

NBER WORKING PAPER SERIES

INVESTMENT AND THE CROSS-SECTION OF EQUITY RETURNS

Gian Luca Clementi
Berardino Palazzo

Working Paper 21064
<http://www.nber.org/papers/w21064>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2015, Revised April 2018

We thank the Editors, Kenneth Singleton and Stefan Nagel, one associate editor, and two anonymous referees for suggestions that greatly improved the paper. We are also grateful to Frederico Belo, Martin Boileau, Andrea Buffa, Ilan Cooper, João Gomes, Dirk Hackbart, Xiaoji Lin, Lars Lochstoer, Sydney Ludvigson, and Stijn Van Nieuwerburgh, as well as seminar attendants at NHH Bergen, Bank of Italy, Boston University, Federal Reserve Bank of St. Louis, NYU, University of Minnesota, University of Southern California, Wharton, University of Wisconsin, CAPR Workshop (Oslo), CIREQ Workshop (Montreal), Duke-UNC Asset Pricing Conference, “Firm Dynamics and the Aggregate Economy” at the Becker-Friedman Institute, ENSAI Economic Day (Rennes), EUI Macro Conference (Florence), Midwest Macro Meeting, NY/Philly Workshop on Quantitative Macroeconomics, and meetings of American Finance Association, European Finance Association, Canadian Macro Study Group, and Society for Economic Dynamics for their comments and insights. Marco Casiraghi and Peifan Wu provided excellent research assistance. Neither of us has relevant or material financial interests that relate to the research described in this paper. All remaining errors are our own responsibility. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2015 by Gian Luca Clementi and Berardino Palazzo. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Investment and The Cross-Section of Equity Returns
Gian Luca Clementi and Berardino Palazzo
NBER Working Paper No. 21064
April 2015, Revised April 2018
JEL No. D24,D92,G12

ABSTRACT

We confront the one-factor production-based asset pricing model with the evidence on firm-level investment, to uncover that it produces implications for the dynamics of capital that are seriously at odds with the evidence. The data shows that, upon being hit by adverse profitability shocks, large public firms have ample latitude to divest their least productive assets and downsize. In turn, this reduces the risk faced by their shareholders and the returns that they are likely to demand. It follows that when the frictions to capital adjustment are shaped to respect the evidence on investment, the model-generated cross-sectional dispersion of returns is only a small fraction of what documented in the data. Our conclusions hold true even when either operating or labor leverage are modeled in ways that were shown to be promising in the extant literature.

Gian Luca Clementi
Department of Economics
Stern School of Business
New York University
44 West Fourth Street
New York, NY 10012
and NBER
clem@nyu.edu

Berardino Palazzo
Boston University
bpalazzo@bu.edu

1 Introduction

A noteworthy goal of modern applied economics is to devise internally consistent theories of choice, production and exchange, whose predictions for both quantities and prices are consistent with the evidence. For example, much of the research in macroeconomics and asset pricing over the last twenty-five years has been guided by the objective to conceive a *unified* model of the business cycle, able to generate empirically appealing *time-series* behavior of both aggregate quantities – output, consumption, investment, and employment – and financial assets returns.¹

Coming to the *cross-sectional* evidence on investment and financial asset returns, it is the broadly intended neoclassical model of firm optimization that has become the analytical framework of choice for scholars interested in rationalizing either one or the others. On the one hand, [Cooper and Haltiwanger \(2006\)](#), among others, have shown us how plant-level investment data restricts modeling choices on capital adjustment costs. On the other, a series of papers that started with [Carlson et al. \(2004\)](#), [Zhang \(2005\)](#), and [Cooper \(2006\)](#) have investigated the restrictions on the same choices coming from equity returns. The upshot of these distinct strands of literature is that the same theoretical framework, when appropriately specified, has empirically sensible implications for quantities (investment rate) and prices (equity returns).

Our interest lies in establishing whether the parametric restrictions imposed by investment data are compatible with those deriving from equity returns data. We find that, for a large class of one-factor models, they are not. Loosely speaking, the capital adjustment costs implied by the investment data are too small to justify the observed dispersion in returns.

To judge the relevance of our finding, consider that over the last decade or so the one-factor investment-based model has been the dominating paradigm in the quest to understand the drivers of cross-sectional heterogeneity in returns. For example, [Livdan et al. \(2009\)](#) used it to study the effect of financial constraints on stock returns, [Gomes and Schmid \(2010\)](#) adopted it to investigate the role of financial leverage, [Schmid and Kuehn \(2014\)](#) assessed its ability to rationalize credit spreads, while [Tuzel and Zhang \(2017\)](#) argued that it can rationalize the impact of local factors on asset prices.

We start by documenting investment behavior among publicly traded US firms, i.e. the subset of firms at the attention of asset pricing scholars. This exercise is akin to

¹See [Rouwenhorst \(1995\)](#) for a diagnosis of the inability of the prototypical real business cycle model to generate sensible predictions for asset returns, and [Jermann \(1998\)](#) for an early attempt to write down a model that came to terms with the evidence on both quantity and asset return dynamics.

that conducted by [Cooper and Haltiwanger \(2006\)](#) on manufacturing plants. The data displays a substantial cross-sectional dispersion of the investment rate – 28.5% at the annual frequency, almost twice the unconditional mean. Furthermore, every quarter an average 18.2% of firms record negative gross investment. We take the latter as strong evidence against the assumption of irreversibility.

We then consider a variant of the neoclassical investment model, very close to that entertained by [Zhang \(2005\)](#). Our environment features decreasing returns to scale, mean-reverting idiosyncratic shocks to profitability, and a flexible formulation of capital adjustment costs. We select parameters to match the cross-sectional dispersion and autocorrelation of the investment rate, as well as the fraction of firms undertaking negative investment. Future cash-flows are priced by means of an exogenous stochastic discount factor parameterized to replicate the first two unconditional moments of the risk-free rate and the Sharpe ratio.

In the cross-section, risk and returns are decreasing in productivity. Because of mean-reversion, low-productivity firms derive most of their value from cash-flows with long maturity, which in this framework are riskier. Conversely, the value of high-productivity firms comes mostly from short-run cash flows, which are less risky. This rationalizes the size premium, i.e. the finding that firms with small market capitalization (low productivity, in this environment) elicit higher returns.

Due to decreasing returns to scale, on average value firms (high book-to-market) feature higher idiosyncratic productivity than growth firms (low book-to-market). Therefore, the model produces a value discount. The equilibrium association between productivity and book-to-market is inverted by assuming that firms incur operating costs (invariant to firm size) that are large enough.

When calibrating the operating cost to match the unconditional mean firms' exit rate, we find that indeed the unconditional cross-sectional correlation between productivity and book-to-market turns negative. A value premium obtains.

However, value firms are saddled with a large capital stock, resulting from a recent history of good shocks. The ability to relinquish capital – disciplined by the evidence on investment – makes them less risky than small firms. Growth firms feature high productivity but, thanks to a recent history of bad shocks, a relatively low capital. The ability to quickly grow their operating assets – once again disciplined by the evidence on investment – makes them riskier than large firms.

It follows that the resulting value premium is smaller than the size premium and, in response to our research question, equal to only 1/3 of the value we estimate from the

data. A higher dispersion of equity returns can be obtained by raising the adjustment cost of capital, but at the cost of a counterfactually lower volatility of investment.

This is the tension we focus upon. We show that for the value premium to get close to its empirical counterpart, both the volatility of investment and the fraction of firms undertaking negative investment must be very close to zero. Consistent with these results, we document that for the model in [Zhang \(2005\)](#) to generate a non-trivial value premium, the implied investment process must be counterfactual. In particular, the unconditional standard deviation of investment is one order of magnitude lower than in the data.

[Zhang \(2005\)](#), [Carlson et al. \(2004\)](#) and [Obreja \(2013\)](#) among others, emphasize the amplification role of operating leverage. In our baseline model, the magnitude of the operating cost is disciplined by matching the exit rate. Raising the cost to levels associated with counterfactually higher exit rates has two countervailing effects on the dispersion of returns: On the one hand, it raises the variation among firms that stay in the sample. On the other hand, the enhanced exit selection compresses the distribution. In our environment, the two effects cancel each other out.

We then go on to assume that operating costs rise with firm size. Using COMPUSTAT data, we estimate such costs to be increasing and concave in installed capital, with an elasticity of about 0.65. Consistent with [Carlson et al. \(2004\)](#), we find that – everything else equal – scaling up operating costs by a multiplicative factor leads to a wider variation in returns. However, such modification also raises the dispersion of the marginal cost of investment, leading to a counterfactually higher variation in investment rates. Recalibrating the model to hit the investment rate targets has the effect of reducing the variation in returns to the starting level.

A direct comparison with the results in [Carlson et al. \(2004\)](#) is not possible, since they estimate the adjustment costs parameter by matching equity returns and do not report the implications for investment. Key assumptions of their framework, however, indicate that their investment process is at odds with the evidence. In their first model, they assume complete investment irreversibility. In their quantitative framework, investment expenditures are completely sunk.

In recent years, [Marfè \(2017\)](#), [Donangelo et al. \(2018\)](#), and [Favilukis and Lin \(2016\)](#) have argued that labor leverage may play an important role in shaping the cross-sectional variation of equity returns. It is not known, however, what is the role of labor leverage under our requirement that the model be consistent with the evidence on investment. In the first two papers capital is constant by assumption. The latter allows for variation in capital, but unfortunately does not report the model’s implications for firm-level

investment.

To find out, we relax the Cobb–Douglas assumption to allow capital and labor to be complement in production. This leads the labor share to covary negatively both with aggregate productivity in the time-series and with idiosyncratic productivity in the cross-section. Notice that these are also the necessary conditions for labor leverage to matter for equity returns.

Wages and aggregate productivity follow a joint VAR(1) process, whose parameter we estimate from the data. For an elasticity of substitution equal to 0.6, the model generates a very limited cross-sectional variation of the labor share and a counterfactually low value premium. Once more, this result suggests that ignoring the adjustment of capital is not without loss of generality.

Lowering the elasticity of substitution to 0.4 leads to a substantially higher value premium, but only because the aggregate labor share is 13 percentage points higher than in the data. Reducing the capital intensity to reign in the labor share also reduces the value premium to the same level obtained for an elasticity of 0.6.

The path-breaking contributions by [Berk et al. \(1999\)](#) and [Gomes et al. \(2003\)](#) first highlighted that firm choice can be reconciled with a range of cross-sectional regularities for risk and returns. A long series of quantitative studies inspired by [Zhang \(2005\)](#) and [Carlson et al. \(2004\)](#) singled out assumptions on functional forms and parameter values under which the one-factor investment-based models delivers an empirically sensible cross-sectional variation of asset returns.

Our main contribution is to show that this happens at the expense of the model’s ability to rationalize investment behavior. The investment-based one-factor model does not explain investment, because it requires heroic level of capital adjustment costs. Contrary to what required by this framework, public firms have a relatively easy time disposing of assets when hit by persistent shocks to their profitability. This substantially reduces their risk and the financial returns they elicit, along the lines suggested by [Guthrie \(2011\)](#).

Fiddling further with this framework does not look promising. The reason is that, as hinted by our analysis of the roles played by operating and labor leverage, any friction introduced with the goal of limiting firms’ chances to adjust to shocks is likely to have empirically implausible consequences for investment behavior.

In general, our conclusions do not extend to two-factor models such as those considered by [Papanikolaou \(2011\)](#), [Kogan and Papanikolaou \(2014\)](#), [Belo et al. \(2014\)](#), and [Garlappi and Song \(2017\)](#).

The remainder of the paper is organized as follows. In Section 2, we present our evidence on investment across public companies. The baseline model and the calibration procedure are introduced in Section 3. The roles of operating and labor leverage are analyzed in Sections 4 and 5, respectively. Section 6 concludes.

2 Evidence on investment at U.S. public companies

Studies of the plant-level investment process such as Doms and Dunne (1998) and Cooper and Haltiwanger (2006) provide the empirical evidence needed to discipline quantitative studies of the role of cross-sectional heterogeneity in macroeconomic models. In an analogous fashion, in this section we carefully describe investment at U.S. public companies, for the purpose of informing modeling choices in the quantitative analysis of production-based asset pricing models.

Casual observation as well as academic studies (see Doms and Dunne (1998)) hint that capital accumulation at public firms is likely to be very different from that emerging from the analysis of a representative sample of manufacturing establishments. Most public firms are very large entities, are often multi-plants, and operate in a variety of industries. While comprehensive, our study emphasizes features, such as the volatility and reversibility of investment, that play a particularly important role in the class of models considered below.

2.1 Description of the data

Our data is from the quarterly Compustat database for the period 1975q1–2016q4. We start in 1975, because quarterly data on capital expenditures is sparse in earlier years. We exclude financial firms (SIC codes 6000 to 6999), utilities (SIC codes 4900 to 4999), and other unclassified firms (SIC codes greater or equal to 9000), as well as firms not incorporated in the United States or not traded on either NYSE, AMEX, or NASDAQ.

We proxy the capital stock with the item *PPENTQ*, defined as the net value of property, plant, and equipment. Net quarterly investment is the difference between two consecutive values of that variable.²

Our convention amounts to assume that accounting depreciation is an accurate proxy for economic depreciation. This is short of ideal. One would rather recover gross investment, and then subtract the best estimate of economic depreciation. We do not follow this route, because it would entail a loss of about 35% of our observations.³ The gross

²We replace missing observations for *PPENTQ* with a linear interpolation if values of that variable in the immediately contiguous quarters are available. Imputed entries account for 0.6% of the total.

³The reason lies in the very large number of missing observations for the variables *PPEGTQ* and *DPACTQ* – the gross value of PPE and the change in accumulated depreciation, respectively.

investment rate is set equal to

$$\frac{PPENTQ_t - PPENTQ_{t-1}}{PPENTQ_{t-1}} + \delta_j,$$

where δ_j is the average depreciation rate of industry j estimated using data from the Bureau of Economic Analysis.⁴

We further exclude (i) companies that have less than 12 quarters of data, so that our observations effectively start in 1978q1, (ii) firm-quarter observations associated with acquisitions larger than 5% of assets in absolute value and those yielding investment rates in the top or bottom 0.5% of the distribution, and (iii) observations with missing values for the investment rate or the book-to-market ratio. Our cleaned dataset consists of 296,218 firm-quarter observations.

2.2 Results

Figure 1 below illustrates the pooled distribution of gross investment rates. The data displays a substantial amount of cross-sectional variation, right-skewness, and a large fraction of negative observations:

- The time-series average of the cross-sectional standard deviation is 9.5%, close to three times the pooled mean investment rate of 3.5%. The average autocorrelation coefficient is 0.26.⁵
- In an average quarter, 16% of firms report a gross investment rate in absolute value lower than 1%. In line with the convention established by [Cooper and Haltiwanger \(2006\)](#), we call such firms *inactive*.
- A further 18% of firms have an investment rate lower than -1%. It turns out that plenty of firms downsize, at all times.

The latter finding corroborates the ample evidence against investment irreversibility. For example, [Eisfeldt and Rampini \(2006\)](#) find that every year the average firm in their dataset sells physical assets for 9.16 million in 1996 dollars, or about 10% of capital expenditures. According to [Eckbo and Thorburn \(2008\)](#), in 2006 alone U.S. corporation announced 3,375

⁴The data consists of the total values of the stock of private fixed assets and its depreciation, across 3-digit SIC industries. If a company is assigned by COMPUSTAT only a 2-digit (1-digit) SIC code, we use the average depreciation rate at the 2-digit (1-digit) level.

⁵The cross-sectional autocorrelation at time t ($t = 1, \dots, T$) is the estimate of α_{1t} in the equation $x_{i,t} = \alpha_{0t} + \alpha_{1t}x_{i,t-1} + \varepsilon_{i,t}$, $i = 1, \dots, N_t$, where $x_{i,t}$ is the investment rate of firm i at time t and N_t is the total number of firms for which we have data for both $x_{i,t}$ and $x_{i,t-1}$. Pooled ordinary least squares deliver a very similar result.

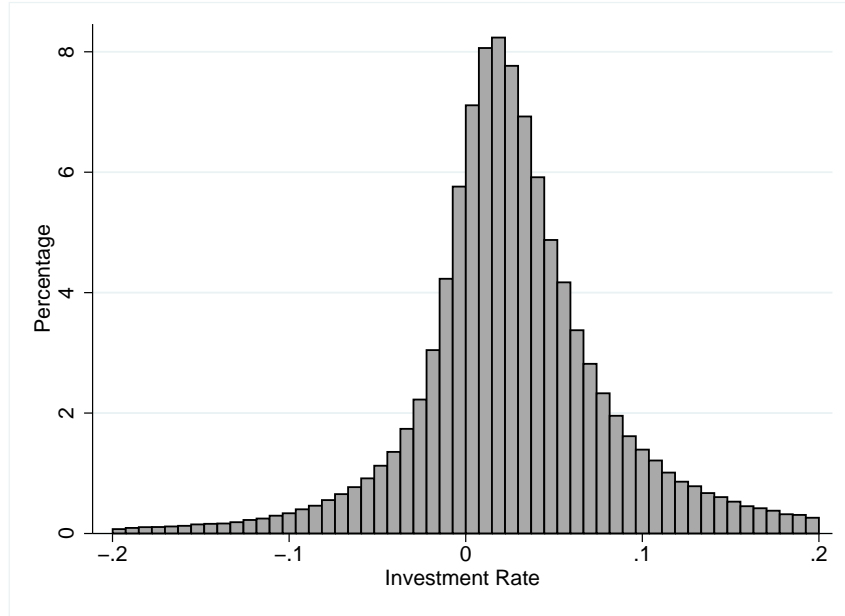


Figure 1: Cross-sectional Distribution of Gross Quarterly Investment Rates.

divestitures – sales of a portion of firm’s assets to a third party – for a total value of \$342 billion.⁶

In order to facilitate the comparison between our results and the available evidence, in the right column of Table 1 we report the annual versions of our summary statistics. The mean investment rate is close to the values reported by Gomes (2001), 0.145, Barnett and Sakellaris (1998), 0.16, and Kogan et al. (2018), 0.175. The same is true for the serial correlation. The three studies report values of the autocorrelation coefficient of 0.239, 0.22, and 0.223, respectively.

Larger are the differences when considering measures of volatility and negative investment. Gomes (2001) and Barnett and Sakellaris (1998) report standard deviations of 0.14 and 0.24, respectively. The fractions of firms reporting negative investment in the two studies are instead 8% and 3%, respectively.⁷ To the extent that our results differ from those reached in the extant literature, they do largely because of the different approach to measuring investment. In Gomes (2001), investment is *CAPX* net of capital retirements (item *PPEVR*). For Barnett and Sakellaris (1998), it is *CAPX* net of the sale of property, plant, and equipment (item *SPPE*). Kogan et al. (2018) measure investment as the percentage change in the gross value of plant, property, and equipment.

The variable *PPEVR* employed by Gomes (2001) is available only through 1997. The

⁶Eckbo and Thorburn (2008) report that the number of transaction was relatively stable between 1980 and 2005, but grew at a fast pace until the start of the Great Recession.

⁷Kogan et al. (2018) omit to report both values.

Table 1: Gross Investment Rate: Summary statistics.

	Quarterly	Annual
Mean investment rate	0.035	0.156
Standard deviation	0.095	0.285
Autocorrelation	0.262	0.276
Negative investment	0.182	0.190
Inaction rate	0.165	0.039
Positive spikes	0.038	0.295
Negative spikes	0.012	0.045
Observations	296,218	65,724

reason for not limiting ourselves to netting *SPPE* from *CAPX*, as in [Barnett and Sakellaris \(1998\)](#), is that *SPPE* understates the value of divested assets, as it only captures alienations that originate cash inflows. However, firms often exchange assets for equity or debt. See [Slovin et al. \(2005\)](#). A further issue is that *SPPE* only imperfectly accounts for capital retirements.⁸

To gain some insight into the shortcoming caused by measuring investment as the difference between *CAPX* and *SPPE*, consider the case of ADC Telecommunications, a Minneapolis-based company that operated independently until 2010. At the end of fiscal 2002, its reported value for net property plant and equipment was about \$408 million short of its 2001 value. The 2002 10-K form⁹ reveals that such decline resulted from a substantial downsizing that included (i) the sale or closure of product lines, (ii) the disposition of equipment, and (iii) a reduction in the workforce from 12,042 to 7,600. Unfortunately, however, *CAPX* net of *SPPE* was only \$25.6 million!

The extent of winsorization also matters. To be consistent with the algorithm used in constructing the quarterly investment rate, when computing the moments in the right column of Table 1 we rule out observations of the annual investment rate in the top and bottom 0.5% of the distribution. When trimming the distribution at the top and bottom 1%, our annual volatility declines to 24%, the same value reported by [Barnett and Sakellaris \(1998\)](#).

In summary, our empirical analysis indicates that for US public firms, investment displays substantial volatility and no sign of irreversibility. In fact, in every quarter a

⁸See [Bayraktar \(2002\)](#) for a careful discussion.

⁹Available on the SEC's website: <https://www.sec.gov/edgar/searchedgar/companysearch.html>

large fraction of firms reduce their deployment of plant, property, and equipment. In the remainder of the paper, we will require that the investment process implied by the asset-pricing models under consideration conforms closely with this evidence.

3 The Baseline Model

Time is discrete and is indexed by $t = 1, 2, \dots$. The horizon is infinite. At every time t , a positive mass of firms produce an homogenous good by means of the production function $y_t = e^{z_t + s_t} [k_t^\alpha l_t^{1-\alpha}]^\nu$, $\alpha, \nu \in (0, 1)$. Here k_t denotes physical capital and l_t measures employment.

The supply of labor is infinitely elastic at the wage rate w_t . Capital depreciates at the random rate δ_t , distributed over the open interval $(0, 1)$ according to the time-invariant c.d.f. Q . The variables z_t and s_t are aggregate and idiosyncratic random disturbances, respectively. They are orthogonal to each other.

The common component of productivity z_t is driven by the stochastic process

$$z_{t+1} = \rho_z z_t + \sigma_z \varepsilon_{z,t+1},$$

where $\rho_z \in (0, 1)$, $\sigma_z > 0$, and $\varepsilon_{z,t} \sim N(0, 1)$ for all $t \geq 0$. The conditional distribution of z_{t+1} will be denoted as $J(z_{t+1}|z_t)$.

The idiosyncratic component s_t evolves according to

$$s_{t+1} = \rho_s s_t + \sigma_s \varepsilon_{s,t+1},$$

where $\rho_s \in (0, 1)$, $\sigma_s > 0$, and $\varepsilon_{s,t} \sim N(0, 1)$ for all $t \geq 0$. The conditional distribution of s_{t+1} will be denoted as $H(s_{t+1}|s_t)$.

When gross investment is limited to replacing depreciated capital, i.e. $x_t \in [0, \delta_t k_t]$, we refer to it as maintenance investment. We assume that its installation involves no further costs beyond the acquisition of the new equipment at the unit price of 1. Alternatively, when either $x_t > \delta_t k_t$ or $x_t < 0$, firms incur an adjustment cost

$$g(x, k_t) \equiv \phi_1 \left(\frac{x_t}{k_t} \right)^2 k_t, \quad \phi_1 \geq 0.$$

We further assume that the purchase of capital goods in excess of maintenance investment, i.e. $x_t > \delta_t k_t$, occurs at the unit price $1 + \phi_0$, with $\phi_0 > 0$. Conversely, the sale of capital equipment (the scenario for $x_t < 0$) yields a unit revenue $1 - \phi_0$. The net payment induced by a given level of gross investment is pictured in Figure 2.

By assuming – along the lines of [Abel and Eberly \(1994\)](#) and [Cooper and Haltiwanger \(2006\)](#) – that the selling price of capital is lower than the buying price, we allow our model

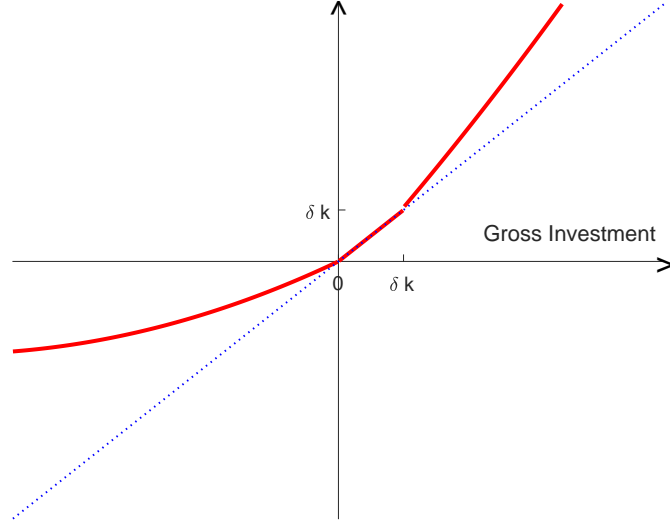


Figure 2: Total expenditures implied by different levels of gross investment.

the flexibility to generate an empirically plausible value for the fraction of observations with negative gross investment, a key indicator of investment reversibility.

In [Abel and Eberly \(1994\)](#), the gap between buying and selling price of capital induces a discontinuity in the marginal cost of adjusting capital at $x = 0$. Compared to a scenario with only convex adjustment costs, the simulation of their model generates greater inaction, i.e. more observations with identically zero gross investment, and correspondingly less observations with either negative or positive investment.

We depart slightly from [Abel and Eberly \(1994\)](#), because in our data inaction takes the shape of a large fraction of observation with small, but strictly positive gross investment. By assuming no adjustment cost for maintenance investment, our model can reproduce this evidence.

Firms evaluate future cash flows by means of the discount factor $M(z_t, z_{t+1})$. Similarly to [Gomes and Schmid \(2010\)](#), we posit that

$$\log M(z_t, z_{t+1}) \equiv \log \beta + \gamma_0 z_t + \gamma_1 z_{t+1},$$

where $\beta > 0$ is the time discount factor, $\gamma_0 > 0$, and $\gamma_1 < 0$. This specification implies that the conditional risk-free rate equals

$$R_{f,t} = \frac{1}{\beta} e^{[-z_t(\gamma_0 + \rho_z \gamma_1) - \frac{1}{2} \gamma_1^2 \sigma_z^2]}.$$

Notice that $R_{f,t}$ is counter-cyclical if and only if $\gamma_0 > -\rho_z \gamma_1$. The price of risk is constant, as

$$\frac{std(M_{t+1})}{E_t(M_{t+1})} = \sqrt{e^{\gamma_1^2 \sigma_z^2} - 1}.$$

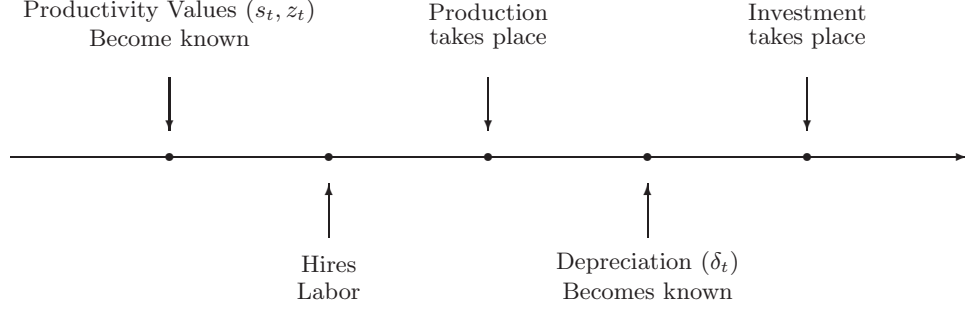


Figure 3: Timing.

The timing is summarized in Figure 3.

3.1 Optimization

For the sake of clarity, we abandon the time notation. State variables are capital in place and the current realizations of aggregate and idiosyncratic productivity, respectively. Conditional on a realization of the depreciation rate δ and a continuation value function $V(z, k, s)$, a firm that grows its capital stock will have a value

$$\begin{aligned} \overline{V}(z, k, s, \delta) = \max_{x > \delta k} \phi_0 \delta k - x(1 + \phi_0) - g(x, k) + \int_{\mathfrak{R}} \int_{\mathfrak{R}} M(z, z') V(z', k', s') dH(s'|s) dJ(z'|z), \\ \text{s.t. } k' = k(1 - \delta) + x. \end{aligned} \quad (1)$$

The first term to the right of the max operator reflects our assumption that maintenance investment occurs at price 1. Conditional on relinquishing capital, firm value will be

$$\begin{aligned} \underline{V}(z, k, s, \delta) = \max_{x < 0} -x(1 - \phi_0) - g(x, k) + \int_{\mathfrak{R}} \int_{\mathfrak{R}} M(z, z') V(z', k', s') dH(s'|s) dJ(z'|z), \\ \text{s.t. } k' = k(1 - \delta) + x. \end{aligned} \quad (2)$$

Finally, the value of firms limiting themselves to maintenance investment can be written as

$$\begin{aligned} V_m(z, k, s, \delta) = \max_{x \in [0, \delta k]} -x + \int_{\mathfrak{R}} \int_{\mathfrak{R}} M(z, z') V(z', k', s') dH(s'|s) dJ(z'|z), \\ \text{s.t. } k' = k(1 - \delta) + x. \end{aligned} \quad (3)$$

The conditional firm value $V(z, k, s)$ is the fixed point of the functional equation

defined by (1), (2), (3), and

$$V(z, k, s) = \max_l e^{s+z} (k^\alpha l^{1-\alpha})^\nu - wl + \int_0^1 \max \{ \underline{V}(z, k, s, \delta), V_m(z, k, s, \delta), \overline{V}(z, k, s, \delta) \} dQ(\delta). \quad (4)$$

Our main object of interest is the expected return on equity, defined as the ratio of expected cum-dividend value at the next date to the current ex-dividend value. With some abuse of notation, let $k^*(\delta)$ denote the optimal choice of capital conditional on the quadruplet (z, k, s, δ) . Then, the conditional return writes as

$$R_e(z, k, s, \delta) = \frac{\int_{\mathbb{R}} \int_{\mathbb{R}} V(z', k^*(\delta), s') dH(s'|s) dJ(z'|z)}{\int_{\mathbb{R}} \int_{\mathbb{R}} M(z, z') V(z', k^*(\delta), s') dH(s'|s) dJ(z'|z)}.$$

3.2 Calibration

We approximate the policy function for capital by means of an algorithm based on the value function iteration method. Together with the stochastic processes of idiosyncratic and aggregate productivity, the policy function is deployed to generate a 5,000-quarter long time-series of the distribution of firms. Except for the first 500 periods of the simulation, that are discarded, the resulting panel is our approximation of the model's ergodic distribution.

Our calibration strategy sets our study apart from any other investigation of equity prices in production-based models, as we do not target any feature of the cross-section of returns. Rather, we require the model to be consistent with our evidence on investment and we evaluate its implications for equity returns. Key moments are listed in Table 2 along with their empirical counterparts. All parameter values are listed in the column labeled “CD” in Table 3.¹⁰

One period is assumed to be one quarter. Unless noted otherwise, all magnitudes reported below are expressed at the quarterly frequency.

The wage rate is normalized to 1. Following Basu and Fernald (1997), the span-of-control parameter ν is set to equal to 0.87. We also posit $\alpha = 0.3$, implying a labor share of approximately 0.6, a value very close to that recovered from the Flow-of-Funds data for the nonfinancial corporate business sector.¹¹ The stochastic process driving the common productivity component is parameterized based on Cooley and Prescott (1995), who set $\rho_z = 0.95$ and $\sigma_z = 0.007$.

¹⁰The column CES lists the parameters of the model considered in Section 5.

¹¹Over the period 1976q1-2016q4, the ratio of compensation of employees to gross value added exhibits a quarterly mean of 0.62.

Table 2: Moments

	Data	Model
Targeted		
Labor share	0.62	0.61
Investment rate		
Standard Deviation	0.095	0.097
Autocorrelation	0.261	0.262
Negative	0.182	0.180
Risk-free rate		
Average (%)	0.4	0.5
Standard Deviation (%)	0.9	0.8
Sharpe ratio	0.22	0.22
Not targeted		
Average investment rate	0.035	0.049
Inaction	0.165	0.306
Book-to-Market	0.811	0.612

The parameters of the stochastic discount factor are set to match the first two unconditional moments of the risk-free, as well as the mean Sharpe ratio. Because of nonlinearities in the map between parameters and moments, there are two distinct sets of values for the triplet $\{\beta, \gamma_0, \gamma_1\}$ that match the targets. One produces a counter-cyclical risk-free rate, while the other generates a pro-cyclical rate. To be consistent with the evidence,¹² we decide to go with the former.

The discrete representation of Q , reported at the bottom of Table 3, was constructed to approximate the frequency distribution of depreciation rates in our sample of firms.¹³ The mean quarterly depreciation rate implied by our calibration equals its empirical counterpart of 2.5%. The standard deviation is 0.96%, a value only slightly smaller than the empirical value of 0.98%. We verified that neither the moments of investment growth nor the asset pricing implications described below change in any appreciable way when we posit a constant depreciation rate equal to its mean. The added heterogeneity has the

¹²See Beaudry and Guay (1996) and Cooper and Willis (2014) among others.

¹³The BEA's depreciation rates employed in Section 2 range between 0.575% and 4.425%. We partition the interval between the two values into five equally sized sub-intervals. The average values within each sub-interval are selected as grid points and the fractions of observations as the corresponding probabilities.

Table 3: Parameters

	Symbol	CD	CES
Not Calibrated			
Span of control	ν	0.87	0.87
Elasticity of substitution	φ	-	-0.67
Autocorrelation aggregate productivity	ρ_z	0.95	0.95
Standard deviation aggregate productivity	σ_z	0.007	0.007
Autocorrelation wage process	ρ_w	-	0.90
Standard deviation wage process	σ_w	-	0.008
Loading of wage on productivity	ρ_{zw}	-	0.07
Correlation aggregate shock with wage shock	$\frac{\sigma_{zw}}{\sigma_z \sigma_w}$	-	0.27
Autocorrelation idiosyncratic productivity	ρ_s	0.97	0.97
Calibrated			
Capital elasticity	α	0.30	-
Capital intensity	ζ	-	0.70
Standard deviation idiosyncratic productivity	σ_s	0.017	0.017
Investment price wedge	ϕ_0	0.0023	0.0022
Capital adjustment cost	ϕ_1	0.00066	0.00056
Parameter pricing kernel	β	0.969	0.969
Parameter pricing kernel	γ_0	31.70	31.70
Parameter pricing kernel	γ_1	-33.00	-33.00

Distribution of depreciation rate δ					
Value	0.0113	0.0180	0.0247	0.0322	0.0408
Probability	0.11	0.30	0.32	0.05	0.22

only effect of getting rid of gaps in the simulated distribution of investment rates, which are caused by the discreteness of our numerical representation for productivity, together with our assumption about the asymmetry in capital prices.

It is left to parameterize the stochastic process driving idiosyncratic productivity and the capital adjustment costs. Given our methodological approach, the ideal targets are mean, standard deviation and autocorrelation of the investment rate, along with the fraction of firms that disinvest. However, as illustrated by [Clementi and Palazzo \(2016\)](#) for a simplified version of the model considered here, the log-normal neoclassical framework does not allow to independently pin down mean and standard deviation of the investment rate.¹⁴

This is why we resort to setting ρ_s arbitrarily and then find the triplet $\{\sigma_s, \phi_0, \phi_1\}$ that matches all targets except the mean investment rate. This is always a feasible strategy.

We set the autocorrelation coefficient to 0.97, as in [Zhang \(2005\)](#). To our knowledge, this value is larger than any available direct estimates. Our rationale for selecting it is to give the model the chance to generate a large cross-sectional variation in productivity. In turn, as will become clear below, this will allow us to interpret the implications for the variation in returns as upper bounds. The implied mean investment growth rate is 4.9%, higher than then its empirical counterpart of 3.5%.

3.3 Results

We illustrate the model’s implications for equity returns by means of a simple methodology commonly used in the empirical asset pricing literature. In every quarter, we form portfolios of firms based on the values assumed by size (ex-dividend firm value) and book-to-market (ratio of current capital to firm value), respectively, and we compute their realized returns. We then compare the time-series means of the returns earned by the different portfolios. These moments are reported in [Table 4](#), along with mean values of size, book-to-market, investment rate, capital in place, and idiosyncratic productivity.

When sorted according to size, stocks in the *bottom* category are those that fall in the two bottom deciles of the distribution in the period of portfolio formation. We will refer

¹⁴[Clementi and Palazzo \(2016\)](#) show analytically that in the absence of capital adjustment costs, (i) the standard deviation of the investment rate is a simple non-linear function of the mean, which only depends on the depreciation rate, and that (ii) there exists an uncountable set of pairs $\{\sigma_s, \rho_s\}$ consistent with a given value of the standard deviation of the investment rate. It follows that that framework, once set the depreciation rate, cannot replicate any arbitrary first and second moments of investment growth. Furthermore, the two moments of investment growth do not identify the pair $\{\sigma_s, \rho_s\}$. While these properties do not hold exact in our model, numerical results reveal that similar restrictions apply.

Table 4: Portfolio Sorts

	Size Sorted	Book-to-Market Sorted
Excess Returns (%)		
Bottom	1.058	0.933
Medium	0.874	0.848
Top	0.694	0.780
Top-Bottom	-0.364	-0.153
Top-Bottom (Data)	-1.248	1.374
Size		
Bottom	0.034	0.042
Medium	0.059	0.061
Top	0.109	0.095
Book-to-Market		
Bottom	0.511	0.486
Medium	0.616	0.612
Top	0.704	0.741
Investment Rate		
Bottom	0.045	0.121
Medium	0.048	0.048
Top	0.056	-0.018
Capital		
Bottom	0.021	0.024
Medium	0.042	0.043
Top	0.085	0.077
Idiosyncratic Shock		
Bottom	-0.096	-0.072
Medium	0.000	0.002
Top	0.097	0.067

to them as small. Stocks in the *top* category, labeled as large, are those falling in the top two deciles.

Consistent with the empirical evidence, on average small firms earn higher returns. The key driver of size heterogeneity is idiosyncratic productivity. Firms can be conceived as portfolios of assets with different systematic risk. Idiosyncratic productivity becomes relevant for risk, because it pins down the loadings of the different assets, exactly as in [Babenko et al. \(2016\)](#).

Since idiosyncratic productivity is persistent, firms characterized by a low value of s will choose a low capital stock and will have lower value. Firms with high s will have higher capital and large market size. Since s is also mean-reverting, small firms derive most of their value from cash-flows to accrue far in the future. Conversely, the value of large firms is mostly due to cash-flows with very short maturity. The size premium follows from the finding that the risk of cash flows is increasing in their maturity.

We refer the reader interested in a more punctual description of this mechanism to appendix [B](#), where we lay out a three-period version of our model that is amenable to analytical characterization. That analysis suggests that as long as aggregate productivity is persistent and returns to scale are reasonably close to constant, the result that more distant cash-flows are riskier is very robust.

Let's now turn to portfolios sorted according to book-to-market. Contrary to the empirical evidence, growth stocks (those in the bottom two deciles of the distribution) fetch higher returns than value stocks (top two deciles). The reason is that growth stocks feature lower productivity and are therefore riskier.

From a theoretical perspective, the relevance of the book-to-market ratio for asset pricing analysis stems from the well-known result that under constant returns to scale and convex adjustment costs, its reciprocal – known as Tobin's q – equals the marginal product of capital (marginal q). Since marginal q and the investment rate, tied to each other by the Euler equation for investment, are positively associated with idiosyncratic productivity in stationary distribution, one would expect value firms to have a higher average idiosyncratic productivity and therefore command lower returns.

However, with decreasing returns – the relevant case here – Tobin's q does not equal marginal q . Most importantly, in stationary distribution Tobin's q is negatively associated with idiosyncratic productivity.

According to the model, the productivity of value firms rises ahead of portfolio formation and declines thereafter. Conversely, the productivity of growth firms declines ahead of the formation date, to recover in the aftermath. This is why growth firms are riskier

and command higher returns.

The data suggests the opposite. As shown by [Imrohoroglu and Tuzel \(2014\)](#), low book-to-market are firms that are expanding their operations in order to catch up with recently risen productivity levels. Conversely, high book-to-market firms consist of large operations that were recently hit by negative news about their productivity.

Our finding of a value discount is clearly at odds with the conclusions reached in influential contributions to the literature. [Zhang \(2005\)](#), for example, shows that in a setup very close to ours, assuming costly investment reversibility and countercyclical price of risk yields a substantial value premium. How to reconcile our results with his?

3.4 The role of capital adjustment costs

[Zhang \(2005\)](#) obtains a value premium – a large one – by imposing strong frictions to the capital adjustment process. With large enough adjustment costs, the association between book-to-market and idiosyncratic productivity is the opposite of the one described above. That is, on average value firms will have lower idiosyncratic productivity and command a higher return on equity.

To envision why this is the case, consider an extreme scenario where adjustment costs are so large that capital is essentially fixed. Then, forming portfolios based on size and book-to-market is equivalent, as the two indicators are monotone increasing and monotone decreasing in idiosyncratic productivity, respectively.

Consistent with this narrative and with the message of [Zhang \(2005\)](#), even in our model there exist values of the marginal capital adjustment costs that yield a value premium instead of a value discount. For example, increasing the parameter ϕ_1 by four orders of magnitude generates a value premium of 0.135% – about one order of magnitude smaller than in the data. Unfortunately, however, as a result the standard deviation of the investment rate plunges to 0.007, the autocorrelation is 0.95, and only a negligible fraction of firms ever reduce the size of their operations. In other words, firms exhibit an investment behavior grossly at odds with the evidence.

In appendix [C](#), we show that Zhang’s model, once disciplined to yield data-conforming values for the volatility of the investment rate, generates a value discount. Consistent with the results shown in [Zhang \(2005\)](#), raising the capital adjustment cost to the levels considered there leads to a sizeable value premium, but at the cost of essentially no cross-sectional variation in investment rates. We also show that, at the margin, asymmetry in capital adjustment costs plays no discernable role.

A further insight of [Zhang \(2005\)](#) is that under certain conditions, operating leverage

– absent in treatment so far – will have non-trivial implications for the cross-sectional dispersion of asset returns. For this reason, we now turn to studying its effects in our setup.

4 Operating Leverage

We begin by following [Zhang \(2005\)](#) in assuming that every period firms pay a cost $c_f > 0$, independent from the scale of operations. An immediate implication is that for low enough realizations of productivity, firm value will turn negative. In turn, this precipitates the need to explicitly model exit.

At the beginning of each period, after the realization of the aggregate and firm-level shocks but before the realization of the depreciation rate, firms have the opportunity to cease operations. If they do, they receive current operating profits along with the proceeds from the sale of the undepreciated capital, net of the adjustment cost required to uninstall it. It follows that the value of exit is

$$\max_l e^{s+z} (k^\alpha l^{1-\alpha})^\nu - wl + \int_0^1 [(1 - \phi_0)(1 - \delta)k - g((1 - \delta)k, k)] dQ(\delta).$$

Upon exit, a firm is replaced by another, randomly drawn from the stationary distribution. This ensure that the simulated panel stays balanced.

We set c_f to generate an average exit rate of 0.5%, equal to the average quarterly rate of delisting for poor performance in our sample.¹⁵ The implied average book-to-market ratio is 0.888, rather close to our empirical estimate of 0.811.

On average, modeling operating leverage in this fashion increases cash-flow risk for firms with low idiosyncratic productivity and reduces it for those with high productivity. This is the case because introducing the fixed cost is akin to having shareholders take a short position on the risk-free asset.

For small firms, operating assets command a higher return than the risk-free rate. As result, leveraging makes equity riskier. The opposite is true for high-productivity firms, whose operating assets yield on average a return lower than risk-free. The reason is that for such firms, negative aggregate shocks trigger a sale of assets and an increase in payouts, turning them into good hedges.

A further effect deriving from the introduction of operating leverage is the reshaping of the variation of the book-to-market ratio over the ergodic set. Recall that in our baseline scenario, which features $c_f = 0$, book-to-market decreases with productivity in stationary

¹⁵More details on the calculation of the exit rate can be found in [Appendix A.2](#).

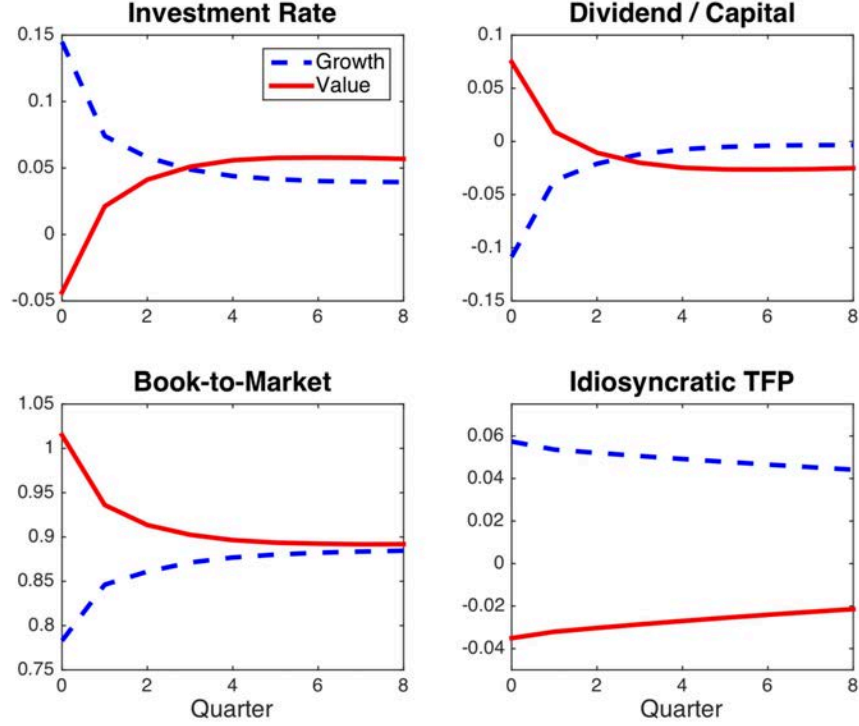


Figure 4: Firm dynamics after portfolio formation – With Operating Leverage

distribution, yielding a value discount. As c_f rises, the slope of such association decreases in absolute value and eventually changes sign, allowing for a value premium. The reader interested in a firmer illustration of this feature is referred to Appendix B.

Figure 4 helps us appreciate the dynamics of growth and value stocks in the aftermath of portfolio formation. We report the time-series averages of the portfolio means of investment rate, dividend to capital ratio, book-to-market, and idiosyncratic TFP over the 8 quarters following portfolio formation.

At portfolio formation, value firms are endowed on average with capital stocks substantially greater than the optimal level dictated by their efficiency. Growth firms, on the other hand, have less capital in place than warranted by their relatively high idiosyncratic productivity. As a result, for a number of periods growth firms invest heavily, requiring new resources from their shareholders, while value firms divest and pay out dividends. This occurs while productivity mean-reverts for both sets of firms.

These dynamics are in line with the evidence. The evolution of idiosyncratic productivity is consistent with Imrohoroglu and Tuzel (2014): TFP is higher for growth firms than value firms, but the gap between the two declines after portfolio formation. Asset sellers have below average productivity, in line with the evidence presented by Maksimovic

and Phillips (2001) and Schoar (2002). Value firms have shorter cash-flow duration, as documented by Dechow et al. (2004).

The takeaway is that value firms are riskier, and command a greater equity return, because their idiosyncratic productivity is low. The divesting activity tends to reduce their risk, since the associated payouts occur even in bad aggregate states of nature. By the same token, equities of growth firms are safer and yield a lower return, because their idiosyncratic productivity is high. The investment activity contributes positively to their risk, as it requires resources from shareholders in all aggregate states of nature.

The above discussion also clarifies why, consistent with the empirical findings of Xing (2008) among others, an investment strategy calling for a long position on low investment-rate stocks and a short position on high investment-rate stocks yields a positive return on average.

With the help of Table 5, we now turn to a quantitative evaluation of the role played by operating leverage.¹⁶ Unfortunately, the magnitude of the value premium is only 0.454% per quarter, or about 1/3 of the value we estimated from our data.¹⁷

Small firms earn, on average, a value-weighted excess return of around 0.6% per quarter over large firms, or about half the estimated average quarterly value weighted size premium over the period 1976:q1–2016:q4.

Notice that, counterfactually, the value premium is smaller than the size premium. This is a robust result, which can be rationalized by two observations. First, the dispersion of idiosyncratic productivity – the main determinant of variation in risk – is greater across firms of different size than across firms of different book-to-market. Second, and more interesting, on average value firms shed capital after portfolio formation, regardless of the aggregate state of nature. This makes them less risky. Growth firms, on the other hand, tend to invest – drawing resources from shareholders – in all aggregate states. This feature makes them riskier.

One may argue that the model’s failure to generate a sizable value premium follows directly from its inability to generate enough cross-sectional dispersion of returns. Our response is that as long as we require it to be consistent with the cross-sectional evidence on investment, the model simply cannot generate a greater dispersion of returns!

In fact, the value premium we report is an upper bound. Because of cross-moment

¹⁶In Table 5, portfolios labeled *bottom* identify small firms in the first column and growth firms in the second column. Analogously, portfolios labeled *top* include large firms in the first column and value firms in the second.

¹⁷Over the period 1976:q1–2016:q4, Kenneth French’s data generate an average quarterly value-weighted value premium of 1.305%. See http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

Table 5: Portfolio Sorts with Operating Leverage

	Size Sorted	Book-to-Market Sorted
Excess Returns (%)		
Bottom	1.244	0.551
Average	0.839	0.879
Top	0.650	1.005
Top-Bottom	-0.594	0.454
Top-Bottom (Data)	-1.248	1.374
Size		
Bottom	0.026	0.063
Average	0.051	0.057
Top	0.102	0.046
Book-to-Market		
Bottom	0.930	0.783
Average	0.882	0.878
Top	0.858	1.016
Investment Rate		
Bottom	0.039	0.144
Average	0.046	0.044
Top	0.056	-0.044
Capital		
Bottom	0.023	0.050
Average	0.045	0.050
Top	0.089	0.046
Idiosyncratic Shock		
Bottom	-0.067	0.058
Average	0.016	0.022
Top	0.107	-0.035

Table 6: When operating costs increase with capital

	Low a_0	High a_0
Investment rate		
Standard deviation	0.095	0.094
Autocorrelation	0.266	0.266
Negative	0.181	0.181
Excess Returns (%)		
Growth	0.623	0.725
Medium	0.942	1.027
Value	1.030	1.105
Value-Growth	0.407	0.379
Idiosyncratic Shock		
Growth	0.036	0.021
Medium	0.015	0.008
Value	-0.021	-0.013
Value-Growth	-0.057	-0.034

restrictions endemic to the log-normal environment – see Section 3.2 – the autocorrelation of idiosyncratic productivity ρ_s is a free parameter. We set it equal to 0.97, much higher than any available estimates, in order to give the model the best chance to generate dispersion in returns. Lower values of ρ_s are going to imply a lower dispersion and a lower value premium.

Once more, our conclusions appear to be at odds with the extant literature. For example, Carlson et al. (2004) argue that imbedding operating leverage in a production model accounts both qualitatively and quantitatively for the book-to-market effect.

Key differences between our framework and theirs are in the constraints to investment activity and in the formulation of operating costs, which in their case are increasing in the size of capital in place. In order to identify which of the two items is chiefly responsible for the difference in the conclusions, we now assume that operating costs are of the form $c_f + a_0 k_t^{a_1}$. The curvature parameter a_1 is estimated to be 0.65.¹⁸ Regarding the value

¹⁸The value for a_1 results from a pooled non-linear estimation of the cost function $c = \psi + \psi_0 k^{\psi_1}$, where c is the real value of selling, general, and administrative expenses (COMPUSTAT item *xsgaq*) and k is the real value of net property, plant, and equipment (*ppentq*). To dampen the effect of extreme observations,

of the scaling factor a_0 , we experiment with a range of parameter values. We begin with $a_0 = 0.0039$ – see the column labeled *low* a_0 in Table 6. The parameters governing the investment rate process $\{\sigma_s, \phi_0, \phi_1\}$, as well as c_f , are re-calibrated to match the usual targets for the cross-sectional investment moments and the exit rate.

The value premium is essentially unchanged with respect to the exercise conducted above. The comparative statics with respect to a_0 helps understanding why.

Holding investment behavior constant, raising a_0 by 50 percent does increase the dispersion in returns. However, by impacting the marginal cost of accumulating capital, a change in a_0 affects firms’ investment. Since $a_1 < 1$, a hike in a_0 raises the rate at which the marginal cost of investment declines with k . In turn, this leads to a higher cross-sectional variation in investment rates. When the conditional volatility of idiosyncratic productivity σ_s is adjusted downward to bring the volatility of investment back to its target value, both the dispersion of returns and the value premium drop. See the right column in Table 6, labeled *high* a_0 .

We conclude that, as long as we require the model to be consistent with the cross-sectional evidence on investment, allowing operating leverage to scale up with operating assets does not lead to an increase in the dispersion of equity returns. In turn, this also suggests that the reason why [Carlson et al. \(2004\)](#) obtain a much larger value premium is that the implied firm-level investment dynamics is at odds with the evidence.

This is transparent in the case of the very stylized setup that [Carlson et al. \(2004\)](#) use to illustrate the economic mechanism of interest. Since profitability is persistent, a good sequence of shocks leads firms to accumulate a large capital stock, in turn associated with large operating costs. Since investment is irreversible by assumption, firms whose fortune has reversed – value firms – are riskier. Unfortunately, the results we reported in Section 2 suggest that the irreversibility assumption has no empirical support.

In the model utilized by [Carlson et al. \(2004\)](#) for their quantitative analysis, shedding capital is feasible but is heavily discouraged by the assumption that capital expenditures are completely sunk. The other parameters shaping adjustment costs, as well as the volatility of idiosyncratic productivity, are estimated by matching the mean return on decile portfolios of size-sorted and book-to-market-sorted returns. Unfortunately, the model’s implications for investment are not provided.

we eliminate firms with a value of the cost-to-capital ratio ($xsgaq/ppentq$) in the top and bottom 1% of the distribution. The estimated value of ψ_1 is 0.654 with a standard error of 0.015.

5 Labor leverage

In Section 3.4 we have argued that when capital is fixed, production is Cobb–Douglas, and the only source of firm–level heterogeneity is in idiosyncratic, mean–reverting productivity, value firms are riskier simply because they are less productive. Since the labor share is constant, employment decisions have no effect on risk and returns.

As highlighted in Donangelo et al. (2018), relaxing the Cobb–Douglas assumption in such environment, to allow for an elasticity of substitution between capital and labor lower than one, leads the labor share to move counter–cyclically and to covary negatively with idiosyncratic productivity in the cross–section. In turn, this means a wider cross–sectional variation in risk – i.e. a greater value premium.

Does labor leverage contribute to enhancing the value premium even in a scenario, such as ours, where capital is not fixed and firms’ investment process is forced to conform with the evidence? To address this question, we generalize the production function to

$$y_t = e^{s_t + z_t} [\zeta k_t^\varphi + (1 - \zeta) l_t^\varphi]^{\nu/\varphi},$$

where $\zeta \in (0, 1)$ measures capital intensity, $\nu \in (0, 1)$ is still the span–of–control parameter, and $1/(1 - \varphi)$ is the elasticity of substitution between capital and labor. The Cobb–Douglas case obtains in the limit as φ goes to 0.

In spite of the large volume of research on the subject, there is still wide disagreement about the value of the elasticity of substitution between capital and labor. Here we follow Chirinko (2008) in considering values in the interval $[0.4, 0.6]$. We begin with the upper bound, which is tantamount to assuming $\varphi = -2/3$. In turn, ζ is set to match the average aggregate labor share. The span of control parameter is unchanged.

We further allow for realistic times–series variation in wages. We assume that the pair $\{z_t, w_t\}$ evolves according to the process

$$\begin{bmatrix} z_{t+1} \\ w_{t+1} \end{bmatrix} = \begin{bmatrix} \rho_z & \rho_{wz} \\ \rho_{zw} & \rho_w \end{bmatrix} \begin{bmatrix} z_t \\ w_t \end{bmatrix} + \begin{bmatrix} \varepsilon_{z,t+1} \\ \varepsilon_{w,t+1} \end{bmatrix}, \quad (5)$$

where the vector of innovations is jointly normal with mean zero and variance–covariance matrix $\begin{bmatrix} \sigma_z^2 & \sigma_{wz} \\ \sigma_{wz} & \sigma_w^2 \end{bmatrix}$.

The empirical counterparts for z_t and w_t are the deviations from their respective log–linear trends of John Fernald’s business sector total factor productivity and the BLS’ seasonal adjusted real compensation per hour in the non–farm business sector, respectively, over the period 1964q1–2016q1.

Our estimation procedure, carefully described in appendix A.3, leads us to set $\rho_z = 0.95$, $\rho_{wz} = 0$, and $\sigma_z = 0.007$, exactly as in Section 3. The wage is also highly persistent,

Table 7: Moments – CES production

		(1)	(2)	(3)	(4)	(5)
	Data	Baseline	OL	OL	OL	OL
				$\varphi = -1.5$	$\varphi = -1.5$	$\varphi = 0$
					$\zeta = 0.95$	$\zeta = 0.15$
Targeted						
Labor share	0.62	0.63	0.63	0.75	0.64	0.74
<u>Investment rate</u>						
Standard Deviation	0.095	0.094	0.093	0.092	0.093	0.092
Autocorrelation	0.261	0.262	0.263	0.259	0.265	0.261
Negative	0.182	0.176	0.180	0.194	0.187	0.186
<u>Risk-free rate</u>						
Average (%)	0.4	0.4	0.4	0.4	0.4	0.4
Standard Deviation (%)	0.9	0.9	0.9	0.9	0.9	0.9
Sharpe ratio	0.22	0.20	0.21	0.23	0.21	0.22
Not targeted						
Average investment rate	0.035	0.054	0.051	0.034	0.047	0.042
Inaction	0.165	0.282	0.288	0.251	0.278	0.270
Book-to-Market	0.811	0.606	0.892	0.771	0.888	0.794

with $\rho_w = 0.9$, $\rho_{zw} = 0.07$, and $\sigma_w = 0.008$. The correlation between the two innovations is 0.276.

In the numerical implementation, we approximate the bi-variate VAR process by means of the finite-state Markov chain approximation methodology due to [Gospodinov and Lkhagvasuren \(2014\)](#).¹⁹

We set the remaining parameters following the same procedure as in Section 3. In particular, we strive to match key features of the cross-sectional distribution of investment.

5.1 Results

We start by considering the case for $c_f = 0$. Targeted and non-targeted moments are listed in column (1) of Table 7, while the parameter values can be found in column “CES” of Table 3.

Regressing the labor share on the logarithm of aggregate TFP growth over the pooled

¹⁹See appendix A.3 for an assessment of the performance of our approximation.

simulated data, yields a coefficient of -0.69, confirming that labor leverage is a risk-factor. For the sake of comparison, the same regression on actual data by [Donangelo et al. \(2018\)](#) yields a coefficient of -0.45.

The implications for the cross-section of asset prices are best appreciated with the help of column (1) in Table 8, which lists mean expected return, labor share, and idiosyncratic shock across portfolios with low, medium, and high book-to-market, respectively. Once again, the model generates a value discount. The reason is the same as in Section 3: Growth firms are characterized by lower idiosyncratic productivity, and are therefore riskier.

A second observation is that the labor share varies very little across portfolios characterized by different book-to-market ratios. This is the case because an increase in idiosyncratic productivity is accommodated by changes in both labor and capital. The flexibility of capital mutes dramatically the effect that reducing the elasticity of substitution has on the labor share variation in the scenario of [Donangelo et al. \(2018\)](#), where capital is assumed to be fixed.

Column (2) refers to the scenario in which we re-introduce operating leverage in the form of a value for c_f that generates the same exit rate as in the data. As it was the case in Section 4, the sign of the cross-sectional association between risk and book-to market changes and a value premium appears.

Operating leverage and labor leverage interact to generate a non-negligible variation in the *fixed cost share*, i.e. the ratio of c_f to operating profits $y_t - w_t l_t$. Since they are less productive, value firms feature a greater labor share and a greater fixed cost share, which in turns implies a higher risk. The value premium, however, is still substantially smaller than in the data.

We turn next to consider a value for the elasticity of substitution equal to 0.4, the level favored by [Chirinko and Mallick \(2017\)](#). See column (3). The dispersion of the labor share across portfolios increases, and so does the value premium. However, the former finding is not the driver of the latter. The dispersion in returns increases, because the average labor share is higher, counter-factually so. A higher labor share is responsible for the greater dispersion in the fixed cost share and, ultimately, expected returns.

To see that this is indeed the case, consider the scenario of which in column (4). When raising ζ in order to lower the model-implied average labor share to its targeted value, the average fixed cost shares across portfolio and the value premium revert to the same values as in column (2).

Finally, in column (5) we reconsider the Cobb-Douglas scenario, taking care of setting

Table 8: Book-to-Market Portfolios

	(1)	(2)	(3)	(4)	(5)
	Baseline	OL	OL	OL	OL
			$\varphi = -1.5$	$\varphi = -1.5$	$\varphi = 0$
				$\zeta = 0.95$	$\zeta = 0.15$
Excess Returns (%)					
Growth	0.985	0.674	0.684	0.674	0.657
Medium	0.882	0.863	1.005	0.874	0.959
Value	0.805	1.036	1.226	1.040	1.227
Value-Growth	-0.18	0.362	0.542	0.366	0.570
Labor share					
Growth	0.626	0.623	0.743	0.629	0.740
Medium	0.628	0.629	0.751	0.637	0.740
Value	0.631	0.634	0.759	0.645	0.740
Fixed cost share					
Growth	0.000	0.145	0.230	0.149	0.223
Medium	0.000	0.178	0.313	0.181	0.306
Value	0.000	0.253	0.485	0.258	0.467
Idiosyncratic Shock					
Growth	-0.072	0.059	0.075	0.058	0.076
Medium	0.002	0.022	0.023	0.022	0.023
Value	0.068	-0.036	-0.043	-0.036	-0.043

ζ to generate an average labor share as counter-factually high as the one reached in column (3). The model generates the dispersion in fixed cost shares needed to replicate the value premium obtained when the elasticity of substitution is 0.4.

The upshot is that, differently from what happens in simpler setups with constant capital, amending the neoclassical framework to allow labor leverage to be a risk factor yields very little payoff in terms of cross-sectional dispersion of equity returns. In particular, the value premium is substantially smaller than in the data. This result reflects in large part the limited cross-sectoral variation in labor share that obtains when capital can adjust optimally to innovations in idiosyncratic productivity.

Cognizant of this outcome, a valuable research direction may consist of departing further from the neoclassical model. This may mean allowing for labor adjustment costs,

firing taxes, or alternative means to drive a wedge between wages and the marginal product of labor.

Indeed, recently [Favilukis and Lin \(2016\)](#) have argued that allowing firms to reset wages for only a small fraction of their workforce substantially per period, boosts the value premium. [Marfè \(2017\)](#) obtains a sizeable value premium in a model in which wages are the outcome of an intra-firm risk-sharing scheme between shareholders and workers.

Crucially, neither of these contributions confront their results with the evidence on firm-level investment that disciplines our exercises. In [Marfè \(2017\)](#), capital is fixed exogenously. In [Favilukis and Lin \(2016\)](#), it is allowed to vary. However, these authors retain the same parametric assumptions about the process driving idiosyncratic productivity across all of their models. As a result, they end up comparing allocations characterized by different cross-sectional dispersions in investment rates.

Our own results suggest that a fraction of the difference in returns volatility they report across models simply reflects the difference in investment growth dispersion. What fraction, we do not know.

6 Conclusion

In the one-factor neoclassical investment model with mean-reverting idiosyncratic productivity, value firms command greater expected returns than growth firms because, in common with small firms, they derive a larger fraction of their value from cash-flows accruing far in the future, which are riskier. Unlike small firms, however, value firms are burdened with excess capacity. Therefore their risk also depends on their ability to reduce such capacity by divesting operating assets.

The data strongly suggests that US public firms do adjust to adverse profitability shocks by relinquishing capital. When capital adjustment costs are parameterized to reflect this feature of the data, states of nature characterized by low aggregate productivity (high marginal utility) see value firms to disinvest. This makes them safer, leading to a lower value premium.

These conclusions, which hold true in scenarios where either operating leverage or labor leverage are modeled in ways that were deemed promising in the extant literature, cast a doubt on the quantitative significance of what in the last decade has become the prevalent theoretical paradigm for the study of cross-sectional variation of asset returns. Research strategies that aim at obtaining greater variation in returns by limiting firms' chances to adjust to shocks are likely to have empirically implausible consequences for

investment behavior and therefore don't appear promising.

Our conclusions do not necessarily extend to two-factor models such as those introduced by Papanikolaou (2011), Kogan and Papanikolaou (2014), Belo et al. (2014), and Garlappi and Song (2017). Adding one risk factor provides one more degree of freedom. If the objective is to rationalize both the size and value premium, one can, at least in principle, assign one factor to each source of variation in returns.

For example, consider Belo et al. (2014). The stochastic discount factor loads on aggregate shocks to productivity and to the adjustment cost of capital, respectively. The loadings are set to match the Shape ratio and the value premium.

Importantly, the loading on the shock to the adjustment cost of capital is negative: In states of nature where adjustment costs are low, state prices are (conditionally) high. It follows that, upon a negative innovation to the adjustment cost, firms whose capital is lower than efficient (growth firms) will increase their value disproportionately exactly when marginal utility is relatively high. In turn, this makes growth firm less risky and contributes to increasing the value premium.

What is not entirely clear to us is the extent to which the claims of success in the two-factor literature also depend on frictions to the capital adjustment process and whether the resulting variation in investment rate is consistent with our evidence. For example, Belo et al. (2014) argue that even in their scenario, costly reversibility is important to generate a sizeable value premium. Unfortunately, they do not discuss whether such assumption impairs firms in the model from reacting to adverse shocks as in the data, i.e. by divesting assets – large portions of them, if needed.

References

- Abel, A. and J. Eberly (1994). A unified model of investment under uncertainty. *American Economic Review* 84(5), 1369–1384.
- Babenko, I., O. Boguth, and Y. Tserlukevich (2016). Idiosyncratic cash flows and systematic risk. *Journal of Finance* 71(1), 425–455.
- Barnett, S. A. and P. Sakellaris (1998). Nonlinear response of firm investment to q: Testing a model of convex and non-convex adjustment costs. *Journal of Monetary Economics* 42(2), 261–288.
- Basu, S. and J. Fernald (1997). Returns to scale in us production: Estimates and implications. *Journal of Political Economy* 105(2), 249–283.

- Basu, S., J. Fernald, and M. Kimball (2006). Are technology improvements contractionary? *American Economic Review* 96(5), 1418–1448.
- Bayraktar, N. (2002). Effects of financial market imperfections and non-convex adjustment costs in the capital adjustment process. Working Paper.
- Beaudry, P. and A. Guay (1996). What do interest rates reveal about the functioning of real business cycle models? *Journal of Economic Dynamics and Control* 20, 1661–1682.
- Belo, F., X. Lin, and S. Bazdresch (2014). Labor hiring, investment, and stock return predictability in the cross section. *Journal of Political Economy* 122(6), 129–177.
- Berk, J. B., R. C. Green, and V. Naik (1999). Optimal investment, growth options, and security returns. *Journal of Finance* 54(5), 1553–1607.
- Campbell, J., J. Hilscher, and J. Szilagyi (2008). In search of distress risk. *Journal of Finance* 63(6), 2899–2939.
- Carlson, M., A. Fisher, and R. Giammarino (2004). Corporate investment and asset price dynamics: Implications for the cross-section of returns. *Journal of Finance* 59(6), 2577–2603.
- Chirinko, R. (2008). σ : The long and short of it. *Journal of Macroeconomics* 30(2), 671–686.
- Chirinko, R. and D. Mallick (2017). The substitution elasticity, factor shares, and the low-frequency panel model. *American Economic Journal: Macroeconomics* 9(4), 225–253.
- Clementi, G. L. and B. Palazzo (2016). On the calibration of competitive industry dynamics models. Working paper, Boston University.
- Cooley, T. F. and E. C. Prescott (1995). Economic growth and business cycles. In T. F. Cooley (Ed.), *Frontiers of Business Cycle Research*, Chapter Chapter 1, pp. 1–37. Princeton University Press.
- Cooper, I. (2006). Asset pricing implications of nonconvex adjustment costs and irreversibility of investment. *Journal of Finance* 61(1), 139–170.
- Cooper, R. and J. Willis (2014). Discounting: Investment sensitivity and aggregate implications. Working Paper, Penn State.

- Cooper, R. W. and J. C. Haltiwanger (2006). On the nature of capital adjustment costs. *Review of Economic Studies* 73(3), 611–633.
- Dechow, P., R. Sloan, and M. Soliman (2004). Implied equity duration: A new measure of equity risk. *Review of Accounting Studies* 9, 197–228.
- Dimopoulos, T. and S. Sacchetto (2017). Merger activity in industry equilibrium. *Journal of Financial Economics* 126(1), 200–226.
- Doms, M. and T. Dunne (1998). Capital adjustment patterns in manufacturing plants. *Review of Economic Dynamics* (1), 409–429.
- Donangelo, A., F. Gourio, M. Kehrig, and M. Palacios (2018). The cross-section of labor leverage and equity returns. *Journal of Financial Economics*, Forthcoming.
- Eckbo, E. and K. Thorburn (2008). Corporate restructuring: Breakups and lbo's. In E. Eckbo (Ed.), *Handbook of Corporate Finance: Empirical Corporate Finance*, Volume 2, Chapter 16, pp. 431–496. North-Holland/Elsevier.
- Eisfeldt, A. and A. Rampini (2006). Capital reallocation and liquidity. *Journal of Monetary Economics* 53, 369–399.
- Favilukis, J. and X. Lin (2016). Wage rigidity: A quantitative solution to several asset pricing puzzles. *Review of Financial Studies* 1(1), 148–192.
- Garlappi, L. and Z. Song (2017). Can investment shocks explain the cross section of equity returns? *Management Science* 63(11), 3829–3848.
- Gomes, J. and L. Schmid (2010). Levered returns. *Journal of Finance* 65(2), 467–494.
- Gomes, J. F. (2001). Financing investment. *American Economic Review* 91(5), 1263–1285.
- Gomes, J. F., L. Kogan, and L. Zhang (2003). Equilibrium cross section of returns. *Journal of Political Economy* 111(4), 693–732.
- Gospodinov, N. and D. Lkhagvasuren (2014). A moment-matching method for approximating vector autoregressive processes by finite-state markov chains. *Journal of Applied Econometrics* 29, 843–859.
- Guthrie, G. (2011). A note on operating leverage and expected equity returns. *Finance Research Letters* 8, 88–100.

- Imrohoroglu, A. and S. Tuzel (2014). Firm level productivity, risk, and return. *Management Science* 60(8), 2073–2090.
- Jermann, U. (1998). Asset pricing in production economies. *Journal of Monetary Economics* 41(2), 257–275.
- Kogan, L. and D. Papanikolaou (2014). Growth opportunities, technology shocks, and asset prices. *Journal of Finance* 69(2), 675–718.
- Kogan, L., D. Papanikolaou, and N. Stoffman (2018). Winners and losers: creative destruction and the stock market. *Journal of Political Economy*, Forthcoming.
- Livdan, D., H. Sapriza, and L. Zhang (2009). Financially constrained stock returns. *Journal of Finance* 44(4), 1827–1862.
- Maksimovic, V. and G. Phillips (2001). The market for corporate assets: Who engages in mergers and asset sales and what are the efficiency gains? *Journal of Finance* 56, 2019–2065.
- Marfè, R. (2017). Labor rigidity and the dynamics of the value premium. Working Paper, Collegio Carlo Alberto.
- Obreja, I. (2013). Book-to-market equity, financial leverage, and the cross-section of stock returns. *Review of Financial Studies* 5(1), 1146–1189.
- Papanikolaou, D. (2011). Investment shocks and asset prices. *Journal of Political Economy* 119(4), 639–685.
- Rouwenhorst, G. K. (1995). Asset pricing implications of equilibrium business cycle models. In T. F. Cooley (Ed.), *Frontiers of Business Cycle Research*, Chapter Chapter 10, pp. 294–330. Princeton University Press.
- Schmid, L. and L.-A. Kuehn (2014). Investment-based corporate bond pricing. *Journal of Finance* 6(69), 2741–2776.
- Schoar, A. (2002). Effects of corporate diversification on productivity. *Journal of Finance* 57, 2379–2403.
- Slovin, M. B., M. E. Sushka, and J. A. Poloncheck (2005). Methods of payment in asset sales: Contracting with equity versus cash. *Journal of Finance* 60(5), 2385–2407.

Tuzel, S. and M. B. Zhang (2017). Local risk, local factors, and asset prices. *Journal of Finance* 72(1), 325–370.

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

A Data

In this appendix, we provide details on a number of methodological choices we made when estimating financial returns, the exit rate, and the VAR(1) process introduced in Section 5.

A.1 Returns

The risk-free rate is the difference between the return on the 90-days US Treasury Bill (t90ret) and the rate of change in the Consumer Price Index (cpiret) – both from CRSP – over the period 1976q1-2016q4. The market excess return is the difference between the total value-weighted market return (vwret) and return on the 90-days Bills (t90ret). The Sharpe ratio is the average market excess return divided by its standard deviation.

When sorting stocks based on either book-to-market or market capitalization,²⁰ portfolios are formed at the beginning of January, April, July, and October using firm’s accounting data available at least three months prior. The first portfolio is formed in July 1978, the last one in October 2016. Once the portfolios are formed, we track their monthly returns for the next three months.

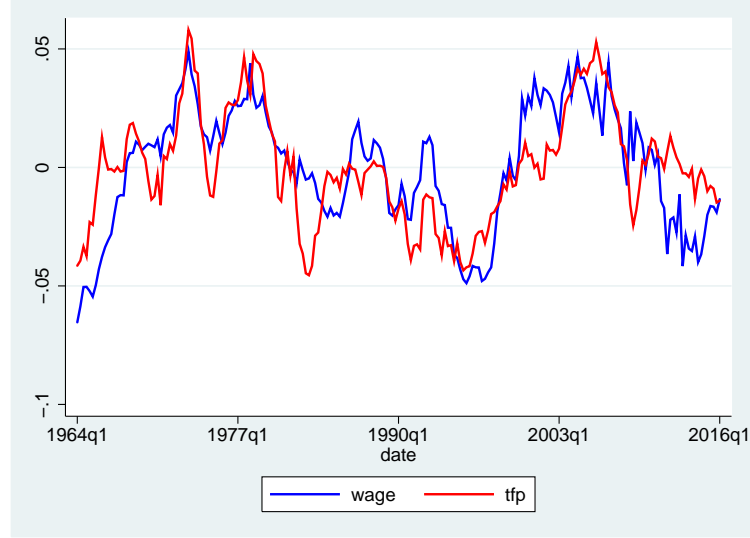
The size premium is the difference between the value-weighted average return among stocks in the bottom quintile of the size distribution (small size stocks) and the value-weighted average return among stocks in the top quintile (large size stocks). Similarly, for the value premium.

A.2 Exit Rate

We identify exits as the performance delistings associated to the following CRSP codes: 560 (insufficient capital, surplus, and/or equity), 561 (insufficient float or assets), 574 (bankruptcy, declared insolvent), 580 (delinquent in filing, non-payment of fees), and 584 (does not meet exchanges financial guidelines for continued listing). A quarter’s exit rate is the fraction of firms that are in the sample at the start and are delisted over the next three months.

²⁰ The book value of equity is equal to shareholders’ equity (item *SEQQ*) plus deferred taxes and investment tax credit (item *txditcq*, if available) minus the book value of preferred stock (item *PSTKRG*). If shareholders’ equity is not available, we use common equity (item *CEQQ*) plus the carrying value of the preferred stock (item *PSTKQ*). If common equity is not available, we measure shareholders’ equity as the difference between total assets (item *ATQ*) and total liabilities (item *LTQ*). The book-to-market ratio is the book value of equity divided by the market capitalization. Market capitalization is calculated using data from CRSP and it is equal to the number of shares outstanding (item *SHROUT*) multiplied by the share price (item *PRC*).

Figure 5: Wage and TFP



The average quarterly exit rate of 0.5% implies an annual average exit rate of 2%, which is roughly mid-way between the rates computed by [Campbell et al. \(2008\)](#) for the period 1976-2003 – 1.1% – and [Dimopoulos and Sacchetto \(2017\)](#) for the period 1981-2010 – 3.7%.

A.3 Estimation and approximation of the VAR process for wages and TFP

The wage rate is proxied by real compensation per hour in the nonfarm business sector, provided by the the U.S. Bureau of Labor Statistics.²¹ For TFP, we use the series for the U.S. business sector available on the Federal Reserve Bank of San Francisco’s website.²² The data is adjusted for variations in factor utilization, labor effort and capital’s workweek, following the methodology introduced by [Basu et al. \(2006\)](#). Both series are quarterly.

We consider deviations from a log-linear trend for the period 1964q1–2016q1. See Figure 5. Estimation of the autoregressive process (5) yields point estimates

$$\begin{bmatrix} \hat{\rho}_z = 0.914^{***} & \hat{\rho}_{wz} = 0.036 \\ \hat{\rho}_{zw} = 0.058^{**} & \hat{\rho}_w = 0.898^{***} \end{bmatrix}.$$

The elements of the variance-covariance matrix are $\hat{\sigma}_z = 0.007$, $\hat{\sigma}_w = 0.008$, and $\sigma_{zw}/(\sigma_z\sigma_w) = 0.276$, all statistically significant at the 1% confidence level.

²¹See <https://fred.stlouisfed.org>, item COMPRNFB.

²²See <https://www.frbsf.org/economic-research/indicators-data/>.

Table A.1: VAR: Assessment of the Approximation

	Actual VAR	Approximation
TFP Standard Deviation	0.023	0.023
TFP Autocorrelation	0.942	0.950
Wage Standard Deviation	0.026	0.024
Wage Autocorrelation	0.936	0.941
$\text{Corr}(\hat{\varepsilon}_z, \hat{\varepsilon}_w)$	0.276	0.278
$\text{Corr}(z, w)$	0.723	0.624

Since $\hat{\rho}_{wz}$ is not significant, we estimate a constrained VAR imposing $\rho_{wz} = 0$. As expected, the point estimates change marginally to

$$\begin{bmatrix} \hat{\rho}_z = 0.942^{***} \\ \hat{\rho}_{zw} = 0.066^{**} \quad \hat{\rho}_w = 0.888^{***} \end{bmatrix}.$$

The estimated variance-covariance matrix is unchanged.

In light of this evidence, we posit $\rho_z = 0.95$, $\sigma_z = 0.007$, as in the baseline model. Furthermore, we set $\rho_w = 0.90$, $\rho_{zw} = 0.07$, $\sigma_w = 0.008$, and $\sigma_{wz} = 0.0000151$. As noted in the main text, we approximate the bi-variate VAR process by means of the finite-state Markov chain approximation methodology developed by [Gospodinov and Lkhagvasuren \(2014\)](#), with 11 gridpoints for each of the variables.

In Table [A.1](#) we compare some key moments for the tfp and wage process with their simulated counterparts obtained simulating a bivariate VAR for 15,000 quarters. The approximated model is able to replicate very closely the standard deviation and the first order autocorrelation of the two process. The implied correlation between ε_z and ε_w is very close to the one observed in the data (0.278 vs 0.276) and this generates a correlation between the tfp and the wage processes equal to 0.624, slightly lower than the actual one (0.723).

B A Three-Period Model

With the purpose of providing intuition for the results in Section [3](#), we characterize analytically the implications for equity returns of a 3-period model that shares all the key features of the more general setup analyzed there.

The time periods are indexed by $t = 0, 1, 2$. Firms produce according to $y_t = e^{s_t + z_t} k_t^\alpha$, where $\alpha \in (0, 1)$. Capital depreciates at the rate $\delta \in (0, 1)$. Dividends equal cash flows minus investment.

Table B.1: Payoffs

	$t = 1$	$t = 2$
Current Asset	$e^{s_1+z_1}k_1^\alpha + k_1(1 - \delta)$	0
Continuation Asset	$-k_2$	$e^{s_2+z_2}k_2^\alpha + k_2(1 - \delta)$

The variables s_t and z_t denote the idiosyncratic and aggregate components of productivity, respectively. Both evolve according to first-order autoregressive processes with independent, normally distributed innovations, exactly as in Section 3. At any time t , firms evaluate cash flows accruing at $t + 1$ according to the stochastic discount factor $M_{t+1} \equiv M(z_t, z_{t+1})$. It follows that, conditional on capital k_1 and productivity levels $\{s_1, z_1\}$, the cum-dividend value of equity at $t = 1$ is

$$V_1(k_1, s_1, z_1) \equiv \max_{k_2} e^{s_1+z_1}k_1^\alpha + k_1(1 - \delta) - k_2 + E_1[M_2[e^{s_2+z_2}k_2^\alpha + k_2(1 - \delta)]],$$

where the linear operator E_s denotes the expectation taken conditional on the information known at $t = s$. As of $t = 0$, the firm's optimization problem is

$$\max_{k_1} -k_1 + E_0[M_1 V_1(k_1, s_1, z_1)].$$

B.1 Characterization

Equity is a portfolio of two risky assets whose payoffs at time $t = 1$ and $t = 2$, respectively, are listed in Table B.1. We will refer to them as *current* and *continuation* asset, respectively.

For simplicity, denote the $t = 1$ conditional payoffs of the two assets as

$$\Gamma_{cu,1} \equiv y_1 + k_1(1 - \delta) \quad \text{and} \quad \Gamma_{co,1} \equiv -k_2 + E_1[M_2[y_2 + k_2(1 - \delta)]],$$

respectively. The sub-indexes *cu* and *co* are mnemonics for *current* and *continuation*. It follows that, with some abuse of notation, we can write the return on equity as

$$E_0[R_1] = \lambda(s_0, z_0) \frac{E_0[\Gamma_{cu,1}]}{E_0[M_1 \Gamma_{cu,1}]} + [1 - \lambda(s_0, z_0)] \frac{E_0[\Gamma_{co,1}]}{E_0[M_1 \Gamma_{co,1}]},$$

where $\lambda(s_0, z_0)$ is the loading on the current asset – the fraction of equity value accounted for by the current asset – or

$$\lambda(s_0, z_0) = \frac{E_0[M_1 \Gamma_{cu,1}]}{E_0[M_1 \Gamma_{cu,1}] + E_0[M_1 \Gamma_{co,1}]}.$$

Since the expected returns on either assets are independent of idiosyncratic productivity, the latter influences the expected return on equity only via its impact on the loading λ . Idiosyncratic productivity being mean-reverting, its expected growth rate is decreasing in s_0 . It follows that λ is strictly increasing in s_0 . The continuation asset accounts for a larger fraction of the value of small firms. These claims are formally stated in Lemma 1.

- Lemma 1** 1. *Equity is a portfolio consisting of current and continuation assets, which pay off exclusively at $t = 1$ and $t = 2$, respectively;*
2. *The excess return of neither asset depends on idiosyncratic productivity;*
3. *The current asset is itself a portfolio of the riskless asset and a risky asset with expected returns $\frac{E_0(e_{z,1}^\varepsilon)}{E_0(M_1 e_{z,1}^\varepsilon)}$. The loadings are both positive and are function of the risk-free rate.*
4. *The loading on the current asset is an increasing function of idiosyncratic productivity.*

Proof. The current asset is itself a portfolio of two assets. One is conditionally riskless, since it pays $k_1(1 - \delta)$ regardless of the state of nature. The other has a payoff $e^{s_1+z_1} k_1^\alpha$. The time-0 expected return of the latter is

$$\frac{E_0[e^{s_1+z_1} k_1^\alpha]}{E_0[M_1 e^{s_1+z_1} k_1^\alpha]} = \frac{E_0[e^{\varepsilon_{z,1}}]}{E_0[M_1 e^{\varepsilon_{z,1}}]}. \quad (6)$$

It follows that the expected return on the current asset is a weighted average of the conditional risk-free rate $R_{f,0}$ and (6), where the weight on the latter is

$$\frac{E_0[M_1[e^{s_1+z_1} k_1^\alpha]]}{E_0[M_1[e^{s_1+z_1} k_1^\alpha + k_1(1 - \delta)]]} = \frac{R_{f,0} - (1 - \delta)}{R_{f,0} - (1 - \delta)(1 - \alpha)}.$$

The weight on the short asset is

$$\frac{E_0[M_1[e^{s_1+z_1} k_1^\alpha + k_1(1 - \delta)]]}{E_0[M_1[e^{s_1+z_1} k_1^\alpha + k_1(1 - \delta)]] + E_0[M_1[-k_2 + E_1[M_2[y_2 + k_2(1 - \delta)]]]]}.$$

The weight will be increasing in s_0 as long as the following quantity is decreasing:

$$\frac{E_0[M_1[-k_2 + E_1[M_2[y_2 + k_2(1 - \delta)]]]]}{E_0[M_1[e^{s_1+z_1} k_1^\alpha + k_1(1 - \delta)]]}.$$

Tedious algebra reveals that the latter can be rewritten as

$$e^{s_0 \frac{\rho_s(\rho_s-1)}{1-\alpha}} \frac{[E[e^{\rho_s \varepsilon_s}]]^{1/(1-\alpha)} (1 - \alpha) E_0 \left[M_1 \left(\frac{\alpha}{1 - \frac{1-\delta}{R_{f,1}}} \right)^{\frac{\alpha}{1-\alpha}} [E_2[M_2 e^{z_2}]]^{\frac{1}{1-\alpha}} \right]}{[E_0(M_1 e^{z_1})]^{\frac{1}{1-\alpha}} \left[\left(\frac{\alpha}{1 - \frac{1-\delta}{R_{f,0}}} \right)^{\frac{\alpha}{1-\alpha}} + \frac{1-\delta}{R_{f,0}} \left(\frac{\alpha}{1 - \frac{1-\delta}{R_{f,0}}} \right)^{\frac{1}{1-\alpha}} \right]} ,$$

which is clearly decreasing in s_0 , as $\rho_s \in (0, 1)$.

■

In order to determine how expected equity returns vary with s_0 , we need to establish whether the current asset commands a greater or lower expected return than the continuation asset.²³ In order to answer this question, we make functional assumptions on the stochastic discount factor.

B.1.1 The Stochastic Discount Factor

Assume that $\log M_{t+1} \equiv \log \beta - \gamma \varepsilon_{z,t+1}$, where $\gamma > 0$ disciplines aversion to risk and $\beta > 0$ is the time discount factor. This choice implies that both the risk-free rate and the maximum Sharpe ratio are constant. For all $t \geq 0$,

$$R_{ft} = R_f = \frac{1}{E_t[M_{t+1}]} = \frac{1}{\beta} e^{-\frac{1}{2}\gamma^2\sigma_z^2} \quad \text{and} \quad \frac{\text{std}(M_{t+1})}{E(M_{t+1})} = \sqrt{e^{\gamma^2\sigma_z^2} - 1}.$$

It follows that the expected return on the risky portion of the current asset is

$$\frac{E_0(e_{z,1}^\varepsilon)}{E_0(M_1 e_{z,1}^\varepsilon)} = \frac{E_0(e_{z,1}^\varepsilon)}{E_0[e^{(1-\gamma)\varepsilon_{z,1}}]} = e^{\gamma\sigma_z^2} R_f,$$

while the expected return on the continuation asset equals

$$\frac{E_0[\Gamma_{co,1}]}{E_0[M_1 \Gamma_{co,1}]} = e^{\frac{1}{2}\gamma^2\sigma_z^2} \frac{E_0\left[e^{\frac{\rho_z}{1-\alpha}\varepsilon_{z,1}}\right]}{E_0\left[e^{\left(\frac{\rho_z}{1-\alpha}-\gamma\right)\varepsilon_{z,1}}\right] / R_f} = e^{\frac{\gamma\rho_z}{1-\alpha}\sigma_z^2} R_f.$$

B.2 Idiosyncratic productivity, size, and expected return on equity

With a constant risk-free rate, the only source of risk is the cash-flow volatility. As long as $\rho_z > 1 - \alpha$, the continuation asset will command a higher expected return than the current asset, and therefore equity returns will be monotonically decreasing in s_0 .

In the cross-section, idiosyncratic productivity is positively associated with both market size and expected returns on equity. This version of the size premium holds both conditionally and unconditionally, since the absence of capital adjustment costs makes installed capital k_0 immaterial for returns.

The parametric condition $\rho_z > 1 - \alpha$ is rather intuitive. As of $t = 0$, the risk of the continuation asset is pinned down by the covariance between time-1 innovations to aggregate productivity ($\varepsilon_{z,1}$) and $\Gamma_{co,1}$, the time-1 conditional expectation of the asset's

²³Note that this is not equivalent to assessing the slope of the equity term structure, as the definition of assets we are using is not standard.

payoff. Such moment is greater, the higher the autocorrelation of the process ρ_z and the higher the returns to scale in production.

Returns to scale are relevant, because they shape the elasticity of the capital choice k_2 to time-1 productivity innovations. In the remainder, we focus on the scenario for $\rho_z > 1 - \alpha$, as it is the empirically relevant one.

We conclude by adding that the impact of modelling interest rate risk would depend on the sign of the covariance between interest rate and the innovation in the aggregate productivity shock. A countercyclical risk-free rate would magnify the risk of the continuation asset. Conversely, a pro-cyclical risk-free rate would lower it.

B.2.1 Book-to-Market, Investment Rate, and Equity Returns

A corollary of the results obtained above is that information on indicators such as book-to-market ratio and investment rate cannot improve upon our characterization of the cross-section of returns. Because of the key role they play in the rest of the paper, however, we find it relevant to characterize the model-implied correlations between the two and expected returns.

Both book-to-market ratio and the investment rate depend on the installed capital k_0 . In order to compute the cross-sectional distribution of both quantities at time $t = 0$, we need to make assumptions about the distribution of k_0 .

We posit that k_0 was chosen optimally by each firm at time $t = -1$, under the assumptions that $z_{-1} = z_0 = 0$ and $s_{-1} \sim N\left(0, \frac{\sigma_s^2}{1-\rho_s^2}\right)$. In other words, we consider the scenario in which the aggregate productivity realization was equal to its unconditional mean in both $t = -1$ and $t = 0$, and that the cross-sectional distribution of idiosyncratic productivity was equal to the unconditional distribution. This is the scenario that more closely resembles the unconditional average of a full-fledged stationary model such as the one analyzed in the main body of the paper.

Under these assumptions, the average growth rate of capital installed by firms experiencing a realization of idiosyncratic productivity s_0 is

$$E\left[\log\left(\frac{k_1}{k_0}\right) | s_0\right] = \frac{\rho_s(1-\rho_s)}{1-\alpha} s_0.$$

This result follows immediately from Lemma 2. In the cross-section, the investment rate is increasing in s_0 .

Lemma 2 *Let $s_{t-1} \sim N\left(0, \frac{\sigma^2}{1-\rho^2}\right)$ and $s_t = \rho s_{t-1} + \varepsilon$, with $\varepsilon \sim N(0, \sigma^2)$, $\sigma > 0$ and $\rho \in (0, 1)$. Then, $E[s_{t-1} | s_t] = \rho s_t$.*

Proof. For simplicity, let f denote the density of a Normal distribution with parameters $(0, \frac{\sigma^2}{1-\rho^2})$. Let also g denote the density of a Normal distribution with parameters $(0, \sigma^2)$. It follows that

$$E[s_{t-1}|s_t] = \int \frac{s_{t-1}f(s_{t-1})g(s_t - \rho s_{t-1})ds_{t-1}}{\int f(s_{t-1})g(s_t - \rho s_{t-1})ds_{t-1}}.$$

To simplify notation further, let $\eta^2 \equiv \frac{\sigma^2}{1-\rho^2}$. Then,

$$f(s_{t-1})g(s_t - \rho s_{t-1}) = \frac{1}{2\pi\sigma\eta} \exp \left[-\frac{1}{2} \left(\frac{s_{t-1}^2}{\eta^2} + \frac{(\rho s_{t-1} - s_t)^2}{\sigma^2} \right) \right].$$

Algebraic manipulations yield

$$\begin{aligned} f(s_{t-1})g(s_t - \rho s_{t-1}) &= \frac{1}{2\pi\sigma\eta} \exp \left(-\frac{1}{2} \frac{(s_{t-1} - \rho s_t)^2}{\sigma^2} \right) \exp \left(-\frac{1}{2} \frac{s_t^2}{\sigma^2 + \eta^2 \rho^2} \right) \\ &= \frac{1}{\sqrt{2\pi}\eta} \exp \left(-\frac{1}{2} \frac{s_t^2}{\sigma^2 + \eta^2 \rho^2} \right) \times \frac{1}{\sqrt{2\pi}\sigma} \exp \left(-\frac{1}{2} \frac{(s_{t-1} - \rho s_t)^2}{\sigma^2} \right). \end{aligned}$$

The latter expression is the product of a constant and the density of a Normal with mean ρs_t and variance σ^2 . It follows that

$$E[s_{t-1}|s_t] = \int s_{t-1} \frac{1}{\sqrt{2\pi}\sigma} \exp \left[-\frac{1}{2} \frac{(s_{t-1} - \rho s_t)^2}{\sigma^2} \right] ds_{t-1} = \rho s_t.$$

■

Finally, we want to understand how the book-to-market ratio, i.e.

$$\frac{E(k_0|s_0)}{E_0[M_1\Gamma_{cu,1}] + E_0[M_1\Gamma_{co,1}]},$$

varies with s_0 . Above we have established that both numerator and denominator are increasing in s_0 . In Lemma 3 we prove that the denominator grows faster.

Lemma 3 *Assume that the discount factor is $M_{t+1} = \beta e^{\gamma \varepsilon_{t+1}}$ and that $s_{-1} \sim N(0, \frac{\sigma_z^2}{1-\rho_s^2})$. Along the path for the aggregate shock $z_{-1} = z_0 = 0$,*

1. *Size and investment rate are increasing in s_0 ;*
2. *Book-to-market is decreasing in s_0 .*

Proof. We limit ourselves to show that the book-to-market is decreasing in s_0 . Rewrite it as

$$\begin{aligned} BM &= \frac{E(k_0|s_0)}{k_1} \frac{k_1}{E_0[M_1[e^{s_1+z_1}k_1^\alpha + k_1(1-\delta)]] + E_0[M_1[-k_2 + E_1[M_2[y_2 + k_2(1-\delta)]]]]} \\ &= \frac{E(k_0|s_0)}{k_1} \widetilde{BM}. \end{aligned}$$

We have that

$$\log \widetilde{BM} = -\log \left(\frac{E_0[M_1[e^{s_1+z_1}k_1^\alpha + k_1(1-\delta)]]}{k_1} + \frac{E_0[M_1[-k_2 + E_1[M_2[y_2 + k_2(1-\delta)]]]]}{k_1} \right).$$

Since

$$\frac{E_0[M_1[e^{s_1+z_1}k_1^\alpha + k_1(1-\delta)]]}{k_1} = \left[\frac{\alpha}{1 - \frac{1-\delta}{R_f}} \right]^{-1} + \frac{1-\delta}{R_f},$$

the first addendum in parenthesis does not depend on s_0 . It follows that

$$\frac{\partial \log(\widetilde{BM})}{\partial s_0} = \frac{\rho_s(1-\rho_s)}{1-\alpha} \Omega_0,$$

where

$$\Omega_0 = \frac{E_0[M_1[-k_2 + E_1[M_2[y_2 + k_2(1-\delta)]]]]}{E_0[M_1[e^{s_1+z_1}k_1^\alpha + k_1(1-\delta)]] + E_0[M_1[-k_2 + E_1[M_2[y_2 + k_2(1-\delta)]]]]}.$$

By Lemma 2, $s_{-1}|s_0$ is normally distributed with mean $\rho_s s_0$ and variance σ_s^2 . It follows that

$$\begin{aligned} \log \left[\frac{E(k_0|s_0)}{k_1} \right] &= \log \left[E \left(e^{\frac{\rho_s s_{-1}}{1-\alpha}} | s_0 \right) \right] - \frac{\rho_s s_0}{1-\alpha} \\ &= \frac{\rho_s(\rho_s - 1)}{1-\alpha} s_0 + \frac{1}{2} \left(\frac{\rho_s \sigma_s}{1-\alpha} \right)^2. \end{aligned}$$

Finally,

$$\frac{\partial \log(BM)}{\partial s_0} = \frac{\rho_s(1-\rho_s)}{1-\alpha} [\Omega_0 - 1] < 0.$$

■

Expected returns exhibit a negative cross-sectional correlation with the investment rate and a positive correlation with book-to-market. The latter is the value discount result that holds true also in Section 3. Value firms earn higher returns because on average they have low market size!

B.3 Operating leverage

Now assume that at $t = 1$ firms incur a fixed operating cost $c_f > 0$. The expected return on equity at $t = 0$ becomes

$$E_0[R_1] = \frac{E_0[V_1(k_1, s_1, z_1)] - c_f}{E_0[M_1 V_1(k_1, s_1, z_1)] - c_f/R_f}.$$

Equity is now the combination of a short position on the risk-free asset and a long position on the current and continuation assets introduced above. The position on the risk-free asset is

$$-\frac{c_f/R_f}{E_0[M_1\Gamma_{cu,1}] - \frac{c_f}{R_f} + E_0[M_1\Gamma_{co,1}]} < 0,$$

which is strictly decreasing in c_f .

Everything else equal, raising the fixed cost is equivalent to expanding the short position – i.e. increasing leverage. It follows that the expected return on equity is increasing in c_f .

A greater level of s_0 is equivalent to a decline in leverage, i.e. a smaller short position on the risk-free asset and a smaller long position on the portfolio of current and continuation assets. It follows that the expected equity return still falls with s_0 . This property is stated formally in Lemma 4.

Lemma 4 *For $c_f > 0$, as long as $\rho_z > 1 - \alpha$, the expected return on equity is still monotonically decreasing in the level of idiosyncratic productivity s_0 .*

Proof. Think of equity as a portfolio including the risk-free-asset as well as a composite of current and continuation business assets. The weight of the risk-free asset $-\frac{c_f/R_f}{E_0[M_1y_1] + E_0[M_1[-k_2 + E_1[M_2y_2]]] - c_f/R_f}$ is negative and clearly increasing in s_0 . That is, the short position on the risk-free asset declines with s_0 . It follows that the long position on the composite of current and continuation assets also declines. Then the result follows from Lemma 1, which ensures that the return on the composite portfolio declines with s_0 . ■

The fixed cost also affects the book-to-market ratio, which becomes

$$\frac{E(k_0|s_0)}{E_0[M_1\Gamma_{cu,1}] + E_0[M_1\Gamma_{co,1}] - c_f/R_f}.$$

The slope of the mapping between s_0 and book-to-market, which was proven to be positive for $c_f = 0$, decreases with c_f . For a large enough value, it will turn negative and deliver an unconditional version of the value premium. On average, high book-to-market firms will be characterized by low productivity, small size, and high equity returns.

C The role of investment adjustment costs in Zhang (2005)

Consisting of a minor variation of the neoclassical investment model, our framework is also very close to that considered in Zhang (2005). Yet, the message of our paper is very different from the one emerging from his work. Zhang claims in his abstract that “*The*

value anomaly arises naturally in the neoclassical model with rational expectations.” In this appendix, we demonstrate that the contradiction between the two messages is only apparent.

We show that when the capital adjustment cost is parameterized in such a way that the volatility of the investment rate is consistent with the evidence, Zhang’s model generates a value discount. Raising the adjustment cost to the levels considered in Zhang (2005) leads to a sizeable value premium, but at the cost of essentially no cross-sectional variation in investment rates. We also show that, at the margin, asymmetry in investment adjustment costs plays no significant role.

We consider a version of Zhang’s economy with counter-cyclical price of risk and symmetric investment adjustment costs. All parameters are set equal to the benchmark values of the original paper.²⁴ We compute allocations for levels of the parameter ϕ_1 , governing the marginal cost of investment, as low as 0 and as high as 15. The latter is the value entertained in Zhang (2005).

The results – expressed at the monthly frequency to ensure consistency with Zhang (2005) – are gathered in Table C.1. The message is the same as that emerging from Section 3.4. As long as adjustment costs are moderate enough to produce an empirically sensible standard deviation of the investment rate (about 0.05 at the monthly frequency, according to our estimation), the model cannot produce a value premium. When ϕ_1 is raised to 15, however, the annualized value premium rises to 4.38%.

A large value premium obtains because the capital stock is essentially fixed. All cross-sectional variation in book-to-market is due to the variation in idiosyncratic productivity. Value firms identify with small-cap firms. Growth firms, on the other hand, are large-cap firms.

In a second exercise, we explore the role played by the asymmetry in investment adjustment cost, which Zhang (2005) identifies as necessary condition for obtaining a sizeable value premium. Table C.2 compares key implications of the model of which in the bottom row of Table C.1 – i.e. symmetric adjustment cost with $\phi_1 = 15$ – with those of a version where, exactly as in the benchmark scenario in Zhang (2005), the marginal cost of negative investment is 10 times larger than the cost of positive investment (which is kept at 15).

²⁴The only minor difference is that we do not allow the price of the final good to vary. We set the price equal to the unconditional mean of its distribution in Zhang (2005). Our own calculation show that the price variation has essentially no role in shaping the cross-section of returns.

Table C.1: Comparative Statics w.r.t. ϕ_1 – Zhang's Economy – Monthly frequency

	Investment Moments		Returns (%)			Investment Rate			Capital			Idiosyncratic Shock		
ϕ_1	mean ik	std ik	Growth	Value	Δ	Growth	Value	Δ	Growth	Value	Δ	Growth	Value	Δ
0	0.039	0.150	1.310	0.968	-0.342	0.177	-0.047	-0.225	2.621	10.677	8.056	-0.426	0.377	0.803
0.0005	0.031	0.125	1.311	0.958	-0.353	0.142	-0.048	-0.190	2.614	11.122	8.509	-0.429	0.415	0.845
0.005	0.020	0.082	1.307	0.959	-0.348	0.081	-0.035	-0.116	2.533	10.236	7.703	-0.434	0.392	0.826
0.05	0.013	0.043	1.227	0.988	-0.239	0.049	-0.020	-0.069	2.547	8.201	5.654	-0.337	0.301	0.638
0.5	0.011	0.019	1.074	1.101	0.027	0.034	-0.009	-0.043	2.581	4.851	2.271	-0.029	0.001	0.030
5	0.010	0.008	0.940	1.209	0.269	0.020	0.002	-0.018	2.120	2.217	0.097	0.421	-0.398	-0.819
15	0.010	0.005	0.911	1.269	0.358	0.016	0.006	-0.011	1.555	1.488	-0.067	0.530	-0.495	-1.025

Table C.2: Asymmetric Adjustment Costs

Moments	No Asymmetric Cost	Asymmetric Cost
Mean Investment Rate	0.0100	0.0100
Std. Dev. Investment Rate	0.0055	0.0043
Autocorrelation Investment Rate	0.7959	0.7278
Growth Firms Return (%)	0.9111	0.9064
Value Firms Return (%)	1.2687	1.2834
Value Premium (%)	0.3576	0.3770

Adding a large degree of asymmetry barely changes the key cross-sectional moments. In particular, the impact on the value premium is a negligible 0.02% per month. The reason is that asymmetry is added to a model – with $\phi_1 = 15$ for both positive and negative investment – that already displays no variation in the capital stock.