.24	0.13	0.10	0.14

Table 4. 25 portfolios sorted on gross profitability

At the end of June of each year t, firms are allocated into twenty-five portfolios based on their fiscal year $t-1$ end gross profitability (GPA). This table reports the value-weighted returns and characteristics of the 25 GPA sorted portfolios. Portfolio characteristics include gross profitability (GPA), log book-to-market ratio (logBM), log June-end market capitalization (logME), cumulative return over past twelve to two months (Mom), Tobin's Q (Q), investment-to-capital ratio (IK), and cash holdings (CHK). The sample period is from July 1964 to June 2020.

GPAGroup	Ret-Rf	GPA	logBM	logME	Mom	Q	IK	CHK
1	4.54	-0.16	-0.80	3.89	-6.44	5.84	0.12	1.16
2	2.48	0.05	-0.47	4.21	-4.26	1.54	0.12	0.32
3	6.58	0.10	-0.22	4.52	0.55	1.40	0.11	0.26
4	4.14	0.14	-0.12	4.81	3.24	0.86	0.10	0.13
5	4.69	0.17	-0.14	5.00	3.45	0.72	0.10	0.08
6	4.89	0.20	-0.17	5.10	3.61	0.73	0.10	0.09
7	5.16	0.22	-0.20	5.04	4.66	0.83	0.11	0.09
8	5.73	0.24	-0.23	4.97	5.04	0.93	0.10	0.09
9	6.96	0.27	-0.28	5.05	5.84	0.94	0.11	0.09
10	8.69	0.29	-0.28	4.93	6.00	1.07	0.10	0.11
11	6.08	0.31	-0.33	4.93	5.94	1.26	0.11	0.14
12	7.91	0.33	-0.35	4.86	7.23	1.54	0.10	0.17
13	5.01	0.35	-0.39	4.82	6.38	1.95	0.11	0.21
14	7.47	0.37	-0.41	4.77	6.42	1.94	0.11	0.19
15	7.26	0.40	-0.46	4.80	6.69	2.48	0.12	0.24
16	5.97	0.42	-0.47	4.81	7.04	2.49	0.12	0.23
17	6.70	0.45	-0.52	4.83	7.55	2.81	0.12	0.28
18	8.36	0.48	-0.54	4.82	7.70	3.02	0.12	0.24
19	7.45	0.51	-0.58	4.83	7.42	3.28	0.12	0.23
20	5.65	0.55	-0.63	4.75	7.97	3.36	0.12	0.23
21	9.58	0.60	-0.64	4.74	8.80	3.26	0.12	0.22
22	8.64	0.65	-0.70	4.76	9.28	3.84	0.14	0.22
23	8.36	0.73	-0.77	4.81	8.73	4.04	0.14	0.23
24	10.07	0.84	-0.79	4.71	8.58	3.55	0.13	0.21
25	9.54	1.09	-0.76	4.27	7.97	2.61	0.13	0.20

Table 5. Fama-MacBeth regressions of returns

This table reports results of Fama and MacBeth (1973) regressions of monthly stock excess returns R on a set of variables that are related with stock returns, including gross profitability GPA, log book-to-market ratio $\log(B/M)$, log June-end market capitalization $\log(ME)$, cumulative return over past twelve to two months Mom. X stands for $\log(B/M)$ in specification (1) and GPA in specification (2). The sample period for stock returns is from July 1964 to June 2020. The slope coefficients are reported after being multiplied by 100. Newey-West t-statistics adjusted for heteroscedasticity and autocorrelation are provided in parentheses.

	(1)	(2)
Intercept	-0.58	-0.66
	(-1.55)	(-1.77)
$X \times GPA$	-0.21	-0.65
	(-2.22)	(-2.56)
GPA	0.90	1.52
	(5.97)	(5.11)
$\log(B/M)$	0.41	0.32
	(5.25)	(4.85)
$\log(ME)$	0.12	0.11
	(3.14)	(3.11)
Mom	0.85	0.84
	(5.15)	(5.11)

Table 6. Parameters values

This table summarizes parameters used in the model. Panel A reports parameters guided by the existing literature; Panel B reports parameters estimated from the simulated method of moments and their standard errors in parentheses. All values are in annual frequency.

Panel A: Calibrated parameters

Symbol	Description	Value
μ_X	Drift of aggregate demand	0.012
σ_X	Volatility of aggregate demand	0.15
μ_Y	Drift of market price of investment	0.096
σ_Y	Volatility of market price of investment	0.06
r	Risk-free rate	0.07
μ_Z	Drift of idiosyncratic productivity	0
σ_Z	Volatility of idiosyncratic productivity	0.001
μ_U	Drift of idiosyncratic investment cost	-0.03
σ_U	Volatility of idiosyncratic investment cost	0.001
ρ_{ZU}	Correlation between idiosyncratic productivity and idiosyncratic investment cost	0
ψ	Capital curvature index	-0.07
ϕ	Profit elasticity to the aggregate demand	1.67
δ	Depreciation rate	0.10

Panel B: Estimated parameters

Symbol	Description	Estimated value	Standard error
θ	Discount of capital resale price	0.414	0.006
γ_X	Price of aggregate demand risk	0.632	0.009
γ_Y	Price of investment price risk	-0.598	0.010

Table 7. Moments in SMM estimation

This table reports the 7 moments used in the SMM estimation. I simulate the model 100 times with 800 firms over 100 years at a weekly frequency. I report the moments in data, and mean, 5th, 25th, median, 75th, 95th percentiles of annualized moments.

	Data	Mean	5th	25th	Median	75th	95th
Market excess return	6.25%	7.50%	2.49%	4.91%	6.59%	9.38%	15.25%
Average profitability premium	5.32%	5.15%	0.35%	1.81%	4.20%	7.31%	14.01%
Average value premium	4.78%	4.64%	0.35%	1.14%	3.22%	6.67%	14.35%
GPA spread over GPA sorted decile portfolios	0.83	0.73	0.22	0.40	0.62	0.93	1.59
logBM spread over GPA sorted decile portfolios	-0.23	-2.42	-3.10	-2.78	-2.52	-2.05	-1.52
GPA spread over BM sorted decile portfolios	-0.27	-0.70	-1.68	-0.81	-0.59	-0.38	-0.16
logBM spread over BM sorted decile portfolios	2.54	2.58	1.72	2.21	2.67	2.91	3.22

Table 8. Portfolio returns in simulation

This table reports the average annualized excess returns and results of CAPM test for 10 gross profitability (GPA) deciles in Panel A and 10 book-to-market (BM) deciles in Panel B from simulated data.

Panel A: Returns of GPA deciles

	Ret-Rf	α(CAPM)	MKT
Low	4.47	-3.01	1.00
	(10.65)	(-5.80)	(12.99)
2	7.58	0.23	0.98
	(17.25)	(0.43)	(13.77)
3	8.25	0.38	1.05
	(18.56)	(0.84)	(21.02)
4	8.67	0.73	1.06
	(17.70)	(1.22)	(12.58)
5	8.28	0.49	1.04
	(17.30)	(0.87)	(13.26)
6	7.60	1.33	0.84
	(17.78)	(2.23)	(9.38)
7	7.07	0.55	0.87
	(16.77)	(1.00)	(11.61)
8	6.48	-0.43	0.92
	(15.79)	(-0.88)	(13.41)
9	7.02	-0.88	1.05
	(15.80)	(-1.98)	(19.32)
High	9.62	0.60	1.20
	(18.99)	(1.20)	(20.55)
High - Low	5.15	3.61	0.21
	(14.46)	(5.51)	(2.17)

Panel B: Returns of BM deciles

	Ret-Rf	α(CAPM)	MKT
Low	4.66	-1.84	0.87
	(10.50)	(-2.94)	(8.79)
2	5.75	0.89	0.65
	(15.16)	(1.54)	(8.45)
3	6.51	0.89	0.75
	(15.71)	(1.50)	(10.33)
4	7.14	0.31	0.91
	(15.70)	(0.48)	(11.18)
5	7.20	-0.62	1.04
	(15.44)	(-1.20)	(15.47)
6	7.75	-1.15	1.19
	(15.07)	(-1.98)	(16.71)
7	8.15	-0.67	1.18
	(15.51)	(-1.16)	(15.18)
8	9.38	0.46	1.19
	(17.41)	(0.75)	(15.68)
9	9.20	0.50	1.16
	(18.20)	(0.81)	(16.20)
High	9.30	1.23	1.08
	(18.34)	(2.00)	(13.75)
High - Low	4.64	3.07	0.21
	(11.94)	(4.42)	(2.22)

Table 9. Cyclicality of profitability premium and value premium

This table reports the profitability premium and the value premium and the 95th percentile-5th percentile spread of GPA and logBM in two states of economy classified by the past year real GDP growth rate. The sample period for stock returns is from July 1964 to June 2020.

State of economy	Profitability premium	Value premium	GPA spread	logBM spread
Good	7.76	1.44	0.95	3.04
Bad	1.47	4.72	0.79	2.44

Figure 1. Portfolio excess returns and characteristics

This figure plots average monthly excess returns and gross profitability (GPA) of 25 portfolios created from five-by-five double sorts on book-to-market ratio (BM) and gross profitability (GPA) (top two panels) and on GPA and GPA (bottom two panels).

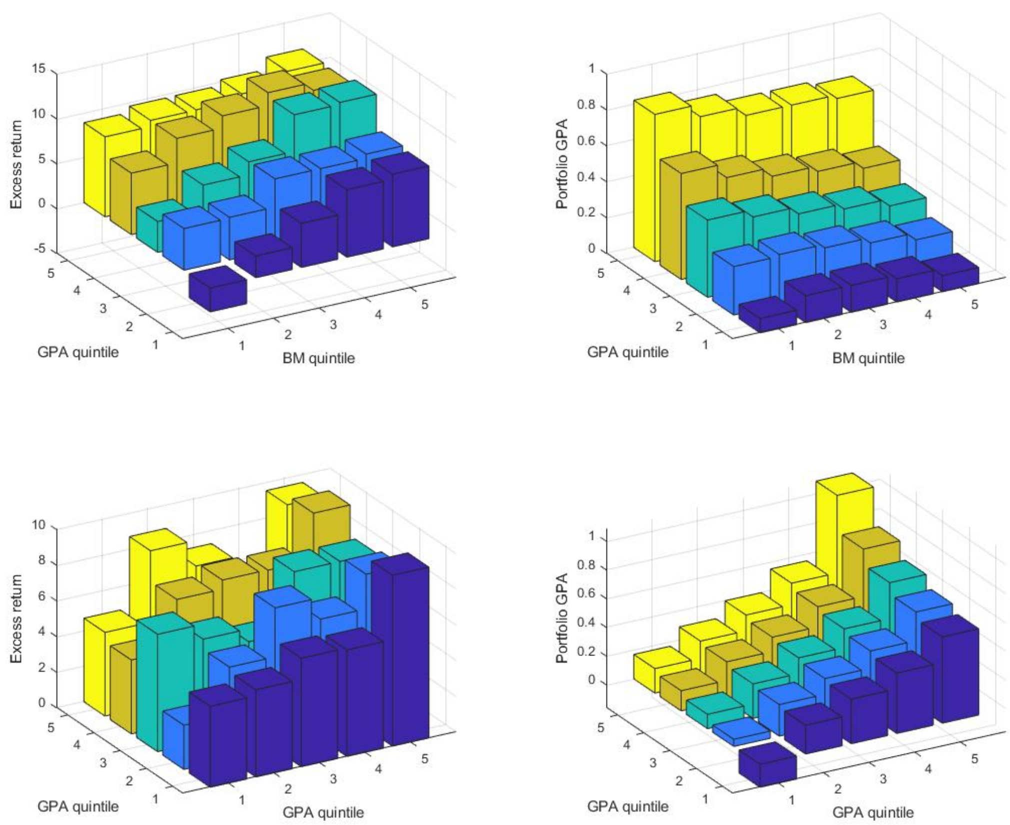


Figure 2. Asset composition and risk premium

This figure plots weights of assets in place, investment options, and disinvestment options in percentage against the ratio of marginal product of capital to investment costs ($\frac{F}{C}$) in Panel A, and plots expected returns against $\frac{F}{C}$ in Panel B.

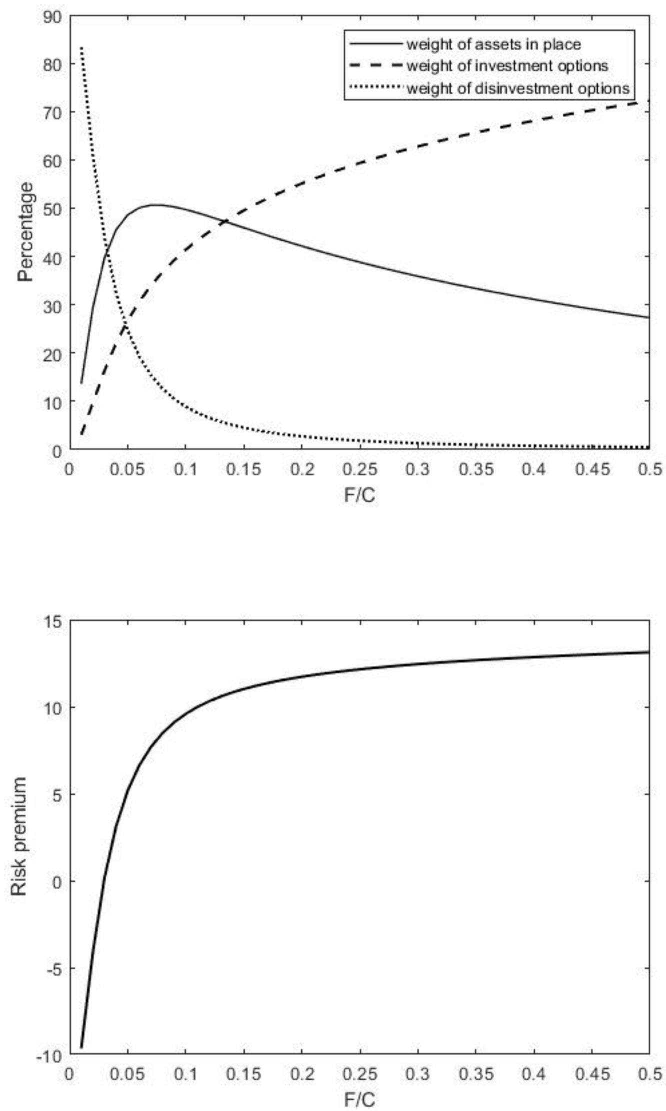
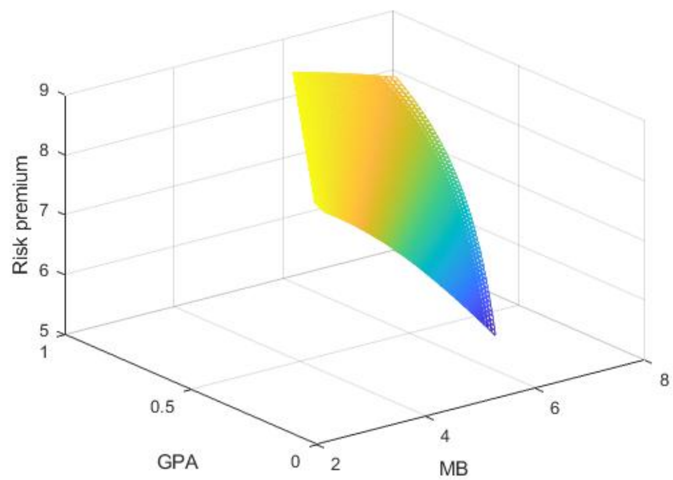


Figure 3. Expected excess return

This figure plots expected returns against profitability (GPA) and market-to-book ratio (MB) in Panel A, and against idiosyncratic productivity (Z) and idiosyncratic dissipation (U) in Panel B.



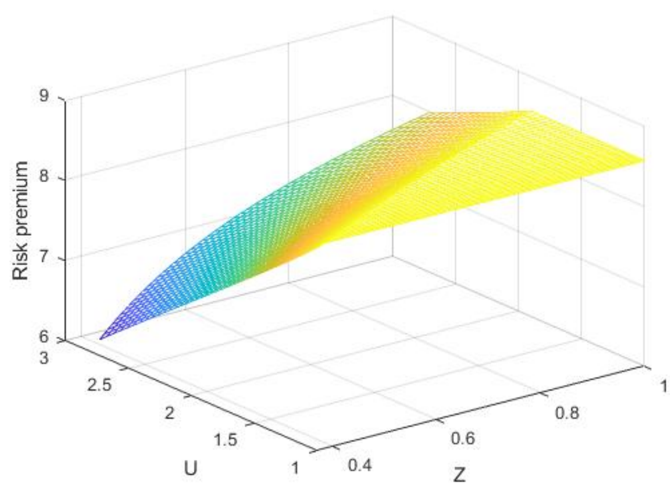


Figure 4. Asset composition and firm characteristics

This figure plots weights of assets in place, investment options, and disinvestment options in percentage against profitability (GPA) and market-to-book ratio (MB).

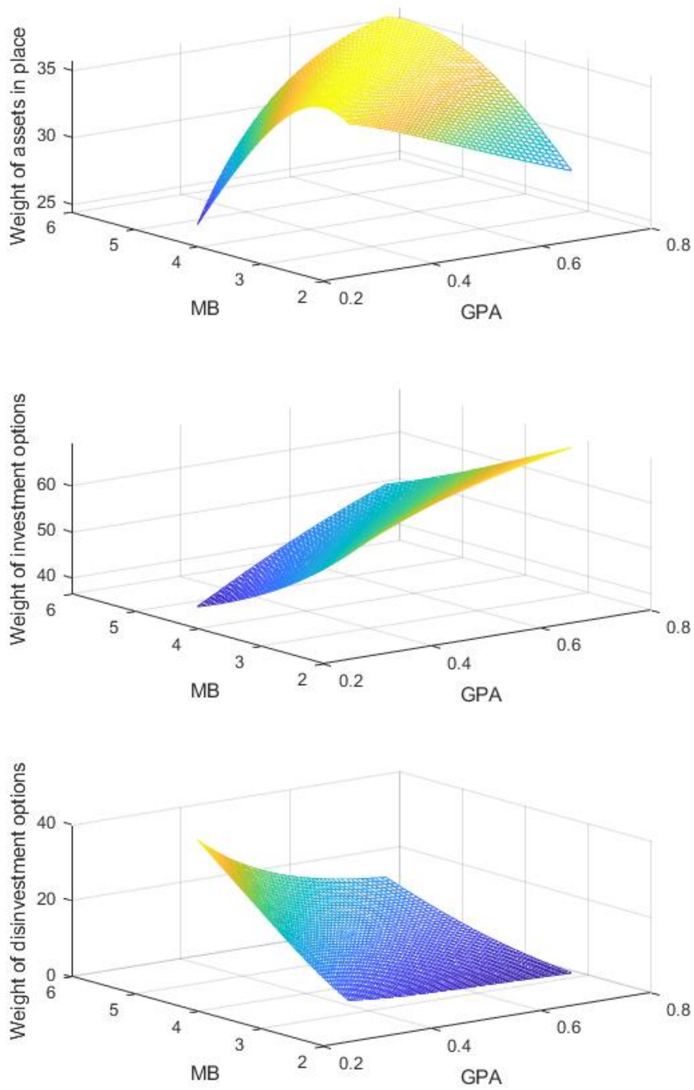


Figure 5. Firm characteristics

This figure plots profitability (GPA) and market-to-book ratio (MB) against idiosyncratic productivity (Z), idiosyncratic dissipation (U), and the ratio of marginal product of capital to investment costs ($\frac{F}{C}$).

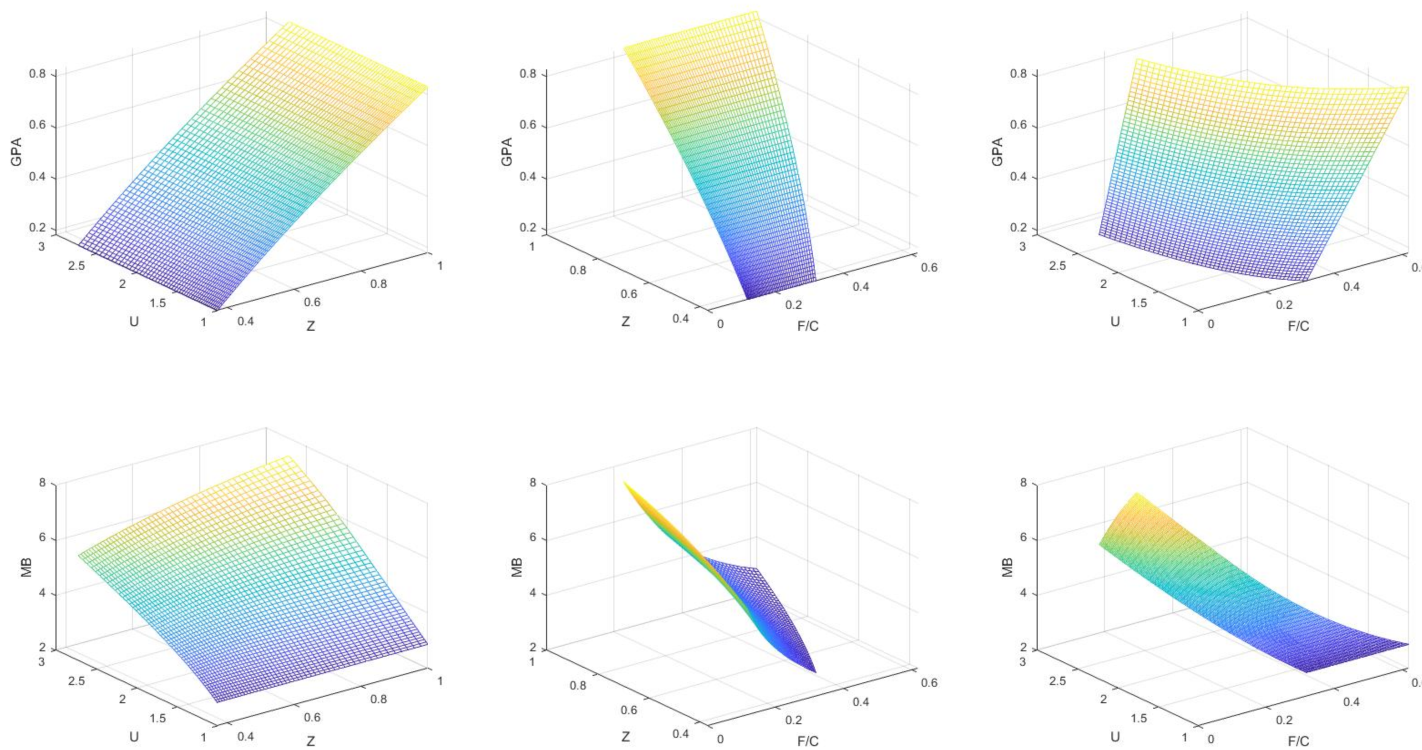


Figure 6. Risk premium and firm characteristics

This figure plots the spread of weights of investment options and disinvestment options against profitability (GPA) and market-to-book (MB) in the good state (solid line) and in the bad state (dashed line).