Profitability, Value and Stock Returns in Production-Based Asset Pricing without Frictions *

Ronald J. Balvers
DeGroote School of Business
McMaster University
Email: balvers@mcmaster.ca

Li Gu
Division of Banking Supervision and Regulation
Federal Reserve Board
Email: li.gu@frb.gov

Dayong Huang
Bryan School of Business and Economics
University of North Carolina at Greensboro
Email: d_huang@uncg.edu

June 3, 2015

Abstract

In a production-based asset pricing model without adjustment costs and with decreasing returns to scale following Brock (1982), stock returns at the firm level are no longer identically equal to investment returns but are determined by profitability, the book-to-market ratio, and the change in future profitability prospects. Although firms with low book-to-market ratios are normally more profitable and profitable firms are predicted to have higher returns, the stylized fact that book-to-market ratios positively forecast returns still holds theoretically, but with specific predicted exceptions. These implications are confirmed empirically.

Profitability, Value and Stock Returns in Production-Based
Asset Pricing without Frictions

1. **Introduction**

We modify the investment-based asset pricing approach of Cochrane (1991, 1996) by building directly on the formulation of Brock (1982) which models capital accumulation on a time-to-build assumption rather than adjustment costs and assumes decreasing returns to scale instead of constant returns. The assumption of decreasing returns allows us to highlight the role of profitability on equity returns and substantially modifies Cochrane's investment-based approach because unlevered stock returns no longer equal investment returns. Investment returns still explain stock returns but only partially: we derive that stock returns are a weighted average of average investment returns (including a profitability markup) and the rate of change in the value of intangible assets, with weights related to the book-to-market ratio. A substantial implication of this result is that it captures theoretically the dual aspects of value documented empirically by Novy-Marx (2013): profitability and the book-to-market ratio have separate positive effects on required stock returns.[1]

Allowing decreasing returns to scale also lets us move away from the convex adjustment costs formulation that has dominated investment-based asset pricing research. While convex adjustment costs are analytically convenient and are helpful in allowing the model to numerically

---

[1] These outcomes may suggest the naïve perspective that the best way to invest is by purchasing "cheap" (high book-to-market) stocks in "good" (high profit rate) companies. However, predictability of stock returns arises here in an efficient market. The positive return impact of book-to-market ratios is a result of tangible asset values being more sensitive than intangible asset values to mean-reverting productivity shocks, and the positive effect of profitability on returns stems from firms becoming more profitable as a reward for taking the riskier production route (as in Berk, Green, and Naik, 1999).

replicate asset price variability and investment dynamics (Jermann, 1998, and Zhang, 2005), they are hard to rationalize as a uniform vital source of economic dynamics. Hall (2004) argues that convex adjustment costs of investment are too small to explain large fluctuations in stock prices. Abel and Eberly (2011) show that a model without adjustment costs can explain both the low sensitivity of investment to the book-to-market ratio and the higher sensitivity of investment to operating profit, which are hard to explain with convex adjustment costs. And the survey by Caballero (1999) makes it clear that investment choices are influenced by convex adjustment costs in some firm-level environments, but by non-convex adjustment costs in others. Convex adjustment costs thus may provide a confounding basis for understanding differences in firm-level investment returns, even if adequate at the aggregate level.[2] The recent work of Bloom (2009), Lin and Zhang (2013), and Belo, Bazdresch and Lin (2013) considers both non-convex and convex adjustment costs in investment and makes progress in replicating firm-level investment dynamics. It, however, relies heavily on numerical methods.

On the other hand, our productivity-based formulation maintains the important strengths of the investment-based formulation.[3] Firms in effect select the riskiness of their operations, and

---

[2] The adjustment cost assumption has been popular in part because it can account for more variability in investment returns and accordingly higher asset price volatility. In our time-to-build framework, however, profitability factors unrelated to investment returns also affect stock prices and these may be highly variable, especially since we interpret the productivity shocks that drive profitability more broadly than total factor productivity. It is not our intent here to calibrate a general equilibrium model to explain the size of the equity premium or match the volatility of asset prices. Remaining outside the confines of the standard real business cycle model, and avoiding explanations that rely on time-series variation in risk premia, we see no a priori reason that our model would have trouble explaining asset price volatility.

[3] To streamline terminology we propose to refer to Brock's (1982) contribution as *productivity*-based asset pricing and to Cochrane's (1991) contribution as *investment*-based asset pricing. Both are special cases of *production*-based asset pricing which focuses on the production side rather than the traditional consumption side to derive implications for asset returns. Production-based asset pricing as broadly interpreted builds on the theoretical work of Brock

observed internal decisions provide a real and current signal as to management's information and intended level of risk exposure. A firm's chosen characteristics provide a timely and accurate measure of risk sensitivities perceived by insiders. As Lin and Zhang (2013) emphasize, it is the promise of investment-based, and more generally production-based asset pricing, to identify the links between production decisions and risk exposure such as to provide estimates of costs of capital that are more precise than those derived from the traditional consumption-based approach. The consumption perspective requires that risk loadings be estimated from past time series of returns with or without ad-hoc conditioning variables; the production perspective determines the structural variables that drive the risk loadings. Lin and Zhang (2013) argue that the production-based view may well turn out to provide superior estimates of costs of capital, with the potential of a paradigm shift in asset pricing.

Several recent papers also relate profitability and stock returns, using elements of the Brock approach. Li, Livdan, and Zhang (2009) present a hybrid of Brock's productivity-based approach and Cochrane's investment-based approach by allowing both decreasing returns to scale and convex adjustment costs. Although they succeed in explaining external financing

---

(1982), Cox, Ingersoll, and Ross (1985), and Berk, Green, and Naik (1999) and has been applied to explain and predict stock returns by Balvers, Cosimano, and McDonald (1990), Hsu (2006), Balvers and Huang (2007), Booth et al. (2008), Lioui and Poncet (2008), Kogan and Papanikolaou (2013), and others. Brock (1982) assumes production with a decreasing-returns-to-scale technology and an array of productivity shocks driving firm decisions. Stock returns depend on production decisions that interact with the firm's exposure to the various productivity shocks. Cochrane (1991) and Restoy and Rockinger (1994), building on the q-theory of Tobin (1969) and Hayashi (1982), show that stock returns are identical to investment returns in an environment with constant returns to scale and convex adjustment costs, thus summarizing the production attributes relevant for determining stock returns as simply the determinants of investment returns. This investment-based approach to asset pricing has stimulated a growing body of empirical work by Cochrane (1996), Gomes, Kogan, and Zhang (2003), Carlson, Fisher, and Giammarino (2004), Zhang (2005), Xing (2008), Liu, Whited and Zhang (2009), Li and Zhang (2010), Lin and Zhang (2013), Hou, Xue, and Zhang (2015), and others. See also the survey of Kogan and Papanikolaou (2012) .

anomalies, they additionally contend that stock returns and profitability (equal to cash flows in their model) are inversely related. This is counterfactual in light of Novy-Marx's (2013) findings and contrasts with the result we obtain in the pure Brock framework. Since the results are largely numerical it is challenging to identify the exact mechanism, but Li et al. (2009) argue along the lines of the pure investment-based approach that higher profitability facilitates investment, and that, in turn, higher investment implies lower investment returns and stock returns. Empirically, Li et al. support this prediction by identifying a positive interaction effect between profit (cash flows) and investment, which affects returns in addition to the investment link by itself. However, they do not control for the effect of profitability by itself, which likely accounts for the discrepancy between their findings and those of Novy-Marx.

Kogan and Papanikolaou (2013) building on Papanikolaou (2011) may be viewed as a structural version of the Brock approach with two types of productivity shocks. By adding a second, investment-specific, productivity shock to traditional approaches (such as Jermann, 1998, Boldrin, Christiano, and Fisher, 2000, Balvers and Huang, 2007, Hsu 2009, and Hsu and Huang, 2009), they generate additional implications and avoid relying on the Solow residual to operationalize productivity shocks. Their model produces a positive link between returns and both book-to-market ratios and profitability: higher growth options imply more exposure to investment-specific technology shocks that mitigate market risk, causing lower average stock returns; and higher profitability implies riskier tangible assets and more exposure to total factor productivity shocks, causing higher average stock returns. But, their model also produces a counterfactual positive correlation between profitability and book-to-market ratios. By explaining the dual dimensions of value and by avoiding the empirical use of Solow residuals,

4

Kogan and Papanikolaou attain some of the objectives of our paper. However, our paper offers a substantially different perspective with alternate economic mechanisms and correlations, distinct testable implications, and a simple model with closed-form solutions for profitability rates, book-to-market ratios, investment returns, stock prices and stock returns. A key implication is that a higher current book-to-market ratio *normally* (i.e., unconditionally) indicates higher future returns, the standard value effect, because the higher book-to-market ratio implies more weight on the tangible value component, which has higher average returns, and less weight on the intangible value component, which has lower average returns.[4] But, *in specific instances* – whenever intangible asset returns are expected to exceed tangible asset returns – the model implies that the book-to-market effect does not hold and is, in fact, reversed as we confirm empirically.

The empirical results support the model predictions. In particular, the book-to-market ratio has the expected positive return effect for the average firm, but as predicted explicitly, is reversed in cases for which returns on intangible assets are expected to exceed returns on tangible assets.

2. **Production, Profitability, and Expected Stock Returns**

We present first an equilibrium model along the lines of Brock (1982), Cox, Ingersoll, and Ross (1985), and Berk, Green, and Naik (1999), that allows prediction of a particular firm's

---

[4] Technically, we show that the risk-adjusted returns on tangible capital must follow a submartingale process, while the risk-adjusted returns on intangible capital must follow a supermartingale process.

required return from market conditions and firm characteristics, utilizing a production-side perspective. The model avoids frictions and makes explicit the impact of decreasing returns, and the associated capital investment decisions, on expected returns and profitability. The impact is time-varying and suggests predictability of realized stock returns based on prior profitability, book-to-market ratios, and growth information.

Firm decisions without frictions

Consider a typical firm maximizing the expected net present value to shareholders with respect to its production and investment choices in each period. The maximized value of the firm V is determined as the present value of dividends to the shareholders:

$$V(\theta_t, K_t) = \underset{I_t}{Max} \left(D_t + E_t[m_{t+1} V(\theta_{t+1}, K_{t+1})] \right), \tag{1}$$

which is the standard Bellman Equation of dynamic programming. The dividends paid by the firm in each period t are denoted by D_t. The productivity variable θ_t and the available capital stock K_t are the state variables which are jointly sufficient for determining firm value.[5] The aggregate stochastic discount factor is given as m_{t+1}, and $E_t(\cdot)$ carries the subscript t to indicate that expectations are conditional on all currently available information.

The value of the firm $V_t \equiv V(\theta_t, K_t)$ is the equity value of the firm before dividends: $V_t = P_t + D_t$, with $P_t = E_t[m_{t+1} V(\theta_{t+1}, K_{t+1})]$ the *ex-dividend* equity value of the firm. Given the

---

[5] It is clear from the specification in equations (1) - (4) that, indeed, the productivity variable θ_t and the available capital stock K_t are sufficient state variables for the value of the firm.

stock return, $r_{t+1}^S = (D_{t+1} + P_{t+1} - P_t)/P_t$, equation (1) can be rewritten as $1 = E_t[m_{t+1}(1 + r_{t+1}^S)]$, implying that risk- and dividend-adjusted stock prices follow a martingale process so that pricing the equity value of the firm according to equation (1) rules out arbitrage opportunities.

In spite of the frictionless model specification, we derive that the (market) value of the firm exceeds its book value (the outstanding value of its tangible capital stock K_{t+1}). The difference is the present value of the residual income that stems from DRTS, implying that only the marginal investment decisions fails to create profit while all infra-marginal investment units generate positive profits. We refer to the difference between firm value and (tangible) asset value as the *intangible asset* value. It is akin to the growth options in a model with frictions and is also related to the disembodied future value created by research and development expenditure, except that we do not model the R&D decisions. (See Lin, 2012, and Li, Liu, and Xue, 2013, for models in which R&D endogenously affects technology). In our setting, intangible "assets" thus generically represent the present value of future economic profits due to DRTS, whether the underlying disembodied asset is the product of research, goodwill, entrepreneurial talent, location, or otherwise.

We assume that net operating income, Y_t, is determined by a simple power function of the firm's current physical capital stock, K_t, and by the stochastic level of productivity, θ_t: $Y_t = \theta_t K_t^\alpha$. Dividends D_t are then given as:

$$D_t = \theta_t K_t^\alpha - I_t, \tag{2}$$

with I_t the current level of capital investment. For later use we define $i_t = I_t / K_t$ as the current level of investment per unit of capital that will contribute to next period's capital stock.

Productivity θ_t here is to be interpreted quite generally. It is a conflation of technology shocks and other disembodied productivity shocks together with miscellaneous exogenous factors that affect the profitability of the firm, such as changes in the competitive environment, and input costs (see, e.g., Novy-Marx, 2013, for a similar characterization of productivity). Following Brock (1982) we may view θ_t as a vector containing a multitude of systematic productivity shocks. We choose not to do so formally to keep the notation simple and because this generalization would add few new insights given our focus.

The interpretation of Y_t as operating income means that we have implicitly accounted for the impact of labor and additional inputs on production and costs, which are only represented by the exogenous influence of wages and other input costs on income, captured by the productivity level θ_t. We assume decreasing returns to scale so that $0 < \alpha < 1$. In this we follow Brock (1982) but differ from Cochrane (1991) and others in the investment-based asset pricing literature, with the critical implication that investment returns are distinct from stock returns. We express the degree of decreasing returns to scale as $\varsigma = (1-\alpha)/\alpha > 0$ so that higher ς represents stronger decreasing returns to scale and higher profitability.

Capital evolves without adjustment costs according to the standard linear specification:

$$K_{t+1} = (1-\delta)K_t + I_t \tag{3}$$

We interpret investment broadly to include changes in working capital so, under clean surplus accounting, capital becomes conceptually equal to total assets and the book value of the firm.

The exogenous productivity indicator follows a linear Markov process:

$$\theta_{t+1} = (1-\rho)\bar{\theta} + \rho\theta_t + \sigma e_{t+1} + \eta_{t+1}, \tag{4}$$

Where η_t is a firm-specific i.i.d. random variable with arbitrary distribution that has mean zero and variance σ_η^2, and e_t is an aggregate i.i.d. variable with an arbitrary distribution that is standardized to have zero mean and unit variance, and is independent of η_t. The autoregressive formulation for productivity is inherited by the process for profitability (although inversely) and captures two features of the data: *persistence of profitability* and *mean reversion of profitability* at the firm level (Fama and French, 2000 and 2006).

The aggregate shock e_t is the only systematic risk (although with minor complications we could model a vector of such systematic productivity shocks as in Brock, 1982) and we specify the stochastic discount factor exogenously as

$$m_t = \frac{1-h e_t}{1+r} . \tag{5}$$

It follows that $1/(E_t m_{t+1}) = 1+r$ so that r represents a constant risk free rate; h is the constant risk premium of the systematic productivity risk. The stochastic discount factor parameters h and r are the only aggregate parameters in the model. *All other parameters may differ across firms.* The exogenous stochastic discount factor is in the tradition of Berk, Green, and Naik (1999) and allows us to focus on differences in factor risk sensitivities of individual firms as determinants of returns, while taking the aggregate values for the risk free rate and the risk premium as given.

9

Optimal investment

Maximization of the value of the firm in equation (1), subject to equations (2), (3), and (4) with respect to investment, yields the first-order condition for the investment choice (barring time subscript t, all subscripts indicate partial derivatives):

$$1 = E_t[m_{t+1}V_K(\theta_{t+1}, K_{t+1})] \ . \tag{6}$$

The optimal capital stock is subsequently found in Appendix A as

$$K_{t+1}^* = [A(z + \rho\theta_t)]^{\frac{1}{1-\alpha}}, \quad \text{with } A \equiv \alpha/(r+\delta), \ z \equiv (1-\rho)\bar{\theta} - h\sigma. \tag{7}$$

The optimal investment ratio i_t^* may be inferred directly from equations (3) and (7).

3. **Implications**

The model solutions allow us to relate stock returns to investment returns, profitability measures, and value measures, providing some new testable implications in the process.

Investment returns and stock returns

In equation (6), $1 = E_t[m_{t+1}V_K(\theta_{t+1}, K_{t+1})]$, which implies that the gross investment return $1 + r_{t+1}^I \equiv V_K(\theta_{t+1}, K_{t+1})$ can be viewed as a regular asset return because $E_t[m_{t+1}(1 + r_{t+1}^I)] = 1$. The excess investment return is derived straightforwardly in Appendix A:

$$r_{t+1}^I - r = \left(\frac{\sigma(h + e_{t+1}) + \eta_{t+1}}{z + \rho\theta_t}\right)(r + \delta). \tag{8}$$

The expected excess investment return equals $E_t r_{t+1}^I - r = (r + \delta)\sigma h/(z + \rho\theta_t)$.

In the constant returns to scale (CRTS) framework of Cochrane (1991, 1996) the marginal investment return equals the average investment return which in turn equals the stock return: $r_{t+1}^I = r_{t+1}^S$. But in our decreasing returns (DRTS) model the two are not identical and a comparison is instructive. Compare $1 + r_{t+1}^I \equiv V_K(\theta_{t+1}, K_{t+1})/p_t^K$ and $1 + r_{t+1}^S \equiv V(\theta_{t+1}, K_{t+1})/P_t$. Under CRTS (with or without adjustment cost) it is necessarily true that both $V_K(\theta_{t+1}, K_{t+1}) = V(\theta_{t+1}, K_{t+1})/K_{t+1}$ and $P_t = p_t^K K_{t+1}$ where, in the case of adjustment costs, p_t^K represents the full price of a capital good – the cost of purchasing (set equal to 1 here) *and* installing the capital good. It follows then that $r_{t+1}^I = r_{t+1}^S$.

Under DRTS there are two sources of difference because (a) average returns to capital and marginal returns to capital differ, $V(\theta_{t+1}, K_{t+1})/K_{t+1} \neq V_K(\theta_{t+1}, K_{t+1})$, and because (b) book-to-market values deviate from one: $b_t = p_t^K K_{t+1}/P_t \neq 1$ implying $V(\theta_{t+1}, K_{t+1})/K_{t+1} \neq V(\theta_{t+1}, K_{t+1})/P_t$. As shown in Appendix B and as also consistent with Abel and Eberly (2011) in a related framework (but focusing on investment rather than asset prices): b_t in this model is always lower than one (setting p_t^K equal to 1 here, $b_t = K_{t+1}/P_t < 1$) because the firm has a strictly positive intangible asset value (in spite of not facing adjustment costs). The intangible asset value is the expected present value of future streams of residual income. The residual income is the abnormal profit, the net income after adjusting for the opportunity cost of capital, and here arises exclusively from the profit markup caused by decreasing returns to scale.Thus, variations in the scale of the firm and in the book-to-market value of the firm cause stock returns to deviate from

11

investment returns. These differences are intriguing because they correspond to the well-documented size and value effects, respectively.

Determinants of stock returns

Appendix B derives an expression for stock returns in relation to investment returns:

$$r_{t+1}^S = \pi_{t+1} b_t + g_{t+1}(1 - b_t), \qquad \pi_t \equiv r_t^I + \varsigma(r_t^I + \delta). \tag{9}$$

The book-to-market ratio $b_t = b(\theta_t) = K_{t+1}/P_t$ represents a predetermined weight $0 < b_t < 1$, and $g_{t+1} = g(\theta_{t+1}, \theta_t)$ is the rate of increase in the firm's intangible asset value. We obtain the *average* investment return as $\pi_t = r_t^I + \varsigma(r_t^I + \delta)$ which equals the *marginal* return on investment r_t^I (stated in equation 8) plus a markup that results from decreasing returns to scale. In their comparable model (which, however, does not allow derivation of stock returns and prices), Abel and Eberly (2011, equation 3) show that the markup can alternatively be interpreted as the profit arising from market power if the assumption of DRTS is replaced by an assumption of finite demand elasticity. The markup here is the product of the degree of decreasing returns, $\varsigma = (1 - \alpha)/\alpha$, and the appropriate user cost of investment equal to the investment return (the risk free rate plus a compensation for investment risk) plus the depreciation rate, $r_t^I + \delta$. Essentially stock returns are a weighted average of the return on tangible assets (the average investment return) and the return on intangible assets (the rate of increase in the present value of future residual income) with the weights equal to b_t and $1 - b_t$. Note that both the mean value of the average investment return and the risk of the average investment return increase directly in the degree of decreasing returns to scale, ς.

12

The intuition for why unlevered stock returns deviate from investment returns is that decreasing returns imply that the (marginal) investment return affecting real decisions is below the average return paid out to stockholders; i.e., average investment returns are larger than marginal investment returns. Additionally, even the average investment return is only the return from tangible assets going to stockholders and should be weighted by the tangible part of total equity value. Another part of the asset's market value affecting stockholder returns derives from profitability prospects. The growth rate of this intangible asset provides a second component of the stock return, weighted by the share of intangible assets in total equity value.

In our model, profit margins and investment returns are high when investment exposes the firm to high systematic risk. High profit margins thus proxy for high σ (see equation 8), signifying high levels of systematic risk and high (productivity) betas.[6]

Profitability and stock returns

The average investment return can be interpreted as a profitability measure. Rearranging the definition in equation (9), $\pi_t + \delta = (1+\varsigma)(r_t^I + \delta)$ represents a gross return which is shown in Appendix B to equal the gross profit of the firm as a fraction of the initial book value, $\pi_t + \delta = Y_t / K_t$, which equals operating income divided by the book value of assets. It is similar in spirit to Novy-Marx's measure of gross profitability, although our theoretical model is not detailed

---

[6] Hou, Xue, and Zhang (2015) find that an investment-based model with an investment factor and a profitability factor in addition to the traditional market and size factor can explain a large variety of financial market anomalies. Their motivation for separate investment and profitability factors is based on a framework in which investment returns and stock returns are equal but they break up the investment return into parts depending on the return on equity and the investment-to-capital ratio. They treat these components as separate systematic risk factors rather than indicators of risk sensitivities.

enough to distinguish this measure from other profitability measures such as the return on assets.[7] Thus,

> **PROPOSITION 1.** Given absence of frictions and with decreasing returns to scale technology, expected stock returns increase in expected gross profitability: $\partial E_t(r_{t+1}^S) / \partial E_t(\pi_{t+1}) = b_t > 0$.

This follows directly from equation (9). Akin to the approach in Berk, Green, and Naik (1999), higher profitability means that the firm is more sensitive to the risk of current productivity shocks and therefore has higher expected return. For given technology and productivity level, the firm has riskier investment projects (higher σ) which imply higher profitability from equation (8), as well as higher expected return from equation (9).

Novy-Marx (2013) finds that the gross profitability of firms – current revenue minus the cost directly attributable to current revenue generation – provides another dimension of value and has significant forecast power for returns that is separate from the traditional book-to-market effect. Empirically he finds in double-sorting US firms from 1963-2010 that the monthly returns between high and low profitability quintiles is 0.68% per month, averaged over all book-to-value quintiles; and the monthly return between high and low book-to-market quintiles is 0.54%, averaged over all profitability quintiles.

---

[7] The model, however, is detailed enough to eliminate cash flow measures (such as Free Cash Flow) as a proxy for profitability in this context because investment expenditures must be subtracted to measure cash flows but are not subtracted here in deriving the average investment returns.

It is important to emphasize that Novy-Marx does not only find that higher profitability raises stock returns, but also that the effect of profitability on returns is larger when controlling for book-to-market values. The reason is the iconic result that book-to-market ratios also affect returns positively, together with a negative correlation between profitability and book-to-market (value firms are less profitable). The negative profit-value correlation is essential for explaining Novy-Marx's double-sorting results but is not always implied theoretically. For instance, Kogan and Papanikolaou (2013) requires a positive profit-value correlation. On the other hand, our model requires the appropriate negative profit-value correlation because high current profitability is associated with a high present value of future profitability, which defines growth firms in our context. Equation (9) shows, as may be seen most easily by setting the expected rate of increase in the firm's intangible asset value equal to its long-run average of zero, $E_t(g_{t+1}) = 0$, that higher expected profitability $E_t(\pi_{t+1})$ raises expected return $E_t(r_{t+1}^S) = E_t(\pi_{t+1})b_t$, holding the book-to-market ratio b_t constant. However, without controlling for b_t, because the higher $E_t(\pi_{t+1})$ tends to be associated with lower b_t this will on average offset part of the simple profitability impact. Thus, in our model, single sorting by $E_t(\pi_{t+1})$ generates lower average returns than double sorting by $E_t(\pi_{t+1})$ and b_t, as Novy-Marx finds.

The effect of profitability on stock returns may also be related to the effect of operating income on stock returns that Li, Livdan, and Zhang (2009) find numerically in their model and confirm empirically. They employ the Brock model but add convex investment adjustment costs. For given positive investment, they find that higher operating income relative to assets has a *negative* impact on stock returns. This result appears to conflict with the empirical Novy-Marx

15

(2013) result and with our theoretical result, both implying a *positive* impact of operating income on stock returns. The likely reason is that in their empirical analysis Li et al. (2009) consider only the interaction between operating income and investment, and not operating income in isolation.

The modification from marginal investment returns to average investment returns in linking to stock returns is analogous to the impact of operating leverage emphasized by Carlson, Fisher, and Giammarino (2004) and Novy-Marx (2011), but here represents a profit markup derived from the extent of the decreasing returns ς measured by $\frac{1-\alpha}{\alpha}$, and has diametrically opposed empirical implications. Novy-Marx (2013, p.16) points out that in existing models (Carlson et al., 2004, Zhang, 2005, and Novy-Marx, 2011) operating leverage and risk, and hence expected returns, are increasing in fixed costs. Therefore, since higher fixed costs lower profits, expected returns are *negatively* linked with profitability, implying that profitable firms should underperform unprofitable firms in the stock market. In our model, however, the "operating" leverage arises directly from the inherent profit markup and is linked *positively* to profitability and stock returns (though negatively to the productivity level).

Kogan and Papanikolaou (2013) likewise generate the observed positive correlation between profitability and expected stock returns that we find in our model. The positive correlation arises because firms that are currently more profitable have less of their value linked to future investment and they are therefore less susceptible to investment-specific productivity shocks. In turn, investment-specific productivity shocks are found to mitigate market risk so that highly profitable firms are riskier. This intuition relies on highly profitable firms typically being

16

value firms and appears to be inconsistent with the negative correlation between profitability and book-to-market ratios identified by Fama and French (1995). In contrast, in our model highly profitable firms, while also associated with higher stock returns, normally continue to be profitable and thus have a high intangible asset value, making them growth firms.

Model predictions relating to the value effect

Separately, b_t, the book-to-market ratio also affects the firm's expected stock return. The effect is positive and magnifies the profitability effect if the average investment return $r_t^I + \varsigma(r_t^I + \delta)$ exceeds the rate of increase in intangible assets, g_{t+1}. This is highly likely since the average return on tangible assets is generally strongly positive while the *rate of increase* in intangible assets must be zero in the unconditional average because the productivity shocks are stationary, and, in fact, should be negative after risk correction as we argue next.

The risk-adjusted stock price must follow a martingale process as is implied by $1 = E_t[m_{t+1}(1 + r_{t+1}^S)]$. Remarkably, while the stock return in equation (9) is a convex combination of the average investment return and the appreciation rate of intangible assets, it is *not* true that the average investment return and the intangible asset appreciation rate follow martingale processes. Because $E_t[m_{t+1}(1 + r_{t+1}^I)] = 1$ it follows from $\pi_{t+1} = r_{t+1}^I + \varsigma(r_{t+1}^I + \delta)$ that $E_t[m_{t+1}(1 + \pi_{t+1})] > 1$ for $\varsigma > 0$. In mathematical terms, the average investment value follows a strict *sub*martingale process. Accordingly, $E_t[m_{t+1}(1 + g_{t+1})] < 1$, so that risk-adjusted intangible assets follow a strict *super*martingale process (see Appendix B). The reason the value of intangible assets (adjusted for risk and time value) is expected to decrease over time is that here

17

the intangible asset consists of the present value of future profit markup which produces profits in the next period that are counted as part of the average investment payoff, not the intangible asset payoff.

The practical implication is that, in typical cases, value effects must be positive: in equation (9) higher b_t implies that stock return is higher by an extent related to the positive difference between the returns on a submartingale and a supermartingale (which is positive in the time series average). The value effect therefore arises because value stocks have more weight on tangible returns which normally are larger than the growth rate of the intangible component. Our model accordingly provides a direct explanation for the Novy-Marx results: the comprehensive value returns arise because profitable firms have high-profit-margin production projects with accompanying high exposure to systematic shocks, and the effect is magnified in "value" firms that also have more tangible assets relative to intangibles assets.

A different implication of our theory, a further refinement of the value effect not tested by Novy-Marx (2013), is that the standard book-to-market premium should be the *opposite* (i.e., negative) for the (presumably small, see Appendix B) segment of firms that has rates of increase in intangible assets expected to exceed the return on tangible assets. For these firms a *higher* book-to-market ratio implies more weight on the relatively low tangible asset return, *lowering* the overall stock return, entailing an inverse value premium. Seen from the risk perspective, these firms necessarily have relatively low exposure to productivity risk in their tangible assets (assets in place) compared to their intangible assets (present value of future residual income), making value stocks less risky within this group of firms. In summary,

18

PROPOSITION 2. Book-to-market ratios affect expected stock returns (given decreasing returns and for given profitability): $\partial E_t(r_{t+1}^S)/\partial b_t = E_t(\pi_{t+1} - g_{t+1})$. The effect is (a) *positive* unconditionally: $E(\pi_{t+1} - g_{t+1}) = E(\pi_{t+1}) > 0$, and (b) *negative* only for particular firms with $E_t(g_{t+1}) > E_t(\pi_{t+1})$, for which the increase in intangible asset values is expected to exceed the return on tangible assets.

The result follows from equation (9) and since θ_t, and thus intangible value $G(\theta_t)$, is stationary, causing $E(g_{t+1})$, the unconditional growth rate of $G(\theta_t)$, to be zero (see Appendix B).

Normally a higher book-to-market ratio implies a higher expected return: Firms with higher book-to-market ratios have relatively more tangible assets. A current productivity shock then has a larger impact on market value because tangible assets are impacted directly and strongly by the shock, whereas intangible assets are affected less strongly because they capitalize future impact on profit which diminishes over time as a result of the mean reversion of θ_t assumed in equation (4).

But, in particular cases, higher book-to-market values generate *lower* expected excess returns. Namely, for the subset of firms that have expected rates of increase in intangible asset value exceeding expected average investment return, it must be true that a larger book-to-market ratio b_t (a greater current proportion of tangible assets) implies less risk and lower expected stock return. The reason is quite subtle. For these firms, intangible assets have *higher* exposure to productivity risk: even though these firms generate more profits at future times when the impact

19

of the current productivity shock is lessened, the high growth rate of the profitability increases the importance of these productivity shocks, even if lessened. Put somewhat imprecisely, firms with a growth rate of future profitability higher than the depreciation rate of a productivity shock have larger exposure to a current productivity shock in their intangible assets (relative to their tangible assets). Hence in these cases higher book-to-market ratios, meaning relatively more tangible assets, imply less exposure to productivity risk and accordingly lower expected returns.

Previous literature explaining the value premium has relied on tangible assets being across the board riskier than intangible assets. In models with adjustment costs, the intangible assets generally take the form of growth options. Zhang (2005) argues that value firms are riskier because they are stuck with high (convex) adjustment costs when marginal utility is high. Cooper (2006) and Cooper, Gerald, and Wu (2005), assuming non-convex adjustment costs, argue that value firms are riskier because they can expand costlessly and benefit more strongly from positive shocks when their production capacity is high. Both arguments hinge on adjustment costs and the conditional variation of risk premia. Li (2014) combines the mechanisms in Zhang (2005), asymmetric convex (quadratic) adjustment costs, and Carlson et al. (2004), non-convex adjustment costs affecting operating leverage, to explain the value premium as generated by larger exposure to productivity risk of firms with a high fraction of tangible assets (value firms).

The explanation of Papanikolaou (2011) and Kogan and Papanikolaou (2013) avoids both adjustment costs and conditional variation of risk premia and maintains reasonably that growth firms are more susceptible to investment-specific productivity shocks which have *negative* risk

20

premia (under plausible assumptions). On the other hand, in Garlappi and Song (2013a) investment-specific technology shocks have a *positive* risk premium while growth firms have smaller loadings on these shocks, also generating a value premium, but from an opposite channel. Garlappi and Song (2013b) argue that the direction of the impact of investment-specific shocks differs by firm depending on the level of competition in the firm's industry. These opposing inferences suggest the likelihood that the value effect is not universally present and may be reversed in particular well-defined cases. Identifying such cases may be helpful in distinguishing alternative theoretical perspectives, but previous literature has not explored this line of inquiry.

Our model provides a complementary perspective that applies without a need to distinguish between the types of productivity shocks. Tangible assets are usually more sensitive to current shocks because their productivity depends fully on these shocks, whereas intangible assets become productive at later times when parts of the current shocks have expired. However, when intangible assets are expected to grow at a substantial rate, the dwindling over time of the current shock is offset by increased importance of future profit realizations. For instance, temporarily lower input costs (a positive 'productivity' shock) are crucial for firms with many assets in place. But for firms with increasing profit opportunities, lower input costs have more future importance even if they revert eventually, and increased weight on this component now raises overall risk.

Novy-Marx (2013) finds that the traditional value effect is not merely separate from the profitability effect, but is in fact enhanced when he controls for profitability. Kogan and

Papanikolaou (2013) also confirm this observation which their model can match quantitatively, though book-to-market ratios and profitability rates are *positively* correlated in their model. We generate the same result but stemming from the fact that book-to-market ratios and profitability rates are *negatively* correlated in our model, as is consistent with Fama and French (1995). The negative correlation arises from a higher productivity level implying simultaneously a lower profitability rate (more capital lowers the average return on capital) and a higher book-to-market ratio (current productivity normally boosts raise tangible asset values more than intangible asset values due to mean reversion). The negative correlation between book-to-market ratios and profitability rates means that a high profit rate, for instance in part due to high σ and in part due to low θ_t, is associated with a lower book-to-market ratio which partly or fully offsets the effect of the profit rate on stock returns. Sorting by book-to-market ratio in addition to profitability, avoids the cases in which profitability is high for reasons having relatively small impact on expected returns (low θ_t for instance), implying a larger profitability premium when portfolios are double sorted.

Productivity, growth, and stock returns

Higher current productivity θ_t at a particular firm forecasts *lower* expected stock returns $E_t(r^s_{t+1})$ for a given book-to-market ratio. This follows from equation (9) since both tangible and intangible asset returns are expected to drop when the productivity level rises. The drop in the tangible asset return with higher productivity, $\partial E_t(\pi_{t+1})/\partial \theta_t < 0$, follows because $\pi_{t+1} + \delta = (1+\varsigma)(r^I_{t+1} + \delta)$ and $E_t r^I_{t+1} - r = (r + \delta)\sigma h/(z + \rho\theta_t)$. The expected investment return (and

therefore expected profitability) is lower because the marginal value of capital is lower at the higher capital level chosen in response to a higher productivity level.

A drop in the intangible asset return with higher productivity will not occur in all cases, but holds at least in the neighborhood of the steady state: $\partial E_t(g_{t+1})/\partial\theta_t\mid_{\theta_t=\bar\theta}<0$, and is simply a consequence of the mean reverting process for θ_t (see Appendix B).

The current productivity level θ_t is *positively* related to the capital stock K_{t+1} as shown in equation (7): the firm increases its capital when it is more productive. Accordingly, since from equation (3) $K_{t+1}=K_t(1-\delta+i_t)$, the proportional level of investment $i_t=I_t/K_t$ and the book value of the firm K_t together determine K_{t+1}, both are positively related to productivity. In addition the only time-varying element (negatively) affecting $E_t(g_{t+1})$ is the productivity level. As shown in Appendix B, these facts imply in combination that

> **PROPOSITION 3.** In the neighborhood of $\theta_t=\bar\theta$ (steady state productivity): (a)
> Investment ratios affect expected stock returns negatively, all else equal:
> $\partial E_t r_{t+1}^s(b_t,E_t\pi_{t+1},i_t,K_t)/\partial i_t<0$. (b) Firm asset size affects expected stock returns
> negatively, all else equal: $\partial E_t r_{t+1}^s(b_t,E_t\pi_{t+1},i_t,K_t)/\partial K_t<0$.

Note that the comparative statics impact of changes in firm asset size K_t and market size P_t is equivalent given that we keep i_t and b_t constant. Hence, an evident corollary of Proposition 3(b) is that $\partial E_t r_{t+1}^s(b_t,E_t\pi_{t+1},i_t,P_t)/\partial P_t<0$.

Explicit solutions for firm-level expected stock returns over time

An explicit solution for the value function, stock price, and expected return requires *(i)* specialization of the model to a linear-quadratic case:

$$\alpha = 1/2, \tag{10}$$

and *(ii)* assuming normal distributions for e_t and η_t. The method of undetermined coefficients in Appendix C then provides an explicit solution for the stock price of the firm and its expected excess return as:

$$P(\theta_t) = B[\rho^2\,\theta_t^2 + c_1\,\theta_t + c_0 + (z + \rho\theta_t)^2(1 + r - \rho^2)/(r + \delta)]. \tag{11}$$

$$E_t r_{t+1}^S - r = h\sigma\left(\frac{2(1+r)(z+\rho\theta_t) + c_1 + 2\rho^2 h\sigma}{\rho^2\,\theta_t^2 + c_1\,\theta_t + c_0 + (z+\rho\theta_t)^2(1+r-\rho^2)/(r+\delta)}\right), \tag{12}$$

with:

$$c_1 = \frac{2\rho z(1+r)}{1+r-\rho}, \quad c_0 = \frac{z^2(1+r+\rho)(1+r)}{r(1+r-\rho)} + \rho^2[(1-h^2)\sigma^2 + \sigma_\eta^2], \quad B = \frac{1}{4(r+\delta)(1+r-\rho^2)}.$$

Interpretation and discussion

For intuition about the closed-form expression in equation (12), we consider a few simple cases.

Some of the comparative statics results are clear cut. Note first that the constant B does not show up in the expected return expression as it multiplies both the price and the expected net payoff. The directional effect of depreciation is clear from equation (12): $d(E_t r_{t+1}^S - r)/d\delta > 0$. An increased depreciation rate raises the user cost of capital, here $r + \delta$, thus lowering the optimal capital stock, future profitability, and the price per share of the firm. The sensitivity to the factor risk increases as a result of the diminishing returns to investment: at a decreased

capital stock, the marginal value and average value of capital are larger so that given productivity shocks have a larger impact.

Higher variance of idiosyncratic productivity shocks σ_η^2 raises c_0 and therefore increases the stock price and decreases expected returns. This is consistent with the empirical results of Ang, Hodrick, Xing, and Zhang (2006) which pose the puzzle that firms with higher idiosyncratic risk have lower average returns. In our model higher idiosyncratic variability enhances the option value arising from a firm's flexibility to ramp production up or down with changes in productivity: a positive shock is amplified by adding capital, and a negative shock is mitigated by shedding capital. As evident from (A13) and (A14) and the definition of c_0 in Appendix C, higher idiosyncratic risk raises stock prices for given systematic operational risk. Firms with higher idiosyncratic risk have lower average stock returns because the higher prices for the given systematic risk imply less systematic operational risk per dollar invested.

For other parameters, the comparative statics are generally ambiguous. Typically, h (the risk premium for the aggregate productivity shock) and σ (the sensitivity of firm productivity to the aggregate productivity shock) increase the expected excess return. In addition, h usually raises the firm's beta (risk sensitivity). The reason is that an increase in h raises the cost of capital which reduces the optimal capital stock, raises marginal and average productivity, and thus the sensitivity to productivity shocks.

Furthermore, an increase in the persistence of the productivity shock, ρ, typically has a positive effect on the expected excess return because it raises the strength of the firm's reaction to initial shocks (raising beta). Increases in both the long-run productivity level $\bar{\theta}$ and the

current productivity level θ_t raise the optimal capital stock, and therefore lower the marginal productivity of capital. The latter reduces the marginal and average impact of productivity shocks and, hence, risk sensitivity and expected stock returns.

4. **The Data**

Monthly firm-level cum-dividend returns are taken from the Center for Research on Stock Prices (CRSP), and the three-month T-bill rate is subtracted to calculate excess returns. All firm-level accounting data are obtained from the COMPUSTAT annual file. Our sample starts in July 1963 and ends in December 2014 and includes an average of 698 firms per month that meet our data criteria. To obtain portfolio-level characteristics, we first calculate annual firm-level characteristics and then take the means of the characteristics of all stocks that belong to each portfolio as the portfolio-level characteristic.

We consider various characteristics studied in Fama and French (2008): size (price per share times shares outstanding, ME), the book-to-market ratio (the book value divided by the market value, BM), the gross profit margin (the gross profit scaled by total assets, GP), and momentum (the cumulative excess return over the past 12 months, MOM). The model's investment-to-capital ratio i (IK) is calculated as the change in property, plant and equipment plus the change in inventory normalized by lagged total assets. Research and Development growth (RDG) is the three year average of, the year-over-year change in R&D expenditures normalized by lagged total assets. We delete firms whose market capitalization are below the NYSE 10 percentile size breakpoints, and use value weighting in constructing portfolios.

26

Table 1 provides descriptive statistics of the key variables. In Panel A we first calculate the means of these variables for each firm and then report their cross-sectional summary statistics. Median GP across all firms equals 48.8%, with 80% of the firms in the range from 16.4% to 126.3%. The median IK ratio is 5.9%, with 80% of the firms in the range from -2.2% to 22.1%. RDG ranges from -0.1% to 3.0% with a median of 0.2%. After fitting an autoregressive model of order one to each firm's BM, GP, IK, and RDG at the cross-section, we find in Panel B that median AR-1 coefficients are 0.44, 0.59, 0.13 and 0.52, respectively, suggesting that the levels of book-to-market, and profitability, and growth rates of R&D spending are quite persistent while the investment-to-capital ratio at the firm level is less so. The correlations between the variables are in Panel C. None of the variables are particularly highly correlated, with highest (in absolute terms) correlation of -0.16 between BM and GP (value firms are less profitable in line with Fama and French, 1995). Annual stock returns RET are negatively correlated with IK and ME, and positively correlated with GP, BM, RDG, and MOM.

5. Empirical Results

Fama MacBeth Regressions

To provide basic tests of our propositions, Table 2 presents Fama and MacBeth (1973) regression results of using year t-1 firm-characteristic variables to predict monthly firm-level excess returns for July of year t to June of year t+1. In Panel A, we first consider Propositions 1, 2(a), and 3(a) asserting that BM and GP affect expected returns positively, and that IK affects expected returns negatively. These propositions are confirmed. IK has a negative coefficient of -0.0088 while BM and GP have positive coefficients 0.0028 and 0.0084, respectively. All are

significant at the 1% level. When size (ME) is added in the regression, conform Proposition 3(b), its coefficient is negative as expected but not significant. We also add momentum MOM to see if our perspective explains the momentum effect. While momentum remains a (marginally) significant indicator of future excess returns, the explanatory power of GP, IK, and BM remains intact, when it is included.

We also present in Panel A a test of equation (9). Taking conditional expectations gives $E_t r_{t+1}^S = E_t(\pi_{t+1})b_t + E_t(g_{t+1})(1-b_t)$. Given the autocorrelation in π_{t+1} (Panel B of Table 1, and Fama and French, 2000), we proxy for $E_t(\pi_{t+1})$ by π_t (GP). We also proxy for $E_t(g_{t+1})$, the expected growth in the firm's intangible value component, by using growth in R&D expenditures of the firm (RDG) as a prior indicator.[8] The motivation is that current R&D spending should directly impact future profitability. The only other variable is b_t (BM). All three variables should affect returns positively, as long as $E_t(\pi_{t+1}) > E_t(g_{t+1})$. This prediction is confirmed strongly for BM and GP. The coefficient on RDG, our proxy for $E_t(g_{t+1})$, is also positive, but not significant. The last equation displayed in Panel A confirms that these results remain even if we include ME and MOM.

---

[8] To validate RDG as a proxy for intangible asset growth we checked if RDG(t) indeed forecasts g(t+1). In results available from the authors we confirm that RDG(t) positively and significantly forecasts g(t+1) (significant at the 1% level). On the other hand, g(t) has no forecast power for g(t+1). To check the robustness of our results we also consider an alternative proxy for (expected) intangible asset growth g(t+1) by regressing realized future intangible asset growth on all available variables. We then use fitted intangible asset growth instead of RDG. The resulting alternative to the three-way sort discussed next also produces the predicted reversal but it is smaller and is subject to the look-ahead bias inherent in utilizing fitted values. The results are available from the authors.

A more specific test of equation (9) is presented in Panel B. If we proxy linearly: $E_t(\pi_{t+1}) = c_0 + c_1\pi_t$, $E_t(g_{t+1}) = c_2 + c_3 RDG_t$ and use $b_t = BM_t$ (level, not log) and $\pi_t = GP_t$, then, taking expectations, equation (9) can be restated as:

$$E_t(r_{t+1}^S) = c_2 + c_3 RDG_t + (c_0 - c_2)BM_t + c_1(GP_t \cdot BM_t) - c_3(RDG_t \cdot BM_t). \qquad (13)$$

Thus, we predict that the coefficient on RDG is positive, the coefficient on the cross-product GP·BM is positive, and the coefficient on the cross-product RDG·BM is negative. Notably, there is no prediction for the impact of BM in this specification. Panel B confirms the signs of these predictions but both the RDG and RDG·BM coefficients are not significant. The coefficient on BM is close to zero implying that its importance in other regressions may alias for interaction effects – in particular the interaction with GP – which is the key variable in this regression as is to be expected by taking unconditional expectations in equation (8) if $E(g_{t+1})$ is close to zero. The results are little changed when we add ME and MOM to the regression.

Panel C presents Fama and MacBeth (1973) results using the same set of variables, but now to predict profitability (measured by Gross Profit Margin, GP). IK has a significantly positive impact which is counter to what we would expect for an individual firm purely based on time-series information. However, when we add BM (and subsequently also ME and MOM) in the regression to adjust for cross-sectional differences in characteristics, IK has a negative coefficient as expected. The effect of BM is significantly negative confirming the result of Fama and French (2006) that growth firms are more profitable. The fact that high-investment firms tend to be (the more profitable) growth firms is responsible for the unexpected sign in the univariate regression.

Portfolio Sorting

While the results for equation (13) are consistent with Propositions 1 and 2(a), even the signs of the cross-product terms are not sufficient to confirm Proposition 2(b) via the Fama-MacBeth regressions. However, it is possible to use portfolio sorting to generate specific evidence regarding Proposition 2(b). Table 3 presents the extreme portfolios from a three-way sorting designed to examine the book-to-market effect for those firms with profitability predicted to be below the appreciation rate of intangible assets. This is uncommon, because of the submartingale property of risk-adjusted tangible asset returns and the supermartingale property of risk-adjusted intangible asset returns, so we choose firms with the lowest GP values as this predicts a *low return on tangible assets* (average investment return); and firms with the highest growth of research and development expenses per unit of capital (RDG) as a predictor of a *high appreciation rate of intangible assets.* While usage of R&D as a sorting variable eliminates a considerable fraction of the data (only 20% of COMPUSTAT firm-year observations have R&D measures), we attempt to identify the specific and presumably small subset of firms anticipated to have a change in intangible assets larger than the tangible return and it seems essential to search for them within the set of firms with high R&D growth: high RDG firms have larger intangible asset growth/returns while their tangible asset returns measured by *gross profits* are not affected (which, as Novy-Marx, 2013, stresses, presents an advantage of sorting by gross profits).

In Panel A we first replicate the basic results of Novy-Marx (2013) which illustrate both a value effect and a profitability effect in one-dimensional sorting on BM and GP. For our full sample of around 402,000 firm-month observations the annualized value premium for top and

bottom quintiles is 4.6% and the gross profitability characteristic premium is 3.2%. Panel B replicates the Novy-Marx results for two-dimensional sorting by BM and GP quintiles. For the high and low quintiles of BM ratios the GP premium equals 4.2% and 6.0%, respectively, exceeding the single-sorting premium of 3.2%. Thus, keeping constant BM, the GP effect is stronger. The annualized premium between the highest BM and GP portfolio and the lowest BM and GP portfolio is 8.6%. The qualitative results of Novy-Marx thus continue to hold in our reduced sample containing only firms with positive R&D expenditures, but they are quantitatively a bit smaller.

To test Proposition 2(b), indicating conditions under which the value effect may reverse, we use a three-dimensional sort. To identify firms with different expected growth in future profitability we first sort firms by RDG – the growth of each firm's R&D expenses, normalized by lagged total assets to reflect the overall importance of R&D for this firm. We put the 30% firms with the lowest RDG in portfolio L and the 30% firms with the highest RDG in portfolio H. The second sorting criterion addresses the profitability, and we again put the firms with the 30% lowest GP in portfolios L and those with the 30% highest GP in portfolios H. The prediction is that the returns of higher RDG and lower GP firms exhibit a smaller value effect, and that the value effect is actually reversed for the highest RDG and lowest GP firms.

Panel C shows that, indeed, the value effect (return difference between highest 30%, H, and lowest 30%, L, book-to-market ratio firms) reverses for firms, with above the 70th percentile (marked by H) in RDG, and below the 30th percentile (marked by L) in GP: portfolios 5 and 6. The premium in this case has an annualized value of -2.4%. This return is not by itself significant. However, the value effect has the standard positive sign in all other cases, with an average

annualized return of 5.5%. The difference between the two is -7.9% which is statistically

significant at the 5% level.[9]

The results here must be interpreted with caution. In restricting the sample to include

only firms with positive R&D expenditures, our sample size was reduced to approximately 20%

of the firms publicly traded in US markets, and the reversal occurs in portfolios of fewer than 10

percent of these firms (only 9% of the firms with positive R&D are in the low GP/high RDG

portfolio). The proxy of RDG for expected intangible-asset growth is imperfect and, even in this

sample, unable to capture those anticipated fluctuations in future profitability caused by factors

such as market power changes when they are not related to R&D.

Nevertheless, even with the relatively small sample and imperfect proxy for identifying

cases in which intangible asset returns may exceed tangible asset returns, our approach identified

a statistically significant difference in value premiums between firms anticipated to have high

and those anticipated to have low tangible returns relative to intangible returns. Identifying cases

in which value premiums are theoretically predicted to be different seems to be a useful new

direction for discriminating between theoretical approaches, of which an increasing number

appears to explain well-known stylized facts. Our paper takes a small step in this direction.

---

[9] In spite of our attempts to control for differences in RDG, the average values for RDG in the portfolios with high
RDG and low GP are almost twice as high for the low BM (portfolio 5) as for the high BM (portfolio 6) portfolios.
(See Table 3, Panel C). In principle, this difference could account for the value premium reversal because Chan,
Lakonishok, and Sougiannis (2001) show that firms with higher R&D have higher returns. To address this issue we
remove firms with the highest RDG from Portfolio 5 and firms with the lowest RDG from Portfolio 6 until the
difference in RDG between the two portfolios is reversed. This occurs after removing 30% of the firms from these
portfolios. Then we examine the return difference for the remaining firms in Portfolios 5 and 6. We find that the
value premium is still reversed at even a marginally higher level. The result is available from the authors.

6. Conclusion

In traditional estimation of required returns, factor sensitivities are obtained from time series of returns and factor realizations. Empirically motivated conditioning approaches aside, the tradition of estimating factor sensitivities from simple time series persists in spite of the theoretical contributions of Cox, Ingersoll, and Ross (1985) and Berk, Green, and Naik (1999) that provide a framework for firm choice of risk sensitivities. Given the prices of risk as determined economy-wide and given market and productivity conditions at the firm level, individual firms choose their production and investment activities which determine sensitivities to risk factors. The resulting sensitivities depend on firm-level characteristics implied by the market and productivity constraints faced by the firm, and can be estimated structurally.

The investment-based approach of Cochrane (1991, 1996) provides a specific market and productivity environment for firm choices that has been limited to constant returns to scale technologies. But the limitation is mostly for the sake of tractability and we avoid it here with the purpose of focusing on endogenous fluctuations in profitability. We find that higher investment choice is associated with a lower marginal and average product of capital which decreases the sensitivity of a firm to the risk of productivity shocks. As a result, risk sensitivities and required stock returns decrease as do investment returns and standard profitability measures.

At the same time, book-to-market ratios have a measurable impact on required returns that is separate from profitability. For given expected profitability of book assets (the tangible capital stock) higher book-to-market ratios imply more risk sensitivity: given the mean reversion of productivity, current shocks to capital productivity have more impact when a large fraction of shareholder equity is tied up in tangible assets. An exception is for uncommon cases in which

33

expected returns on tangible capital are smaller than expected appreciation rates of intangible assets. In these cases we expect an inverted value effect because higher book-to-market ratios now imply less overall risk since the importance of intangible assets is reduced relative to tangible assets and, quantitatively in these uncommon cases, the effect of lower exposure to the risk of reduction of intangible asset values dominates the effect of higher exposure to the risk of tangible assets.

The theoretical results support the finding by Novy-Marx (2013) of dual dimensions to value – profitability levels and book-to-market levels – both raising required returns. Our view is that higher profitability relates to a higher average product of capital making firms more sensitive to current productivity shocks. A higher book-to-market ratio further increases the sensitivity to productivity shocks as this implies less weight on the intangible asset component which has relatively low loadings on the current productivity shocks. Empirically, we confirm that higher book-to-market value is associated with higher required returns, the traditional value effect. But we also provide an indication that the value effect is inverted in the predicted cases – for firms with the (30 percent) lowest expected profitability and (30 percent) highest expected intangible asset appreciation (9 percent of the firms in our sample).

Our results hinge on an intangible asset perspective that parts ways with that of the growth options literature (Berk et al., 1999) and the investment-based asset pricing literature (Cochrane, 1996). The view of this literature is that book-to-market ratios deviate from one because the ability of firms to profitably expand their future activities is not incorporated in book values, yet is priced in the market. But, in our model, firms can expand to their desired size without friction so intangible asset values are not related to expansion options. They arise as the

present value of future residual income stemming from decreasing returns to scale and/or market power (in turn more deeply related to such items as the earnings power of capitalized previous research and development expenses, although we do not formally model this aspect). These intangible asset components add to profitability, but because of the documented mean reversion of profitability (Fama and French, 2000 and 2006), are less sensitive than tangible assets to current productivity shocks.

Discarding the constant returns to scale (or the infinite demand elasticity) formulation in the investment-based framework is vital for generating the positive co-movement between a firm's expected stock returns and profitability in our model. Firms with stronger decreasing returns have higher profit margins resulting from increased leverage of the marginal investment return, but the leverage of the marginal investment return also implies proportionately higher risk, and therefore higher expected stock returns. In contrast, Carlson et al. (2004) and subsequent literature rely on operating leverage to generate fluctuations in risk and return. Operating leverage and risk increase with the fixed costs of capital in place, but the profitability decreases with this fixed cost, generating a counterfactual negative co-movement between expected return and profitability as pointed out also by Kogan and Papanikolaou (2013) and Novy-Marx (2013).

Kogan and Papanikolaou (2013) avoid adjustment costs, thus forgoing endogenous fluctuations in the full price of capital investment, but rely instead on an additional productivity shock that generates exogenous stochastic shocks to the price of capital investment. While their model produces the positive co-movement between a firm's expected stock returns and

profitability, it does so at the expense of a counterfactual positive correlation between book-to-market ratios and profitability.

Despite avoiding modeling frictions or multiple risk sources, our simple model neatly describes the basic facts and interactions between stock returns, book-to-market ratios, and profitability. Profitability is persistent and mean reverting, and is negatively related to book-to-market ratios. Stock returns depend positively on profitability and on book-to-market ratios, such that sorting by one such attribute increases the profitability of the other. But the model also predicts, and confirms empirically, high-profitability-growth cases in which the direction of the effect of book-to-market ratios on stock returns is reversed. The latter is a new notion, also from a methodological perspective. As an increasing number of theoretical models is able to match basic historical facts regarding stock returns, investment, profitability, and value, an additional way of evaluating the cogency of each approach is to identify when these models predict deviations from basic historical tendencies and to confront these predictions.

Appendix

A. The optimal capital stock and investment returns

The solution for optimal investment from equation (6) requires further information about the value function. We substitute equations (3) and (2) into the Bellman equation (1)

$$V(\theta_t, K_t) = \theta_t K_t^{\alpha} + (1-\delta)K_t - K_{t+1}^* + E_t[m_{t+1} V(\theta_{t+1}, K_{t+1}^*)], \tag{A1}$$

where K_{t+1}^* is the capital stock implied by the optimal investment policy. The marginal value of capital, $V_K(\theta_{t+1}, K_{t+1})$ may be obtained straightforwardly by updating the value function by one period and using the envelope theorem:

$$V_K(\theta_{t+1}, K_{t+1}) = \alpha \theta_{t+1} K_{t+1}^{\alpha-1} + (1-\delta). \tag{A2}$$

Thus, $K_{t+1}^{\alpha-1} = [1 - (1-\delta)E_t(m_{t+1})]/\alpha E_t(m_{t+1}\theta_{t+1})$ from equations (6) and (A1). The functional forms in equations (4) and (5) imply straightforwardly that

$$E_t(m_{t+1}\theta_{t+1}) = (z + \rho\theta_t)/(1+r), \quad z \equiv (1-\rho)\bar{\theta} - h\sigma. \tag{A3}$$

Equation (7) follows:

$$K_{t+1}^* = [A(z + \rho\theta_t)]^{\frac{1}{1-\alpha}}, \quad A \equiv \alpha/(r+\delta). \tag{A4}$$

From (A2), (A3), and (A4), using the definition $1 + r_{t+1}^I \equiv V_K(\theta_{t+1}, K_{t+1})$ we obtain equation (8):

$$r_{t+1}^I - r = \left(\frac{\sigma(h + e_{t+1}) + \eta_{t+1}}{z + \rho\theta_t} \right)(r + \delta). \tag{A5}$$

B. Components of stock returns

From (A4) it is clear that K_{t+1}^* is a function of θ_t and not K_t. Hence, we may express the disembodied component of equity value, the intangible assets, as:

$$G(\theta_t) = P_t - K_{t+1}^*, \quad P_t = E_t[m_{t+1} V(\theta_{t+1}, K_{t+1}^*)]. \tag{A6}$$

We derive explicitly the relation between investment returns and stock returns. From (A1), (A2), and (A6), and using the definition $\varsigma = (1-\alpha)/\alpha$:

$$V(\theta_{t+1}, K_{t+1}) = K_{t+1}\left[(1+\varsigma)V_K(\theta_{t+1}, K_{t+1}) - \varsigma(1-\delta)\right] + G(\theta_{t+1}). \tag{A7}$$

To obtain equation (11), the expression for stock returns, divide by the stock price on both sides and then subtract one on both sides (net returns are in lower case: $r_t^X \equiv R_t^X - 1$):

$$r_{t+1}^S = \pi_{t+1} b(\theta_t) + g(\theta_{t+1}, \theta_t)[1 - b(\theta_t)], \quad \pi_t \equiv r_t^I + \varsigma(r_t^I + \delta). \tag{A8}$$

Here $b(\theta_t) = K_{t+1}/P_t$ is the book-to-market ratio and $g(\theta_{t+1}, \theta_t) = [G(\theta_{t+1}) - G(\theta_t)]/G(\theta_t)$ is the rate of increase in the firm's intangible assets.

We define $\pi_t = r_t^I + \varsigma(r_t^I + \delta)$ representing the *average return on investment* which equals the marginal return on investment r_t^I (given in A5) plus a markup due to decreasing returns to scale. Gross average return also equals: $\pi_t + \delta = (1+\varsigma)(r_t^I + \delta) = V_k(\theta_t, K_t) - (1-\delta)$. From (A2) we have $V_K(\theta_t, K_t) - (1-\delta) = \alpha\theta_t K_t^{\alpha-1}$. Then from $Y_t = \theta_t K_t^{\alpha}$ and equation (8) we obtain $Y_t / K_t = \theta_t K_t^{\alpha-1}$. It follows that the gross average return on investment also equals the *gross profit margin*, $\pi_t + \delta = Y_t / K_t$.

To prove that the book-to-market ratio is always less than one in our model with decreasing returns to scale, we need to show that $G(\theta_t) = P_t - K_{t+1}^* > 0$ (which implies $K_{t+1}^* / P_t < 1$). Multiply (A7) by m_{t+1} and take expectations. This yields from $\varsigma = (1-\alpha)/\alpha$ and equation (6) that $E_t[m_{t+1} V(\theta_{t+1}, K_{t+1})] = [1 + \varsigma(r+\delta)/(1+r)] K_{t+1} + E_t[m_{t+1} G(\theta_{t+1})]$. From (A4) and (A6) then

$$G(\theta_t) = \left(\frac{\varsigma(r+\delta)}{1+r}\right) K_{t+1} + E_t[m_{t+1} G(\theta_{t+1})]. \tag{A9}$$

Equation (A9) entails by induction, for all t, that $G(\theta_t) > 0$ since $K_{t+1} > 0$ and $\varsigma > 0$.

Intangible asset values adjusted for risk and time value follow a strict *supermartingale* process: since $E_t(m_{t+1} G_{t+1}) < G_t$ for $\varsigma > 0$ from (A9), we have $E_t[m_{t+1}(1+g_{t+1})] < 1$. Similarly, tangible asset values adjusted for risk and time value follow a strict *submartingale* process: because investment returns are a martingale $E_t[m_{t+1}(1+r_{t+1}^I)] = 1$ it follows from $\pi_{t+1} = r_{t+1}^I + \varsigma(r_{t+1}^I + \delta)$ that $E_t[m_{t+1}(1+\pi_{t+1})] > 1$ for $\varsigma > 0$.

Since $E_t(m_{t+1} G_{t+1}) < G_t$, the *adjusted* value of intangible assets must decrease over time. However, the expected growth rate of intangible assets, $E_t(g_{t+1}) = [E_t(G_{t+1}) - G_t]/G_t$, may be expected to either be positive or negative depending on whether the interest cost and the systematic risk compensation together exceed or fall short of the "supermartingale deficit", $[\varsigma(r+\delta)/(1+r)] K_{t+1}$. Note, however, that the unconditional average growth rate of intangible assets must be zero due to the stationarity of the productivity shock process: $E(g_{t+1}) = 0$.

From $E_t[m_{t+1}(1+r_{t+1}^I)] = 1$ it follows that $E_t(r_{t+1}^I) = r + RP_t^I$ (with RP^I the risk premium for the investment return). So $E_t(\pi_{t+1}) - r = RP_t^I + \varsigma(r+\delta+RP_t^I) = RP_t^I + C_t^I$, with $C_t^I > 0$. Further, $E_t[m_{t+1}(1+g(\theta_{t+1},\theta_t))] < 1$ so that $E_t g(\theta_{t+1},\theta_t) = r + RP_t^G - C_t^G$, with $C_t^G > 0$ from the strict supermartingale nature of the intangible assets. Hence, the effect on the expected excess stock return $E_t(r_{t+1}^S) - r$ of an increase in the book-to-market ratio b_t, taking expectations in equation (15), equals $E_t[\pi_{t+1} - g(\theta_{t+1},\theta_t)] = C_t^I + C_t^G + RP_t^I - RP_t^G$ which is typically positive unless the risk premium on intangible assets, RP_t^G, and thus $E_t g(\theta_{t+1},\theta_t)$, is unusually large.

C. Explicit solutions for firm-level expected stock returns over time

Assume normality and set

$$\alpha = 1/2,$$ (A10)

Proceed by employing the method of undetermined coefficients given a quadratic assumption for $G(\theta_t)$ in equation (A9). It is straightforward, but tedious, from equations (A9), (4), and (5) to confirm the quadratic solution for the following specific parameter values:

$$G(\theta_t) = B(c_2\theta_t^2 + c_1\theta_t + c_0),$$ (A11)

with $\quad c_2 = \rho^2, \quad c_1 = \dfrac{2\rho z(1+r)}{1+r-\rho}, \quad c_0 = \dfrac{z^2(1+r+\rho)(1+r)}{r(1+r-\rho)} + \rho^2[(1-h^2)\sigma^2 + \sigma_\eta^2], \quad$ and

$$B \equiv /4(r+\delta)(1+r-\rho^2).$$

(A1) becomes from equations (A10) and (A11):

$$V(\theta_t, K_t) = \theta_t K_t^{1/2} + (1-\delta)K_t + B(c_2\theta_t^2 + c_1\theta_t + c_0),$$ (A12)

From equations (A4) and (A6) we can also find the stock prices as

$$P(\theta_t) = B[c_2\theta_t^2 + c_1\theta_t + c_0 + (z+\rho\theta_t)^2(1+r-\rho^2)/(r+\delta)].$$ (A13)

The expected return is $E_t r_{t+1}^S - r = h\sigma E_t[V_\theta(\theta_{t+1}, K_{t+1}^*)]/P_t$: From $1 = E_t[m_{t+1}(1+r_{t+1}^i)]$ and the definition of (conditional) covariance, the expected net stock return of the firm is $E_t r_{t+1}^S = r - Cov_t(m_{t+1}/E_t m_{t+1}, r_{t+1}^S)$. Using the definition of return and the stochastic discount factor specification in equation (5) then $E_t r_{t+1}^S - r = h E_t[e_{t+1}V(\theta_{t+1}, K_{t+1}^*)]/P_t$. Given the normality of e_{t+1}, Stein's Lemma implies that $E_t r_{t+1}^S - r = (h\sigma/P_t)E_t[V_\theta(\theta_{t+1}, K_{t+1}^*)]$. Straightforward subsequent derivation provides the closed-form solution for the expected excess stock return as

$$E_t r_{t+1}^S - r = h\sigma\left(\frac{2(1+r)(z+\rho\theta_t)+c_1+2\rho^2 h\sigma}{\rho^2\theta_t^2 + c_1\theta_t + c_0 + (z+\rho\theta_t)^2(1+r-\rho^2)/(r+\delta)}\right).$$ (A14)

References

Abel, A. B., and J. C. Eberly, 2011. How Q and Cash Flow Affect Investment without Frictions: An Analytic Explanation. *Review of Economic Studies* 78, 1179-1200.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang, 2006. The Cross-Section of Volatility and Expected Returns. *Journal of Finance* 61, 259–299.

Balvers, R. J., T. Cosimano and B. McDonald, 1990. Predicting Stock Returns in an Efficient Market. *Journal of Finance*, 1109–1128.

Balvers, R. J., and D. Huang, 2007. Productivity-Based Asset Pricing: Theory and Evidence, *Journal of Financial Economics* 86, 405–445.

Belo, F., S. Bazdresch, and X. Lin, 2013. Labor Hiring, Investment and Stock Return Predictability in the Cross Section, Forthcoming *Journal of Political Economics.*

Berk, J. B., R. C. Green, and V. Naik, 1999. Optimal Investment, Growth Options, and Security Returns. *Journal of Finance* 54, 1553–1607.

Bloom, N., 2009, The Impact of Uncertainty Shocks, *Econometrica* 77, 623–685.

Boldrin, M., L. J. Christiano, and J. D. Fisher, 2001. Habit Persistence, Asset Returns, and the Business Cycle, *American Economic Review* 91, 149-166.

Booth, L., B. Chang, W. Hejazi, and P. Shum, 2008. The Influence of Productivity on Asset Pricing, Working Paper University of Toronto.

Brock, W. A., 1982. Asset Prices in an Exchange Economy. in John J. McCall, ed.: *The Economics of Information and Uncertainty*, University of Chicago Press, Chicago.

Caballero, R. J., 1999. Aggregate Investment, in John B. Taylor and Michael Woodford, eds., *Handbook of Macroeconomics* (New York: Elsevier), 816–862.

Carlson, M., A. Fisher, and R. Giammarino, 2004. Corporate Investment and Asset Pricing Dynamics: Implications for the Cross-Section of Returns. *Journal of Finance* 59, 2577–2603.

Chan, L. K. C., J. Lakonishok, and T. Sougiannis, 2001. The Stock Market Valuation of Research and Development Expenditures. *Journal of Finance* 46, 209-237.

Chen, L., and L. Zhang, 2011, Do Time-Varying Risk Premiums Explain Labor Market Performance?. *Journal of Financial Economics* 99, 385–399.

Cochrane, J. H., 1991. Production-Based Asset Pricing and the Link Between Stock Returns and Economic Fluctuations, *Journal of Finance* 46, 209–237.

Cochrane, J. H., 1996. A Cross-Sectional Test of an Investment-Based Asset Pricing Model. *Journal of Political Economy*, 104, 572–621.

Cooper I., 2006. Asset Pricing Implications of Nonconvex Adjustment Costs and Irreversibility of Investment. *Journal of Finance* 61, 139–170.

Cooper I., B. Gerald, and G. Wu, 2005. Investment Irreversibility, Real Activity and the Value Premium. Working paper.

Cox, J.C., Ingersoll Jr, J.E., Ross, S.A., 1985. An Intertemporal General Equilibrium Model of Asset Prices. *Econometrica* 53, 363–384.

Fama, E. F., and K. R. French. 1995. Size and Book-to-Market Factors in Earnings and Returns. *Journal of Finance* 50, 131-155

Fama, E. F., and K. R. French. 1996. Multifactor Explanations of Asset Pricing Anomalies. *Journal of Finance* 11, 55–84.

Fama, E. F., and K. R. French. 2000. Forecasting Profitability and Earnings. *Journal of Business* 73, 161-175.

Fama, E. F., and K. R. French. 2006. Profitability, Investment, and Average Returns. *Journal of Financial Economics* 82, 491-518.

Fama, E. F., and K. R. French. 2008. Dissecting Anomalies. *Journal of Finance* 63, 1653–1678.

Fama, E.F., and J. MacBeth. 1973. Risk, Return, and Equilibrium: Empirical Tests, *Journal of Political Economy* 91, 607–636.

Garlappi, L., and Z. Song. 2013a. Can Investment Shocks Explain Value Premium and Momentum Profits? Working Paper, University of British Columbia and Cheung Kong GSB.

Garlappi, L., and Z. Song. 2013b. Market Power and Capital Flexibility: A New Perspective on the Pricing of Technology Shocks. Working Paper, University of British Columbia and Cheung Kong GSB.

Gomes, J. F., L. Kogan, and L. Zhang, 2003. Equilibrium Cross-Section of Returns, *Journal of Political Economy* 111, 693–732.

Hall, R. E., 2004. Measuring Factor Adjustment Costs, *Quarterly Journal of Economics* 119, 899–927.

Hayashi, F., 1982. Tobin's Marginal q and Average q: A Neoclassical Interpretation, *Econometrica* 50, 213–224.

Hou, K., Xue, C., Zhang, L., 2015. Digesting Anomalies: An Investment Approach. *Review of Financial Studies* 28, 650-705.

Hsu, P., 2006. Aggregate Technology Shocks and Market Return Predictability. Working Paper. Graduate School of Business, Columbia University.

Hsu, P., 2009. Technological Innovations and Aggregate Risk Premiums. *Journal of Financial Economics* 94, 264-279.

Hsu, P., and D. Huang, 2010. Technology Prospects and the Cross-Section of Stock Returns. *Journal of Empirical Finance* 17, 39-53.

Kogan, L., and D. Papanikolaou, 2012. Economic Activity of Firms and Asset Prices. *Annual Review of Financial Economics* 4, 361–84.

Kogan, L., and D. Papanikolaou, 2013. Firm Characteristics and Stock Returns: The Role of Investment-Specific Shocks. *Review of Financial Studies* 26, 2718-2759.

Li, D., and L. Zhang, 2010. Does q-Theory with Investment Frictions Explain Anomalies in the Cross-Section of Returns? *Journal of Financial Economics* 98, 297–314.

Li, E., L. Liu, and C. Xue, 2014. Intangible Assets and Cross-Sectional Stock Returns: Evidence from Structural Estimation. Working Paper Cheung Kong GSB, Hong Kong UST, and University of Cincinnati.

Li, E., D. Livdan, and L. Zhang, 2009. Anomalies. *Review of Financial Studies* 22, 4301– 4334.

Lin, X., 2012. Endogenous Technological Progress and the Cross Section of Stock Returns. *Journal of Financial Economics* 103, 411–428.

Lin, X., and L. Zhang, 2013. The Investment Manifesto. *Journal of Monetary Economics* 60, 351–366.

Lioui, A., and P. Poncet, 2008. Production-Based Asset Pricing in a Monetary Economy: Theory and Evidence. Working Paper EDHEC Risk and Asset Management Research Centre.

Liu, L.X., T. M. Whited, L. Zhang, 2009. Investment-Based Expected Stock Returns. *Journal of Political Economy* 117, 1105–1139.

Newey, W. K., and K. D. West, 1987. A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703–708.

Novy-Marx, R., 2011. Operating leverage. *Review of Finance*, 15, 103–134.

Novy-Marx, R., 2013. The Other Side of Value: The Gross Profitability Premium, *Journal of Financial Economics* 108, 1–28.

Papanikolaou, D., 2011. Investment Shocks and Asset Prices. *Journal of Political Economy* 119, 639-685.

Pastor, L., and P. Veronesi, 2003. Stock Valuation and Learning about Profitability. *Journal of Finance* 58, 1749–1789.

Restoy, F. and G. M. Rockinger. 1994. On Stock Market Returns and Returns on Investment. *Journal of Finance*, 543–556.

Tobin, J., 1969, A General Equilibrium Approach to Monetary Theory, *Journal of Money Credit and Banking* 1, 15–29.

Xing, Y., 2008, Interpreting the Value Effect Through the *Q*-theory: An Empirical Investigation, *Review of Financial Studies* 21, 1767–795.

Zhang, L., 2005, "The Value Premium," *Journal of Finance* 60, 67–103.

Table 1. Summary Statistics, 1963/07-2014/12

Panel A reports the mean, median, 10th percentile and 90th percentile values, and standard deviation (Mean, Median, P10, P90, Std) of the Book-to-Market ratio, the Gross Profit Margin, the Investment-to-Capital ratio, and the three-year average of R&D growth, (BM, GP, IK and RDG) across all firms. GP, as in Novy-Marx (2013), is the gross profit margin (COMPUSTAT item GP) normalized by total assets. BM is the book-to-market ratio and its calculation follows Fama and French (2008). IK is defined as the sum of changes in gross property, plant and equipment (COMPUSTAT item PPEGT) plus the change in inventory (INVT) scaled by lagged total assets (AT). RDG is a three-year average of, the change in research and development spending (COMPUSTAT item XRD) normalized by lagged total assets. Panel B reports the average of AR1 coefficient estimates across all firms for BM, GP, IK and RDG. Panel C reports correlations of the previous four variables, and, the stock return (RET), the June market capitalization (ME, as in Fama and French, 2008) and each stock's price momentum in the past 12 month (MOM). As in Jegadeesch and Titman (1993), we skip the most recent month when computing momentum. We delete firms whose sizes are below the NYSE 10% size breakpoints. Accounting variables are winsorized at 1% and 99%.

Panel A: Summary of GP, BM and IK

Variable	Mean	Median	P10	P90	Std
BM	0.418	0.397	0.163	0.729	0.262
GP	0.648	0.488	0.164	1.263	0.663
IK	0.085	0.059	-0.022	0.221	0.170
RDG	0.011	0.002	-0.001	0.030	0.045

Panel B: Average AR1 Coefficient across Firms

Variable	Mean	Median	P10	P90	Std
BM	0.427	0.444	-0.081	0.835	0.636
GP	0.525	0.588	0.023	0.900	0.379
IK	0.131	0.133	-0.309	0.575	0.433
RDG	0.452	0.515	0.000	0.846	0.445

Panel C: Correlations

Corr.	GP	BM	IK	RDG	ME	MOM
RET	0.01	0.01	-0.01	0.02	-0.01	0.07
GP		-0.16	0.07	-0.09	-0.01	0.02
BM			-0.07	-0.13	-0.11	0.08
IK				0.09	-0.03	-0.03
RDG					0.00	0.01
ME						-0.01

Table 2. Predicting Future Excess Returns and Gross Profit Margin, 1963/07-2014/12

Panel A reports results from regressing monthly firm-level excess returns from July of year t to June of year t+1 on beginning-of-year t characteristics following Fama and MacBeth (1973). We report the average coefficients and Newey-West t-values with four lags to adjust for heteroskedasticity and autocorrelation. Control variables include ME, and MOM. ME and BM are in natural logs. Panel B uses the level of BM and includes cross products of BM with GP and RDG. Panel C presents Fama and MacBeth (1973) estimation results from regressing annual next-year gross profit margin on combinations of current IK, BM and MOM. Three (two, one) stars indicate significance at the 1% (5%, 10%) level.

Panel A: Predicting Returns with IK, GP, BM and RDG

		Intercept	IK	GP	BM	RDG	ME	MOM	R-square
1	Coeff.	0.0067**	-0.0088***	0.0084***	0.0028***				0.0233
	t-value	2.34	-3.22	4.10	3.28				
2	Coeff.	0.0097*	-0.0095***	0.0071***	0.0020**		-0.0005		0.0400
	t-value	1.86	-3.55	3.44	2.25		-1.06		
3	Coeff.	0.0075	-0.0083***	0.0071***	0.0025***		-0.0004	0.0042*	0.0538
	t-value	1.54	-3.27	3.67	2.94		-0.88	1.83	
4	Coeff.	0.0058**		0.0090***	0.0032***	0.0106			0.0275
	t-value	2.00		4.43	4.10	0.28			
5	Coeff.	0.0063		0.0078***	0.0029***	0.0031	-0.0003	0.0046**	0.0573
	t-value	1.30		4.03	3.79	0.09	-0.77	1.97	

Panel B: Predicting Returns with GP, BM, RDG, and Cross-Products

		Intercept	BM	RDG	BM·GP	BM·RDG	ME	MOM	R-square
1	Coeff.	0.0049*	-0.0002	0.055	0.0112***	-0.099			0.0269
	t-value	1.96	-0.12	1.15	3.42	-0.98			
2	Coeff.	0.0052	-0.0003	0.0322	0.0102***	-0.0567	-0.0003	0.0049**	0.0577
	t-value	1.10	-0.25	0.69	3.07	-0.61	-0.73	2.09	

Panel C: Predicting Gross Profitability with IK and BM

		Intercept	IK	BM	ME	MOM	R-square
1	Coeff.	0.4085***	0.1112***				0.0051
	t-value	44.47	5.44				
2	Coeff.	0.3624***	-0.0243	-0.0971***			0.1349
	t-value	20.75	-0.61	-5.98			
3	Coeff.	0.4132***	-0.0156	-0.1045***	-0.0091***	0.0640***	0.1606
	t-value	15.19	-0.41	-6.01	-3.00	4.51	

Table 3. Reversal of the Book-to-Market Effect, 1963/07-2014/12

In Panel A, at June of each year t, we allocate all firms into five quintiles based on the BM or GP values available at the end of year t-1, and we track monthly excess returns of these firms from July of year t to June of year t+1. Portfolio excess returns are the value-weighted averages of excess returns of firms in each portfolio. Portfolio 1 includes firms with BM or GP values below the 20th percentile, and Portfolio 5 includes firms with GP or BM values above the 80th percentile, etc.

Panel B forms 25 portfolios by the intersections of BM and GP quintile portfolios. We report the bottom and top BM quintile portfolios.

In Panel C, we perform a three-dimensional sequential sorting by RDG, GP, and BM. Portfolio H represents the portfolio of the stocks with the highest 70% RDG, GP or BM firms; Portfolio L represents the portfolio of the stocks with the lowest 30% RDG, IK, or BM firms. We track excess returns from July of year t to June of year t+1, and portfolio returns (RET) are value-weighted. We also report mean values of ME, BM, RDG, IK, and GP for each portfolio. The bottom part of Panel C shows the difference between long and short portfolios.

Panel A: One-Dimensional Sorting Based on BM and GP

		RET	ME	BM	RDG	IK	GP
BM	L	0.0044	49158	0.1964	0.0142	0.1158	0.5921
	2	0.0059	43636	0.3882	0.0067	0.0904	0.4425
	3	0.0054	44932	0.5804	0.0037	0.0818	0.3548
	4	0.0062	33854	0.8061	0.0029	0.0711	0.2916
	H	0.0082	19658	1.2596	0.0028	0.0729	0.2473
GP	L	0.0044	52382	0.8133	0.0031	0.0749	0.1742
	2	0.0067	38169	0.6617	0.0037	0.0867	0.2949
	3	0.0050	46708	0.4938	0.0075	0.0829	0.4002
	4	0.0048	44390	0.3603	0.0086	0.0881	0.5130
	H	0.0071	43440	0.2534	0.0137	0.1165	0.7305

Panel B: Two-Dimensional Sorting Based on BM and GP

BM	GP	RET	ME	BM	RDG	IK	GP
L	L	0.0014	20199	0.2039	0.0207	0.1282	0.1809
L	2	0.0040	42535	0.2170	0.0135	0.1197	0.4018
L	3	0.0035	43041	0.2014	0.0133	0.1171	0.5364
L	4	0.0052	53261	0.1882	0.0158	0.1243	0.6761
L	H	0.0064	30664	0.1733	0.0160	0.1187	0.8600
H	L	0.0051	17047	1.3695	0.0016	0.0744	0.1350
H	2	0.0077	19290	1.2442	0.0030	0.0801	0.2300
H	3	0.0105	14923	1.2175	0.0034	0.0641	0.2967
H	4	0.0112	4760	1.2268	0.0036	0.0637	0.3782
H	H	0.0086	1651	1.2320	0.0036	0.0448	0.5468

Panel C: Three-Dimensional Sorting Based on RDG, GP and BM

Portfolio	RDG	GP	BM	RET	MC	BM	RDG	IK	GP
1	L	L	L	0.0023	14046	0.4447	-0.0042	0.0691	0.1572
2	L	L	H	0.0084	14236	1.5090	-0.0015	0.0456	0.1623
3	L	H	L	0.0085	26946	0.2279	-0.0024	0.0940	0.6399
4	L	H	H	0.0130	1366	0.9451	-0.0030	0.0375	0.5463
5	H	L	L	0.0059	6496	0.2154	0.0420	0.1423	0.1978
6	H	L	H	0.0039	8957	0.9956	0.0229	0.1376	0.2349
7	H	H	L	0.0043	43222	0.1464	0.0287	0.1469	0.7390
8	H	H	H	0.0077	13248	0.5547	0.0221	0.0851	0.6743

Portfolio	(1-2)	(4-3)	(6-5)	(8-7)	(6-5) - [(1-2)+(4-3)+(8-7)]/3
Return	0.0060	0.0045	-0.0020	0.0033	-0.0066
t-value	*2.0961*	*1.7469*	*-0.6105*	*1.1677*	*-2.0543*

R1. Supports Footnote 8: Predicting Intangible Asset Growth, g_{t+1}, which is the ratio of market value minus book value at year t+1 and market value minus book value at year t, minus 1 (conform A6 and the definition below A8). We use the available variables at time t to forecast this growth rate in equation 3.

		Intercept	IK	BM	RDG	GP	ME	MOM	g_t	R^2
					Predicting g_{t+1}					
1	Coeff.	0.0941			6.3131***					0.0028
	t-value	*1.33*			*3.21*					
2	Coeff.	0.1398*							0.0061	0.0008
	t-value	*2.01*							*0.48*	
3	Coeff.	0.2864	-0.152	-0.3319***	1.5365	0.2696*	-0.0106	0.1289*	-0.0007	0.0135
	t-value	*1.55*	*-0.86*	*-4.48*	*1.05*	*1.82*	*-0.64*	*1.83*	*-0.07*	

R2. Supports Footnote 8: EG is the fitted residual from equation (3) in R1 as an alternative proxy for expected intangible asset growth. It is obtained by regressing realized future intangible asset growth on all available variables. We then set EG equal to fitted intangible asset growth and use it as a sorting variable instead of RDG.

Three-Dimensional Sorting Based on EG, GP and BM

Portfolio	EG	GP	BM	RET	MC	BM	RDG	IK	GP
1	L	L	L	0.0046	46831	0.5405	0.0069	0.0786	0.1386
2	L	L	H	0.0060	17317	1.6187	0.0013	0.0621	0.1492
3	L	H	L	0.0031	21501	0.2901	0.0197	0.1361	0.5844
4	L	H	H	0.0072	3073	1.3443	0.0021	0.0377	0.4664
5	H	L	L	0.0070	18494	0.1790	0.0226	0.1433	0.3568
6	H	L	H	0.0063	11977	0.5887	0.0170	0.1381	0.3711
7	H	H	L	0.0064	35561	0.1393	0.0184	0.1316	0.8126
8	H	H	H	0.0099	4452	0.5068	0.0076	0.1019	0.8231

R3. Supports Footnote 9: remove firms with the highest RDG from Portfolio 5 and firms with the lowest RDG from Portfolio 6 until the difference in RDG between the two portfolios is reversed. This occurs after removing 30% of the firms from these portfolios.

Three-Dimensional Sorting Based on RDG, GP and BM

Portfolio	RDG	GP	BM	RET	MC	BM	RDG	IK	GP
1	L	L	L	0.0023	14046	0.4447	-0.0042	0.0691	0.1572
2	L	L	H	0.0084	14236	1.5090	-0.0015	0.0456	0.1623
3	L	H	L	0.0085	26946	0.2279	-0.0024	0.0940	0.6399
4	L	H	H	0.0130	1366	0.9451	-0.0030	0.0375	0.5463
5	H	L	L	0.0043	6703	0.2211	0.0279	0.1370	0.2152
6	H	L	H	0.0018	7180	0.9920	0.0313	0.1385	0.2662
7	H	H	L	0.0043	43222	0.1464	0.0287	0.1469	0.7390
8	H	H	H	0.0077	13248	0.5547	0.0221	0.0851	0.6743

R4. Old formulation Without Utilization: Sorting as in the previous draft but omitting the sort by utilization. Here we use IK to proxy for low expected profitability rather than using GP directly as we do in the current draft.

Three-Dimensional Sorting Based on RDG, IK and BM

Portfolio	RDG	IK	BM	RET	MC	BM	RDG	IK	GP
1	L	L	L	0.0080	15193	0.3989	-0.0036	-0.0428	0.4002
2	L	L	H	0.0104	6556	1.3812	-0.0022	-0.0519	0.2433
3	L	H	L	0.0033	15635	0.2657	-0.0034	0.1867	0.4821
4	L	H	H	0.0100	24086	1.0825	-0.0010	0.1748	0.2449
5	H	L	L	0.0085	46672	0.2247	0.0185	0.0319	0.5917
6	H	L	H	0.0095	5008	0.8990	0.0199	0.0138	0.3819
7	H	H	L	0.0068	31245	0.1466	0.0392	0.2575	0.6189
8	H	H	H	0.0066	11346	0.7333	0.0224	0.2616	0.3735