Misallocation or Risk-Adjusted Capital Allocation?

Joel M. David* Lukas Schmid† David Zeke‡§

USC Duke USC

November 18, 2017

Abstract

Standard, frictionless neoclassical theory of investment predicts that the expected corporate marginal product of capital (MPK) depends on firms' exposure to systematic risk and the price of that risk. This implies that the cross-sectional dispersion in MPK i) depends on cross-sectional variation in risk exposures and ii) fluctuates with the price of risk, and thus is countercyclical. We empirically evaluate these predictions and document strong support for them. In particular, a long-short portfolio of high minus low MPK stocks earns significant and countercyclical excess returns forecastable by standard return predictors. A calibrated investment model suggests that ex ante variation in risk exposure can rationalize permanent dispersion in MPK. These findings suggests that a substantial fraction of dispersion in MPK, often dubbed misallocation, is effectively risk adjusted capital allocation.

*joeldavi@usc.edu.

†lukas.schmid@duke.edu.

‡zeke@usc.edu.

§We thank Leonid Kogan, Deborah Lucas, Ellen McGrattan, Christian Opp, Amir Yaron and seminar participants at McGill, Toronto, USC, Wharton, and at NBER Summer Institute Capital Markets and the Carnegie Mellon/LAEF Macrofinance Conference for comments and discussions.

1 Introduction

A large and growing body of work has documented the 'misallocation' of resources across firms, measured by dispersion in the marginal product of factors of production. The failure of marginal product equalization has been shown to have potentially sizable negative effects on aggregate outcomes, such as productivity and output. Recent studies have found that even after accounting for a host of leading candidates as sources of misallocation - for example, adjustment costs, financial frictions, or imperfect information - a large role is played by firm-level 'distortions,' specifically, of a class that are orthogonal to firm fundamentals and are permanent to the firm. Identifying exactly what - if any - underlying economic mechanisms can lead to this type of distortion has proven puzzling.

In this paper, we propose, empirically test and quantitatively evaluate just such a mechanism. Our approach links capital misallocation to systematic investment risks. To the best of our knowledge, we are the first to make the connection between standard notions of the risk-return tradeoff faced by investors and the resulting dispersion in the marginal product of capital across firms. Our point of departure is a standard model of firm investment in the face of both aggregate and idiosyncratic shocks. Firms discount future payoffs using a stochastic discount factor that is also a function of aggregate conditions. With little more structure than this, the framework gives rise to an asset pricing equation governing the firm's expected marginal product of capital (MPK): firms with higher exposure to the aggregate shock have a permanently higher expected MPK, which appears exactly as what would otherwise be labeled a permanent firm-level distortion. Importantly, this is a statement only about expected MPK; realized MPKs may differ across firms for additional reasons, i.e., uncertainty over future shocks. In fact, the model implies a beta pricing equation of exactly the same form often used to price the cross-section of stock market returns. That equation simply states that a firm's expected MPK should be linked to the exposure of its MPK to systematic risk, and the latter's price.

Although this result is quite general, we provide a number of illustrative examples. If we abstract from aggregate shocks (or assume risk neutrality), expected MPKs are equated across firms. In this case the standard notion obtains that relates any dispersion in expected MPK to misallocation. In any other situation, our beta pricing equation implies that cross-sectional dispersion in risk exposure should be reflected in cross-sectional dispersion in expected MPK. If the utility function is CRRA, for example, expected MPKs are determined by the Consumption CAPM equation, i.e., by the covariance of each firm's MPK with aggregate consumption growth. If aggregate and firm-level conditions are driven by technology shocks, expected MPKs are determined by the covariance of the MPK with innovations in the aggregate process. In a world driven by multiple risk factors, as is typically considered in the current asset pricing literature, the average MPK is linked to exposure to these various factors, as well as the factor prices.

Much of our analysis is devoted to demonstrating that the simple beta pricing equation inherent in the neoclassical model of firm investment has substantial empirical content. More precisely, we state and empirically investigate four predictions of our general framework.

First, the beta pricing equation predicts that exposure to standard risk factors priced in asset markets is an important determinant of expected MPK. We provide empirical evidence supporting this prediction in two ways. First, we directly determine the exposure of firm and portfolio level MPK to various risk factors emerging in current asset pricing models, from consumption growth, the market portfolio, the Fama-French factors, and most recently, the Q-factors, and show that MPKs are indeed significantly related to these factors. Second, given a variety of empirical challenges to directly computing MPK betas at the firm level, we exploit the tight relationship between MPK and investment returns on the one hand, and stock returns on the other hand, as pointed out, for example, by Cochrane (1991) and Restoy and Rockinger (1994), and use an asset pricing approach for testing. Specifically, we sort stocks into decile portfolios based on their MPK and evaluate the portfolio returns. We find that the expected excess returns on these decile portfolios are generally increasing in MPK, so that a high minus low (HML) MPK portfolio earns an annual premium of anywhere from 3% to 6%. We show that the high MPK portfolios have higher exposure to standard risk factors, consistent with the prediction of the beta pricing equation, suggesting that a higher MPK is linked to higher systematic risk.

Consistent with standard asset pricing models, a beta pricing equation does not only entail a cross-sectional prediction regarding cross-sectional variation in expected MPK, but also a time-series prediction linked to movements in factor risk prices. In particular, it suggests that movements in factor risk prices are linked to fluctuations in the conditional expected MPK. Again, we test this prediction on both direct estimates of expected MPK as well as MPK sorted stock portfolios. Consistent with the empirical evidence from the return predictability literature in asset pricing, suggesting a countercyclical price of risk, we indeed find the excess returns on MPK sorted portfolios are predictable and countercyclical, as suggested both by common return predictors such as credit spreads and the price/dividend ratio.

These tests suggest that risk factors are indeed a significant determinant of MPK, both in the cross-section and in the time series. Our next predictions and tests aim at further dissecting the role of risk factors in determining the cross-sectional dispersion in MPK, which has commonly been associated with capital 'misallocation'. In that regard, the beta pricing equation suggests that MPK dispersion should be positively related to beta dispersion. In particular, in the cross-section, industries with higher dispersion in betas should display higher dispersion in MPK. We test this prediction in a two stage procedure that first determines betas with respect to standard risk factors, and then uses the dispersion of the betas as an explanatory variable for industry level MPK dispersion. Consistent with the neoclassical investment model, we find that beta dispersion is a significant determinant of MPK dispersion, explaining a substantial fraction of its inherent variation.

The prediction regarding determinants of MPK dispersion has a natural time series analog, in that movements in factor prices should be linked to fluctuations in MPK dispersion. In particular, given likely countercyclical risk prices, a frictionless neoclassical model predicts a countercyclical dispersion in MPK. We test this notion by evaluating the time series properties of the MPK-HML portfolio, and find that its positive expected excess returns are highly predictable, and in fact countercyclical, as indicated by standard return and macroeconomic predictors such as credit spreads, excess bond premia, and the price/dividend ratio. This suggests that not only do high MPK firms earn higher risk premia, but also that the spread in the implied cost of capital between high and low MPK firms rises in downturns. In other words, high MPK firms, while potentially more productive, become riskier in recessions.

After establishing these empirical results, we interpret them and gauge their magnitudes through the lens of a quantitative model. To that end, we calibrate and structurally estimate a simple neoclassical, dynamic investment model with adjustment costs that we augment with an exogenously specified stochastic discount factor designed to match standard asset pricing moments, as has become standard in the cross-sectional asset pricing literature, as in Zhang (2005) and Gomes and Schmid (2010). Our point of departure from these model is that we allow for ex ante cross-sectional heterogeneity in exposure, that is, beta, with respect to a single source of risk. This extension allows us to gauge what fraction of dispersion in MPK can be rationalized by a realistic amount of permanent ex ante heterogeneity in betas, with or without additional frictions, such as convex adjustements costs. Our preliminary results suggest that in a simple dynamic model such ex ante heterogeneity can rationalize more than fifty percent of the empirical dispersion, the rest being accounted for by other frictions, such as adjustment costs or perhaps financial frictions.

Related Literature. Our paper relates to several branches of the literature. First is the large body of work investigating and quantifying the effects of resource misallocation across firms, seminal examples of which include Hsieh and Klenow (2009) and Restuccia and Rogerson (2008). A number of papers have explored the role of particular economic forces in leading to misallocation. For example, Asker et al. (2014) study the role of capital adjustment costs, Midrigan and Xu (2014), Moll (2014), Buera et al. (2011) and Gopinath et al. (2015) financial frictions, and David et al. (2016) information frictions. David and Venkateswaran (2016) provide a unified theoretical framework and empirical methodology to estimate the contribution of each of these forces to misallocation and find that they can explain only a limited portion of observed dispersion in the marginal product of capital. These authors conclude that firm-specific distortions account for the lion's share of misallocation and, specifically, point out the large role of a permanent component of those distortions. We build on this literature by exploring the implications of a different dimension of financial markets for marginal product dispersion, namely, the risk-return tradeoff faced by risk-averse investors. Importantly, our theory generates what appears to be a permanent firm-specific 'wedge' exactly of the type found by David and Venkateswaran (2016), but which in our framework is a function of each firm's exposure to aggregate risk. The addition of aggregate risk is a key innovation of our analysis - existing work has abstracted from this channel. We show that the link between aggregate risk and misallocation is quite tight in the presence of heterogeneous exposures to that risk.

A growing literature, starting with Eisfeldt and Rampini (2006), investigates the reasons underlying the observation that capital reallocation is procyclical. This indeed seems puzzling as given higher cross-sectional dispersion in MPK in downturns one should expect to see capital flowing to highly productive, high MPK firms in recessions. Our results bear on that observation by noting that given a countercyclical price of risk, and a countercyclical premium on the MPK-HML portfolio, from a risk perspective, capital reallocation to high MPK firms would require capital flow to the riskiest of firms.

In a related effort, Binsbergen and Opp (2017) also investigate the implications of asset market data for the real economic decisions of firms. While they focus on the implications of mispricing in the pricing of financial assets for corporate decisions, we focus on misallocation on the real side. While we investigate the implications of cross-sectional dispersion in expected returns, we remain agnostic about whether that dispersion comes from mispricing or differential exposure to risk.

Our work exploits the insight, due to Cochrane (1991) and Restoy and Rockinger (1994), that stock returns and investment returns are closely linked. Indeed, under the assumption of constant returns to scale, stock and investment returns effectively coincide. Crucially, for our purposes, investment returns are intimately linked with the marginal product of capital. Balvers et al. (2015) explore and confirm the close albeit more complicated relationship under deviations from constant returns to scale. In this context, our work is closely related to the growing literature that examines the cross-section of stock returns by viewing them from the perspective of investment returns, starting from Gomes et al. (2006); Liu et al. (2009), and recently forcefully summarized in Zhang (2017). This literature interprets common risk factors as the Fama-French factors through firms' investment policies, and most recently, shows that risk factors related to corporate investment patterns themselves capture risks priced in the cross-section of returns, culminating in the recent Q-factor model. Our objective is quite different and in some sense turns that logic on its head, in that we examine investment returns and the marginal product of capital as a manifestation of exposure to systematic risk, most readily measured through stock returns.

2 Motivation

We consider a discrete time, infinite-horizon economy. A continuum of firms, indexed by i, produce output using capital and labor. Labor is chosen period-by-period in a spot market at a competitive wage. At the end of each period, firms choose investment in new capital, which becomes available for production in the following period so that $K_{it+1} = I_{it} + (1-\delta) K_{it}$, where δ is the rate of depreciation.

Let $\Pi_{it} = Y(X_t, Z_{it}, K_{it})$ denote the operating profits of the firm - revenues net of labor costs - where X_t and Z_{it} denote aggregate and idiosyncratic shocks, respectively, and K_{it} the firm's level of capital. The analysis can accommodate a number of interpretations of these shocks, for example, as productivity or demand shifters. We assume the profit function takes a Cobb-Douglas form and is homogeneous in K of degree $\theta < 1$. This structure follows, for example, when the production function is Cobb-Douglas in capital and labor and firms face CES demand curves. In this case, the parameter θ captures both the curvature in production and demand, as well as the relative shares of capital and labor in production, and operating profits are proportional to revenues. The marginal product of capital is equal to $\theta \frac{\Pi_{it}}{K_{it}}$.[1] The payout of the firm in period t is equal to $D_{it} = \Pi_{it} - I_{it}$.

Firms discount future cash flows using a stochastic discount factor (SDF) M_{t+1}, which may be correlated with the aggregate component of firm fundamentals, i.e., with X_t. We can write the firm's problem in recursive form as

$$V(X_t, Z_{it}, K_{it}) = \max_{K_{it+1}} \Pi_i(X_t, Z_{it}, K_{it}) - K_{it+1} + (1-\delta) K_{it} + \mathbb{E}_t\left[M_{t+1} V(X_{t+1}, Z_{it+1}, K_{it+1})\right]$$

where $\mathbb{E}_t[\cdot]$ denotes the firm's expectations conditional on time t information. Standard techniques give the Euler equation

$$1 = \mathbb{E}_t\left[M_{t+1}\left(MPK_{it+1} + 1 - \delta\right)\right] \quad \forall\, i, t \tag{1}$$

[1]For example, if the production function is given by $Y_{it} = F(X_z, Z_{it}) K_{it}^{\theta_1} N_{it}^{\theta_2}$ and the demand function is given by $Y_{it} = P_{it}^{-\eta}$, the curvature parameter θ is equal to $\frac{\theta_1\left(1-\frac{1}{\eta}\right)}{1-\theta_2\left(1-\frac{1}{\eta}\right)}$.

where $MPK_{it+1} = \frac{\partial \Pi_{it+1}}{\partial K_{it+1}}$ is the marginal product of capital of firm i at time $t+1$.

MPK dispersion. An immediate consequence of expression (1) is that expected MPK need not be equated across firms; rather, it is only appropriately discounted expected MPK that is equalized. To the extent that firms load differently on the discount factor, their expected MPKs will differ. Assuming a single source of aggregate risk for the sake of illustration, Appendix A derives the following factor model for expected MPK:

$$\mathbb{E}_t\left[MPK_{it+1}\right] = \alpha_t + \beta_{it}\lambda_t \tag{2}$$

Here, α_t is the 'risk-free' MPK, which equals the riskless user cost of capital $r_{ft} + \delta$ where r_{ft} is the net risk-free rate; $\beta_{it} \equiv -\frac{\text{cov}(M_{t+1}, MPK_{t+1})}{\text{var}(M_{t+1})}$ measures the exposure, or loading, of the firm's MPK on the SDF, i.e., the riskiness of the firm; and $\lambda_t \equiv \frac{\text{var}(M_{t+1})}{\mathbb{E}_t[M_{t+1}]}$ is the market price of that risk. In the language of asset pricing, the Euler equation gives rise to a conditional one-factor model for expected MPK. Expression (2) highlights that expected MPK is not necessarily common across firms and is a function of the risk-free return, the firm's β on the SDF, which may vary across firms, and the market price of risk. The cross-sectional variance of date-t conditional expected MPK is then equal to

$$\sigma^2_{\mathbb{E}_t[MPK_{it+1}]} = \sigma^2_{\beta_t}\lambda_t^2 \tag{3}$$

which shows that the extent to which risk considerations lead to dispersion in the MPK depends on (1) the cross-sectional dispersion in firm-level β's at date t and (2) the level of the price of risk. Taking unconditional expectations, the theory can clearly generate persistent dispersion in MPK, which is equal to the dispersion in required rates of return across firms:

$$\mathbb{E}\left[MPK_{it}\right] = \alpha + \beta_i\lambda + \text{cov}(\beta_{it}, \lambda_t) \tag{4}$$

where $\alpha = \mathbb{E}\left[r_{ft} + \delta\right]$, $\beta_i = \mathbb{E}\left[\beta_{it}\right]$ and $\lambda = \mathbb{E}\left[\lambda_t\right]$ denote the unconditional expectations of the risk-free MPK, conditonal MPK factor betas and factor prices, respectively. So long as the relationship between mean β's and the time-series correlation of those β's with the price of risk is weak, we can write the variance of mean MPK approximately as[2]

$$\sigma^2_{\mathbb{E}[MPK]} \approx \sigma^2_{\beta}\lambda^2 \tag{5}$$

[2]In line with the results in Lewellen and Nagel (2006), we find the time-series variation in β's to be quite small, suggesting the validity of the approximation. In the case of that the β of a firm is constant, for example, which we assume in our quantitative model, the expression is exact.

where σ_β^2 denotes the cross-sectional variance of unconditional expected β's. We note that this observation generalizes in a straightforward manner to environments more recently considered in the cross-sectional asset pricing literature emphasizing the presence of multiple aggregate risk factors. Most prominently, beyond excess returns on the market portfolio and innovations to aggregate consumption growth as considered in the classical CAPM and Breeden-Lucas Consumption CAPM, these risk factors have been linked to excess returns on size, as well as book-to-market sorted portfolios (Fama-French factors), or investment returns or profitability (the Q-factor model of Hou et al. (2015) and Zhang (2017)).

The strength of the mechanism linking persistent dispersion in MPK to exposure to aggregate risk can be understood by inspection of expression (5) - predicted MPK dispersion is increasing in the dispersion in β's and also in the market price of risk, λ. A key observation underlying our analysis is that asset pricing data suggest that risk prices are rather high. A lower bound is given by the Sharpe ratio on the market portfolio, estimated to be around 0.4. However, even easily implementable trading strategies such as those based on value-growth portfolios, or momentum, suggest numbers closer to 0.8, while hedge fund strategies report Sharpe ratios in excess of one. Taken at face value, these numbers suggest the possibility for substantial MPK dispersion - even in frictionless models - after taking risk exposure in account. Our quantitative work in Section 4 quantifies this link using data on risk prices and cross-sectional variation in expected stock market returns.

Empirical Predictions. Even under the general structure we have outlined thus far, the theory has a good deal of empirical content. Specifically, the expressions laid out above contain a number of both (1) cross-sectional and (2) time-series predictions:

1. Exposure to standard risk factors is a determinant of expected MPK. Expression (2) shows that the same factors that determine the cross-section of stock returns - namely, exposure to the SDF - determine the cross-section of MPK. Firms with a higher loading on the SDF, i.e., higher β's, should have higher conditional expected MPK.

2. Predictable variation in the price of risk, λ_t, leads to predictable variation in mean expected MPK. In particular, the mean conditional expected MPK should increase when the price of risk does. This is the time-series equivalent of expression (2) - holding fixed the distribution of β's, movements in λ_t should positively affect the mean expected MPK. Since the price of risk is known to be countercyclical, this implies that the mean expected MPK is as well.

3. MPK dispersion is related to β dispersion. Expression (5) shows that unconditional variation in the cross-section of MPK is proportional to the variation in β. Segments of the economy, for example, industries, with higher dispersion in β should display higher dispersion in MPK.

4. MPK dispersion is positively correlated with the price of risk. Expression (3) also has a time-series prediction linking MPK dispersion to time variation in the price of risk. For a given degree of cross-sectional dispersion in β, when required compensation for bearing risk increases, MPK dispersion should increase as well.

Illustrative examples. Section 3 investigates each of these predictions in detail. Before doing so, however, it is useful to consider a number of more concrete illustrative examples (derivations for this section are in Appendix A).

Example 1: no aggregate risk (or risk neutrality). In the case of no aggregate risk, we have $\beta_{it} = 0 \; \forall \; i, t$, i.e., all shocks are idiosyncratic to the firm. Expressions (2) and (3) show that there will be no dispersion in expected MPK and for each firm, $\mathbb{E}_t\left[MPK_{it+1}\right] = r_f + \delta$, which is simply the riskless user cost of capital (which is constant in the absence of aggregate shocks). This is the standard result from the stationary models widely used in the misallocation literature where without frictions, expected MPK should be equalized across firms.[3] It is straightforward to show this expression also holds in an environment with aggregate shocks but risk neutral preferences, which implies M_{t+1} is simply a constant (equal to the time discount factor).

Example 2: CAPM. In the CAPM, the SDF is linearly related to the market return, i.e., $M_{t+1} = a - bR_{m,t+t}$ for some constants a and b. Because the market portfolio is itself an asset with a β of one, it is straightforward to derive

$$\mathbb{E}_t\left[MPK_{it+1}\right] = \alpha_t + \underbrace{\frac{\text{cov}\left(R_{m,t+1}, MPK_{it+1}\right)}{\text{var}\left(R_{m,t+1}\right)}}_{\beta_{it}} \underbrace{\mathbb{E}_t\left[R_{mt+1} - R_{f,t+1}\right]}_{\lambda_t}$$

i.e., expected MPK is determined by the covariance of the firm's MPK with the market return (i.e., its market β), which is the the risk factor in this environment. The market price of risk is equal to the expected excess return on the market portfolio, i.e., the equity premium.

Example 3: CCAPM. In the case that the utility function is CRRA with coefficient of relative risk aversion γ, standard approximation techniques give the pricing equation from the

[3]With the time-to build for capital and uncertainty over upcoming shocks in our model, there may still be dispersion in *realized* MPK, but not in *expected* terms, and so these forces do not lead to persistent deviations from MPK equalization for a particular firm.

consumption capital asset pricing model:

$$\mathbb{E}_t\left[MPK_{it+1}\right] = \alpha_t + \underbrace{\frac{\text{cov}\left(\Delta c_{t+1}, MPK_{it+1}\right)}{\text{var}\left(\Delta c_{t+1}\right)}}_{\beta_{it}} \underbrace{\gamma \text{var}\left(\Delta c_{t+1}\right)}_{\lambda_t}$$

where Δc_{t+1} denotes log consumption growth. Expected MPK is determined by the covariance of the firm's MPK with consumption growth (its consumption β), which is now the risk factor. The market price of risk is the product of the coefficient of relative risk aversion and the variability of consumption growth.

In Section 4, we follow the recent literature on production-based asset pricing and explicitly model the sources of uncertainty as arising from technology shocks, both at the firm and aggregate level, and quantify the implications of those shocks for *mpk* dispersion.

3 Empirical Results

In this section we investigate the empirical predictions outlined in Section 2.

Data. Our data come primarily from the Center for Research in Security Prices (CRSP) and Compustat. We use data on nonfinancial firms with common equities listed on the NYSE, NASDAQ, or AMEX over the period 1962 to 2014. We supplement this panel with time-series data on market factors and aggregate conditions related to the market price of risk. The market factors we consider are the Fama and French (1992) factors, Hou, Xue, and Zhang (2015) investment-CAPM factors, as well as the growth rate of non-durable and services consumption from the Bureau of Economic Analysis (BEA). We also use data on aggregate macroeconomic and financial market variables from the BEA and the Gilchrist and Zakrajsek (2012) (GZ) credit spread.[4] We measure the firm's capital stock, K_{it}, as the (net of depreciation) value of property, plant and equipment (Compustat series PPENT) and firm revenue, Y_{it}, as reported sales (series SALE). Ignoring constant terms, which will play no role in our analysis, we measure the marginal product of capital (in logs) as $mpk_{it} = y_{it} - k_{it}$.[5]

We can now revisit the main predictions from Section 2.

1. Exposure to standard risk factors is a determinant of expected MPK. To investigate this implication of our framework, Table 1 assesses the relationship between MPK and both contemporaneous and future excess stock returns. We sort firms into 10 portfolios based on their

[4] We obtain measures of the GZ spread from Simon Gilchrist's website.

[5] Recall that in our setup, operating profits are proportional to revenues, making this a valid measure of the *mpk*.

year t MPK, where portfolio 1 contains low MPK firms and portfolio 10 high MPK ones. We then compute the contemporaneous and one-period ahead equal-weighted excess stock return to each portfolio. Following Fama and French (1992), we use the MPK reported by firms in their fiscal-year-end filing in date t-1 with firm returns from July of year t to June of year t+1 when computing future returns. We additionally compute excess returns on on a high-minus-low portfolio (MPK-HML), which is an annually rebalanced portfolio that is long on stocks in the highest MPK portfolio and short on stocks in the lowest.

Table 1: Excess Returns on MPK Sorted Portfolios

	Portfolio										
	Low	2	3	4	5	6	7	8	9	High	MPK-HML
					Panel A: Not Industry-Adjusted						
r^e_t	6.026*	9.288**	9.258**	10.26***	10.65***	12.21***	12.86***	14.57***	15.20***	17.69***	11.11***
	(1.68)	(2.44)	(2.42)	(2.72)	(2.86)	(3.13)	(3.11)	(3.36)	(3.39)	(3.74)	(4.05)
r^e_{t+1}	6.632*	10.48***	11.19***	12.68***	12.59***	13.25***	13.46***	13.01***	13.22***	13.58***	6.583**
	(1.87)	(2.83)	(2.99)	(3.44)	(3.45)	(3.36)	(3.30)	(3.12)	(3.03)	(3.00)	(2.46)
					Panel B: Industry-Adjusted						
r^e_t	8.909*	8.208*	9.408**	9.386**	10.06***	11.58***	11.05***	13.80***	16.03***	17.67***	8.870***
	(1.71)	(1.95)	(2.41)	(2.55)	(2.76)	(3.05)	(2.86)	(3.28)	(3.45)	(3.60)	(5.17)
r^e_{t+1}	10.11*	11.48***	10.88***	11.73***	11.15***	11.95***	11.39***	13.33***	13.08***	13.20***	3.286**
	(1.96)	(2.79)	(2.85)	(3.15)	(3.11)	(3.21)	(2.98)	(3.35)	(3.03)	(2.81)	(1.99)

Notes: $r^{\mathbb{E}}_t$ denotes equal-weighted contemporaneous annual excess stock returns (over the risk-free rate) measured in the year of the portfolio formation from January to December of year t. r^e_{t+1} denotes the analogous future returns, measured in the year following the portfolio formation, from July of year $t+1$ to June of year $t+2$. Industry adjustment is done by de-meaning returns by industry-year, where an industry is defined as a 4 digit SIC code. t-statistics in parentheses. Significance levels are denoted by: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1 reveals a strong relationship between MPK and stock returns - portfolios with higher MPK earn higher excess returns. Panel A shows that the difference in contemporaneous returns between high and low MPK firms, i.e., the excess return on the MPK-HML portfolio is about 11% annually and remains high, about 6.5%, for one-period ahead returns. Both contemporaneous and future spreads are statistically different from zero at the 95% level. Firms that offer high stock returns tend to also have MPKs, both in a realized and an expected sense.

The focus in the misallocation literature is generally on within-industry variation in the MPK. Panel B of Table 1 reports within-industry results, defined at the 4-digit SIC level. To compute these values, from each return observation we subtract the mean return within that industry-year.[6] Although the magnitudes fall somewhat, the relationship between MPK and stock returns remains strong even when comparing across firms within a particular industry, both in an economic and statistical sense - the MPK-HML contemporaneous excess return is almost 9% annually and the future excess return almost 3.5%. Both are statistically significant at the 95% level.

[6] We define an industry as a 4-digit SIC code and examine industry-year pairs with at least 10 observations.

2. Predictable variation in the price of risk λ_t leads to predictable variation in expected MPK. Expression (2) implies that the market price of risk, λ_t, is positively related to the level of expected MPK in the following period. To test this, we estimate regressions of firm *mpk* on three lagged (by one year) measures related to the price of risk: 1) the price/dividend ratio; 2) the Gilchrist and Zakrajsek (2012) (GZ) spread, a high-information and duration-adjusted measure of the mean credit spread; and 3) the Excess Bond Premium, which measures the portion of the GZ spread not attributable to default risk. We control for the changing composition of firms in the following way: using only those firms where our measure of *mpk* is observed for the firm in consecutive quarters, we compute changes in mean *mpk* for every pair of years. We then use those changes to construct a synthetic composition-adjusted mean *mpk* which is unaffected by new additions or deletions from the dataset. Table 2 reports the results of these regressions. In line with the theory, column (3) and (2) show that the GZ spread and the excess bond premium (which are likely positively correlated with the market price of risk) predict higher future *mpk*, while column (1) shows that the price-dividend ratio (likely negatively correlated with the market price of risk) predicts lower future *mpk*.

Table 2: Predictability of $\mathbb{E}_t\left[MPK_{it}\right]$

	(1)	(2)	(3)
PD Ratio	-0.0114***		
	(-4.61)		
GZ Spread		0.0452***	
		(3.01)	
Excess Bond Premium			0.0603**
			(2.59)
Constant	-0.00653	-0.0803***	-0.00595
	(-0.41)	(-2.73)	(-0.36)
Observations	148	148	148
R^2	0.142	0.094	0.056

Notes: Table reports time-series regressions of average composition-adjusted *mpk* on lagged (by one year) measures of the price of risk. Long-term trends in *mpk* and the price/dividend ratio are removed using a one-sided hp filter. t-statistics are in parentheses. t-statistics in parentheses, which are computed using Newey-West standard errors. Significance levels are denoted by: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All observations are observed at the quarterly frequency.

3. MPK dispersion is related to β dispersion. Expression (5) implies that for particular groups of firms, dispersion in expected *mpk* should be positively related to the dispersion in β. In particular, this suggests that dispersion of *mpk* within an industry, a common measure of misallocation, is positively correlated with dispersion in expected stock returns and β's. We investigate this prediction using variation in the dispersion of firm-level β's across industries.

For each industry in each year, we compute the standard deviation of mpk, $\sigma(mpk)$, expected returns, $\sigma(\mathbb{E}[ret])$ and β's, $\sigma(\beta)$ and estimate a pooled regression of industry-level mpk dispersion on the dispersion in stock returns and β's. To avoid potential biases from the realization of shocks, we lag the independent variables (dispersion in expected stock returns and β's) by a year. We detail our computation of firm-level measures of β and excess returns in Appendix B.

Table 3 reports the results of these regressions and demonstrates that indeed, industries with higher dispersion in expected stock returns and β's exhibit greater dispersion in mpk. Column (1) reveals this fact using expected returns calculated from the Fama-French 3 factor model. Column (2) shows this relationship continues to hold using expected returns predicted using β's only. The Fama-French model explains between about 25% and 30% of the variation in MPK dispersion across industry-years. Column (3) estimates a multiple regression of mpk dispersion on each of the three individual factors - dispersion in each is significantly related to mpk dispersion. In column (4) we take a slightly different approach - we estimate more direct measures of "mpk β's" by regressing firm-level mpk directly on the Fama-French factors (rather than stock returns). For each industry-year, we compute the standard deviation of these β's. The results in column (4) show that dispersion in these alternative measures of β are also significantly related dispersion in mpk. The relationships are highly statistically significant and the R^2 remains close to 25%. In Table 9 in Appendix B, we report results from related regressions where we average our dispersion measures across years for each industry. The findings there are broadly similar (indeed, slightly stronger).[7]

4. MPK dispersion is positively correlated with the price of risk. Expression (2) implies that the price of risk is positively related to mpk dispersion. We investigate this prediction in two ways. First, we show that the measures of the market price of risk considered before (the PD ratio, GZ spread, and excess bond premium) predict time series variation in measures of MPK dispersion. Second, we show that the future expected return on a long-short MPK portfolio are also predicted by these measures of the market price of risk.

We show that both the unconditional dispersion in mpk, and the dispersion of mpk within industries are positively correlated with the lagged price of risk. We control for the changing composition of firms in the following way: using only those firms where our measure of mpk is observed for the firm in consecutive quarters, we compute changes in the standard deviation of mpk (or of industry-demeaned mpk for the within-industry dispersion) for every pair of

[7] Our results are also robust to using a number of different asset pricing models to compute measures of β and expected returns, including CAPM, the Hou et al. (2015) Investment-CAPM, and the Consumption-CAPM models. This relationship is robust to a variety of different controls and industry definitions as well. Table 10 in Appendix B displays the same regression as in Table 3, but with year fixed-effects (reporting within-year R^2), which generates similar results as well.

Table 3: Industry-level Dispersion in *mpk*, Stock Returns and β

	(1)	(2)	(3)	(4)
$\sigma(E[ret])$	2.542***			
	(34.14)			
$\sigma(E_\beta[ret])$		11.63***		
		(31.43)		
$\sigma(\beta_{MKT})$			0.244***	
			(12.22)	
$\sigma(\beta_{HML})$			0.120***	
			(10.63)	
$\sigma(\beta_{SMB})$			0.116***	
			(8.54)	
$\sigma(\beta_{CAPM,MPK})$				0.137***
				(9.30)
$\sigma(\beta_{HML,MPK})$				0.0412***
				(3.90)
$\sigma(\beta_{SMB,MPK})$				0.0549***
				(7.87)
Observations	2721	2746	2734	1427
R^2	0.300	0.265	0.306	0.219

Notes: $\mathbb{E}[ret]$ is the expected return computed from a Fama-Macbeth regression. $\mathbb{E}[ret(\beta)]$ is the expected return predicted from the β's of that regression alone. β' denotes the stock return β on the FF factors and β_{MPK} the *mpk* β on the same factors. t-statistics are in parentheses. Significance levels are denoted by: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

years. We then use those changes to construct a synthetic composition-adjusted measure of the dispersion of *mpk* which is unaffected by new additions or deletions from the dataset. Table 4 displays a regression of the standard deviation of $log(mpk)$ (both within industries and unconditional) on lagged (by one year) measures of the price/dividend ratio, GZ spread, and excess bond premium. All three measures of the business cycle and the market price of risk significantly predict *mpk* dispersion, and in the direction our theory would suggest: The GZ Spread and excess bond premium predict greater *mpk* dispersion, while the PD ratio predicts lower *mpk* dispersion.

As a final test of this prediction, we construct a long-short MPK portfolio and investigate its relation with market price of risk. The portfolio is long the top decile of MPK firms and short the bottom decile, re-balancing every every June based on MPK from the previous year. Table 5 reports a regression of the cumulative twelve month returns on the long-short MPK portfolio on the Pd ratio, GZ spread, and excess bond premium. The GZ spread and excess bond premium rate predict higher future returns on the MPK portfolio, while the PD ratio predicts lower future returns.

Table 4: Time-series Regression of MPK Dispersion

$\sigma(LMPK)$	Within Industry			Unconditional		
PD Ratio	-0.00188***			-0.00502***		
	(-3.01)			(-7.24)		
GZ Spread		0.0123***			0.0221***	
		(3.69)			(3.53)	
EB Premium			0.0271***			0.0472***
			(4.41)			(4.72)
Constant	0.000700	-0.0198***	-0.000349	0.0000136	-0.0362***	-0.00110
	(0.16)	(-2.96)	(-0.08)	(0.00)	(-2.86)	(-0.20)
Observations	148	148	148	148	148	148
R^2	0.054	0.095	0.156	0.200	0.161	0.244

Notes: We regress our measure of composition-adjusted MPK Dispersion (both within industry dispersion or unconditional) on time-series factors. Long-term trends in *mpk* dispersion and the price/dividend ratio are removed using a one-sided hp filter. t-statistics are in parentheses, computed using Newey-West standard errors. Significance levels are denoted by: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All observations are quarterly.

4 Quantitative Analysis

In this section, we use a more detailed version of the investment model laid out above to quantitatively explore the extent to which risk considerations can lead to MPK dispersion. The model is kept deliberately simple in order to isolate the impact of our basic mechanism, namely dispersion in exposure to systematic risk. The model consists of two building blocks: (1) a stochastic discount factor, which we directly parameterize to be consistent with salient patterns in financial markets and (2) a cross-section of heterogeneous firms, which make optimal investment decisions in the presence of firm-level and aggregate risk, given the stochastic discount factor. Specifying the stochastic discount factor exogenously allows to sidetrack challenges with generating empirically relevant risk prices in general equilibrium, and focus on gauging the quantitative strength of our mechanism.

Heterogeneity in risk exposures. Consistent with our assumption Section 2, firm operating profits are given by

$$\Pi_{it} = X_t^{\beta_i} Z_{it} K_{it}^{\theta}, \quad \theta < 1 \tag{6}$$

Firm productivity (in logs, denoted by lowercase) is equal to $\beta_i x_t + z_{it}$, where β_i captures the exposure of firm i to the aggregate shock. Heterogeneity in this exposure is a key element of our framework - cross-sectional variation in β_i will lead directly to dispersion in expected *mpk*.

Table 5: Predictive Regression of Equal Weighted MPK Portfolio Return, Industry Adjusted

Year-Ahead Cumulative Return			
PD Ratio	-0.000456*		
	(-1.70)		
GZ Spread		0.00384**	
		(2.31)	
Excess Bond Premium			0.00580**
			(2.36)
Constant	0.00921***	0.00256	0.00892***
	(8.05)	(1.01)	(8.42)
Observations	152	152	152
R^2	0.048	0.128	0.097

Notes: The dependent variable is the equal-weighted returns from going long firms in the top decile of MPK (after demeaning by sic4) and short the bottom decile, for the following twelve months. Long-term trends in the price-dividend ratio are removed using a one-sided HP filter. t-statistics are in parentheses, computed using Newey-West standard errors. Significance levels are denoted by: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The shocks follow AR(1) processes[8]

$$\begin{aligned} x_{t+1} &= \rho_x x_t + \varepsilon_{t+1}, \quad \varepsilon_{t+1} \sim \mathcal{N}\left(0, \sigma_\varepsilon^2\right) \\ z_{it+1} &= \rho_z z_{it} + \varepsilon_{it+1}, \quad \varepsilon_{it+1} \sim \mathcal{N}\left(0, \sigma_{\tilde{\varepsilon}}^2 - \beta_i^2 \sigma_\varepsilon^2\right) \end{aligned}$$

Stochastic discount factor. In line with the vast literature on cross-sectional asset pricing in production economies, we parameterize directly the pricing kernel without explicitly modeling the consumer's problem. In particular, we follow Zhang (2005) and specify the SDF (in logs) as

$$m_{t+1} = \log \rho + \gamma_t (x_t - x_{t+1}),$$

where

$$\gamma_t = \gamma_0 + \gamma_1 x_t$$

and $\gamma_0 > 0$ and $\gamma_1 < 0$. This formulation allows us to capture in a simple manner a high and time varying, and as a matter of fact, countercyclical (since $\gamma_1 < 0$) price of risk, as observed in the data. Additionally, directly parameterizing γ_0 and γ_1 enables the model to be quantitatively

[8] The variance of the idiosyncratic shock ensures that all firms have the same expected value of the innovations in productivity $\mathbb{E}_t\left[e^{\beta_i \varepsilon_{t+1} + \varepsilon_{it+1}}\right] = e^{\frac{1}{2}\sigma_{\tilde{\varepsilon}}^2}$.

consistent with key moments of asset returns, which are important for our analysis.

MPK dispersion This setup is consistent with the key features of the data revealed in Section 2. Specifically, we show in Appendix A that the realized *mpk* of firm i in period $t+1$ is equal to

$$mpk_{it+1} = \alpha_t + \varepsilon_{it+1} + \beta_i \varepsilon_{t+1} + \beta_i \gamma_t \sigma_\varepsilon^2 \tag{7}$$

where α_t is a time-varying term that is constant across firms, so does not lead to MPK dispersion. Dispersion in realized MPK can be due to uncertainty over the realization of shocks, as well as a risk premium term that is persistent at the firm level and depends on (1) the firm's exposure to the aggregate shock β_i (and is increasing in β_i) and (2) the time-t price of risk, $\gamma_t \sigma_\varepsilon^2$, which is the conditional volatility of the SDF.

Expression (7) also makes clear that cross-sectional differences in the expected MPK only depend on the last (risk-premium) term:

$$\mathbb{E}_t \left[mpk_{it+1} \right] = \alpha_t + \beta_i \gamma_t \sigma_\varepsilon^2 \quad \Rightarrow \quad \sigma^2_{\mathbb{E}_t[mpk]} = \sigma_\beta^2 \left(\gamma_t \sigma_\varepsilon^2 \right)^2 \tag{8}$$

Dispersion in expected MPK depends only on dispersion in firm-level β's, i.e., differences in exposure to the aggregate shock and the price of risk. Intuitively, from expression (7), dispersion in the realized MPK is composed of both transitory components due to uncertainty and a persistent component due to the risk premium. The transitory components, however, are iid over time and thus lead to purely temporary deviations in MPK (this is true even if shocks are autocorrelated); the risk premium, on the other hand, leads to persistent deviations, in which firms that are more exposed to aggregate shocks, and so are riskier, will feature persistently high mpk deviations. Further, dispersion will be greater when the market price of risk $\gamma_t \sigma_\varepsilon^2$ is higher. Notice that the price of risk is countercyclical in the model, since γ_t is higher in bad times, i.e., when x_t is low. These are the two key implications from Section 2 - cross-sectional *mpk* dispersion is increasing in (1) dispersion in firm-level β's and (2) the market price of risk. The remainder of the analysis primarily goes towards quantifying these two objects.

4.1 The Cross-Section of Expected Stock Returns and MPK

In this section, we use a simplified version of the model - namely, with a constant price of risk - to derive a sharp link between a firm's expected *mpk* its expected stock return. This connection provides a novel empirical strategy to quantify the *mpk* dispersion that arises from risk considerations. In Section 4.4 we add back in a time-varying price of risk to capture the countercyclicality of the *mpk* dispersion, but there we show that very similar insights go through (and further, the implications of our theory for the average degree of dispersion remains similar to the results from the simpler model here). Our key finding in this section is that the firm's expected stock return is an affine transformation of its expected *mpk*. This link, first, justifies our use of data on expected stock returns and stock return β's as a proxy for expected *mpk* in Section 3 and second, shows that the dispersion expected stock returns puts tight empirical discipline on the dispersion in expected *mpk* arising from risk channels - indeed, they are proportional to one another. We use this link to provide intuitive, transparent estimates of expected *mpk* dispersion that is due to variation in risk premia.

Consider a simplified version of the model in which $\gamma_1 = 0$, i.e., the price of risk is constant.[9] From the previous section, cross-sectional dispersion in expected *mpk* is given by

$$\mathbb{E}_t\left[mpk_{it+1}\right] = \alpha_t + \beta_i \gamma_0 \sigma_\varepsilon^2 \quad \Rightarrow \quad \sigma^2_{\mathbb{E}_t[mpk]} = \sigma^2_\beta \left(\gamma_0 \sigma^2_\varepsilon\right)^2 \tag{9}$$

Next, Appendix A.3 derives a log-linear approximation to the stock market return, which reveals that the dispersion in expected excess stock returns is given by

$$\mathbb{E}_t\left[r^e_{it+1}\right] = \frac{1}{\psi_1}\beta_i \gamma_0 \sigma_\varepsilon^2 \quad \Rightarrow \quad \sigma^2_{\mathbb{E}_t[r]} = \left(\frac{1}{\psi_1}\right)^2 \sigma^2_\beta \left(\gamma_0 \sigma^2_\varepsilon\right)^2 \tag{10}$$

where ψ_1 is a (constant) function of parameters. Comparing expressions (9) and (10) shows first, that expected stock returns are simply an affine transformation of expected *mpk*, and second, the dispersion in expected returns is proportional to the dispersion in expected *mpk*, with the constant of proportionality given by ψ_1^2 (in variance space). Thus, we have:

$$\sigma^2_{\mathbb{E}_t[mpk]} = \psi_1^2 \sigma^2_{\mathbb{E}_t[r]} \tag{11}$$

Expression (11) reveals a tight connection between cross-sectional variation in expected stock returns and expected *mpk*. The intuition is that the effects of risk premia are all embedded in the cross-section of expected stock returns. Once that is object is known, the link to expected *mpk* only depends on transforming those returns into returns on productive capital. The expression points to a simple way of quantifying dispersion in the expected MPK due to risk - measure the cross-sectional variance of expected stock market returns, the parameters inside the constant ψ_1 and apply the formula in the expression.

A key object is the multiplier ψ_1. It depends only on four parameters - the rate of time discount, ρ, the persistence of aggregate shocks, ρ_x, the curvature of the profit function, θ, and the rate of depreciation, δ. (12) expresses the multiplier in terms of our parameters, while

[9] Deriving expressions for expected stock returns is possible, though more complicated, with a time-varying price of risk. However, similar insights go through in that case as well.

Proposition 1 denotes its properties with respect to these parameters.

$$\psi_1 = \left(\frac{1-\rho\rho_x}{1-\rho}\right)\frac{\frac{1}{\rho}+(1-\theta)\,\delta-1}{\frac{1}{\rho}+\delta-1} \tag{12}$$

Proposition 1. *If $\rho, \theta \in (0,1)$, $\delta, \rho_x \in [0,1]$, then*

1. $\psi_1 > 0$
2. ψ_1 *is increasing in* ρ
3. ψ_1 *is decreasing in* ρ_x
4. ψ_1 *is decreasing in* δ
5. ψ_1 *is weakly decreasing in* θ *(strongly decreasing if* $\delta > 0$*)*

Intuitively, expected stock market returns reflect the long-lived effects of persistent (but not permanent) realized shocks, i.e., their effects on the totality of discounted expected future profits. For patient investors, i.e., with a high ρ, these shocks lead to smaller changes in the value of discounted profits, since more of this value is derived from future profits, which are in the long-run independent of current shocks, and so smaller variation in expected returns, increasing the multiplier. The higher is the persistence of the shocks, ρ_x, the greater effect they have on the sum of discounted profits, increasing return variation and reducing the multiplier. Finally, the lower is curvature (i.e., the larger is θ), the smaller are future profits (e.g., in the limit, as $\theta \to 1$, profits go to zero), reducing the multiplier.[10] (13) shows that the multiplier can also be written as a function of the ratio of market cap to profits (in steady-state), rho_x, and rho.

$$\psi_1 = \left(\frac{1}{\rho}-\rho_x\right)\frac{P}{\Pi} \tag{13}$$

Importantly, the multiplier is completely independent of parameters that govern the aggregate risk premium, i.e., γ_0 and and σ_e^2, and idiosyncratic variation in risk premia, i.e., σ_β^2 , as well as of the parameters that govern the idiosyncratic shock process. The second property is due to the fact that idiosyncratic risk is not priced. The first is because all risk considerations are implicitly embedded in the cross-section of stock market returns. In other words, to quantify risk considerations using expression (11) we do not need to explicitly take a stand on the

[10] As mentioned, the multiplier is decreasing in δ, but it turns out to be not very sensitive to this parameter.

nature of risk and its price or the cross-sectional variation in risk exposure - all of these forces are already reflected by the dispersion in expected stock market returns.

Risk-based dispersion in expected mpk. To use expression (11) to quantify *mpk* dispersion, we must parameterize ρ, ρ_x, θ and δ, as well as estimate the cross-sectional dispersion in expected stock returns. We set the annual discount rate ρ to 0.97, which is consistent with a quarterly discount rate of about 0.99. To estimate the persistence of aggregate shocks, we use data on quarterly TFP growth from John Fernald.[11] Under our AR(1) assumption, this gives a value of 0.095 quarterly (0.815 annually). We set the curvature parameter θ to 0.62, which is a common value in the literature, after taking into account the effects of labor market decisions.[12] We set δ to 0.02 quarterly, which is consistent with an annual depreciation rate of 8%. Using expression (16), these values imply that the (annual) multiplier, ψ_1, is equal to 15.

To estimate the cross-sectional variation in expected stock returns, we must choose an asset pricing model. To be consistent with the broad literature, we use the Fama-French 3 factor model, and estimate a cross-sectional variance of expected returns equal to 0.018.

This suggests that risk-adjusted capital allocation generates a cross-sectional variance of $log(MPK)$ equal to 0.27. This accounts for 41% of the cross-sectional variance of $log(MPK)$, and 57% of the persistent component of the cross-sectional variance in $log(MPK)$.[13]

4.2 Other Distortions

A natural question to ask is to what extent our results are affected by the presence of other frictions/distortions that lead to MPK dispersion (in addition to one-period uncertainty). Recent work has pointed to a number of such factors, including financial frictions, variable markups or policy-induced distortions. Moreover, it has been pointed out that attempts to identify one of these forces - while abstracting from others - may lead to misleading conclusions, and it seems natural to wonder whether this insight applies here. To answer this question, we exploit the tractability of our framework to introduce a 'wedge' a la Hsieh and Klenow (2009) into the firm's first order condition, τ_{it+1}. Specifically, we assume that the firms' revenues are taxed/subsidized at rate $1 - e^{\tau_{it+1}}$ (so that the firm keeps a portion $e^{\tau_{it+1}}$). We follow David and Venkateswaran (2016) by using a flexible specification on this wedge, which allows both for

[11] The data are available at http://www.frbsf.org/economic-research/indicators-data/total-factor-productivity-tfp/.

[12] See, e.g., DAVID et al DAVID/VENKY and COOPER/HALTIWANGER.

[13] The persistent component is defined as the cross-sectional variance of the average MPK firms have other their lifetime.

time-variation and correlation with firm productivity. Specifically, we formulate the wedge as

$$\tau_{it+1} = \nu\left(\beta_i x_{t+1} + z_{it+1}\right) - \xi_{it+1} \tag{14}$$

The wedge is composed of two pieces. The first component is correlated with the firm's productivity, where the degree of correlation is captured by ν. If $\nu < 0$, the wedge discourages (encourages) investment by high (low) productivity firms. If $\nu > 0$, the opposite is true. The second component is uncorrelated with firm characteristics and can be either time-varying or fixed. Low values of ξ spur greater investment by firms irrespective of their underlying characteristics. For simplicity, we assume the firm knows the uncorrelated piece, ξ_{it+1} when it chooses period t investment, i.e., k_{it+1}. Further, we assume that both components of the wedge are uncorrelated with the firm's β. David and Venkateswaran (2016) show that the *mpk* tends to be well-described by this structure and that the correlated piece can capture, for example, models of financial frictions (due, e.g., to liquidity costs) and markups due to monopoly power, in addition to policy-related distortions. We loosely refer to the wedge as a "distortion," although we do not take a stand on whether it stems from efficient factors or not - simply that there are other frictions in the reallocation process. Under this structure, we derive in Appendix A.4 the following expression for the realized MPK:

$$\begin{aligned} mpk_{it+1} &= \alpha_t + \varepsilon_{it+1} + \beta_i \varepsilon_{t+1} - \nu\rho_z z_{it} - \nu\beta_i\rho_x x_{it} + \xi_{it+1} + (1+\nu)\,\beta_i\gamma_0\sigma_\varepsilon^2 \\ &= \alpha_t + \varepsilon_{it+1} + \beta_i \varepsilon_{t+1} - \nu\mathbb{E}_t\left[z_{it+1} + \beta_i x_{it+1}\right] + \xi_{it+1} + (1+\nu)\,\beta_i\gamma_0\sigma_\varepsilon^2 \end{aligned} \tag{15}$$

The distortion has several effects on the realized MPK. The first two terms capture the effects of uncertainty over shocks and are identical to those in the baseline case. Next, the *mpk* includes a component that reflects the severity of the correlation distortion, ν, and depends on the firm's expectations of period $t+1$ productivity - for example, if $\nu < 0$ (which turns out to be the empirically relevant case), the distortion discourages (encourages) investment by high (low) productivity firms, leading to *mrpk* deviations. Next, the *mpk* also depends on the uncorrelated component of distortions ξ - firms with a high realization of ξ_{it+1} will invest more than their fundamentals would dictate. Finally, the last term reflects the risk premium and shows that correlated distortions can scale the risk premium up or down, depending on whether ν is positive or negative. If $\nu < 0$, for a given β_i and price of risk $\gamma_0\sigma_\varepsilon^2$, as the distortion becomes more severe (ν becomes more negative), the magnitude of the risk adjustment term falls (for example, as $\nu \to -1$, the risk adjustment disappears altogether).

From expression (15), we can derive the expected *mpk* as

$$\mathbb{E}_t\left[mpk_{it+1}\right] = \alpha_t - \nu\rho_z z_{it} - \nu\beta_i\rho_x x_t + \xi_{it+1} + (1+\nu)\,\beta_i\gamma_0\sigma_\varepsilon^2$$

and its cross-sectional variance:

$$\sigma^2_{\mathbb{E}_t[mpk]} = \nu^2 \left(\rho_z^2 \sigma_z^2 + \rho_x^2 x_t^2 \sigma_\beta^2\right) + \sigma_\xi^2 + (1+\nu)^2 \sigma_\beta^2 \left(\gamma_0 \sigma_\varepsilon^2\right)^2$$

Dispersion in $\mathbb{E}_t[mpk]$ is generated by three forces - first, the correlated component of the distortion, ν (its contribution to *mpk* dispersion also depends on the cross-sectional variance of expected productivity, which is the term in parentheses); second, the variance of the uncorrelated component; and third, the variance of the (scaled) risk premium, which decreases as ν becomes more negative, i.e., goes towards -1 (realized *mpk*'s have additional sources of variation coming from the dispersion in realized shocks, ε_{it+1} and ε_t).

Appendix A.4 proves that stock market returns are, to a first-order approximation, unaffected by the introduction of other distortions. Thus, a modified version of our approach in Section 4.1 is still valid: dispersion in expected *mpk* stemming from the risk premium term is proportional to dispersion in expected stock market returns, where the factor of proportionality is now $(\psi_1 (1+\nu))^2$, where ψ_1 is as defined in that section. In other words, dispersion in the *mpk*, both in realized and expected terms, can be coming from a host of other sources in addition to the risk-premium terms that is the focus of our analysis. However, even in the presence of these alternative factors, we can still obtain an unbiased estimate of the contribution coming from the risk-premium channel as

$$(1+\nu)^2 \sigma_\beta^2 \left(\gamma_0 \sigma_\varepsilon^2\right)^2 = (\psi_1 (1+\nu))^2 \sigma^2_{\mathbb{E}[r]}$$

If we follow David and Venkateswaran (2016) and specify $\nu = -0.3$, the correlated distortion reduces the amount of cross-sectional MPK dispersion generated by our channel by about half, as firms' responses (in terms of capital choices) to shocks and incentives are dampened.

4.3 Directly Measuring Productivity β's

Our baseline approach to quantifying *mpk* dispersion arising from ex-ante risk exposures used the tight link between expected *mpk* and expected stock market returns outlined in Section 4.1. That method did not require us to directly measure the dispersion in β or the price of risk - both pieces of information were embedded in the observed cross-section of stock returns. In this section, we provide an alternative strategy to estimating the dispersion in ex-ante risk exposures, i.e., σ_β^2 that uses only production-side data and quantify the contribution to *mpk* using moments in aggregate risk premia, specifically, observed Sharpe Ratios. In one sense, this strategy is more direct - there is no need to employ firm-level stock market data to measure risk exposures. On the other hand, computing β's directly from production-side data has its drawbacks - the data are of a lower frequency (quarterly at best) and the time dimension of the panel is shorter. Further, it would be difficult to apply this method to firms in developing countries (where misallocation is larger), since most firm-level datasets there have relatively short panels and are at the annual frequency. For those reasons, we view our results here as an informative check on our baseline findings above.

Our approach is as follows. We can directly measure firm-level productivity as $z_{it} + \beta_i x_t = y_{it} - \alpha k_{it}$ where y_{it} is (the log of) firm revenues. To compute firm-level β's, for each firm we regress measured productivity growth, i.e., $\Delta z_{it} + \beta_i \Delta x_t$ on aggregate productivity growth Δx_t, measured by Fernald's TFP series. It is straightforward to verify that the coefficient from this regression is exactly equal to β_i. Using these estimates, we calculate the cross-sectional variance in β's, σ_β^2. This procedure gives a value of σ_β equal to 25. This is large because aggregate TFP has relatively small volatility, while measures of firm TFP can be much more sensitive to economic downturns (and heterogeneous).

To quantify the contribution to expected mpk dispersion, recall from expression (8) that

$$\sigma^2_{\mathbb{E}_t[mpk]} = \sigma^2_\beta \left(\gamma_t \sigma^2_\varepsilon\right)^2$$

In principal, we can use the calibrated values of γ_0 and γ_1 (along with σ_ε^2) to compute the right hand side of the expression. However, there turns out to be a more direct way, which does not requires us to calibrate the γ parameters. Specifically, the maximal conditional Sharpe ratio in the economy, which is equal to the standard deviation of the (log of the) stochastic discount factor, is given by $\gamma_t \sigma_\varepsilon$.[14] Using this result, we have

$$\sigma^2_{\mathbb{E}_t[mpk]} = \sigma^2_\beta \left(SR_t \sigma_\varepsilon\right)^2$$

where SR_t denotes the period t conditional Sharpe ratio. Thus, using our direct estimates of σ_β^2 to quantify the implications of risk for expected mpk dispersion simply requires an estimate of σ_ε, which we have already seen is available from data on aggregate TFP and values of the maximum Sharpe ratio.

Table 6: Alternate Approach: βs

	$\sigma^2_{\mathbb{E}[mpk]}$	$\frac{\sigma^2_{\mathbb{E}[mpk]}}{\sigma^2_{mpk}}$	$\frac{\sigma^2_{\mathbb{E}[mpk]}}{\sigma^2_{\overline{mpk}}}$
SR=0.5	0.04	0.06	0.08
SR=0.7	0.08	0.12	0.16
SR=1.0	0.16	0.24	0.33

[14] See, e.g., Campbell (2003) for a proof.

We use the value of σ_ε of 0.016 annually. Standard estimates of the annualized Sharpe ratio on the market portfolio range from 0.5 (CITE COCHRANE) to 0.6 (MACKINLAY).[15] There is good reason to believe these are lower bounds on the maximum achievable Sharpe ratio. For example, diversified mutual funds can obtain Sharpe ratios in excess of 0.7, and some asset classes and hedge funds can earn annualized Sharpe ratios in excess of 1. Thus, we report results for three values - 0.5, 0.7, and 1. Table 6 displays the implied cross-sectional variance in $log(MPK)$ implied by our risk-adjusted channel, implied by our measured $Beta_i$ dispersion and varying Sharpe ratios. With higher numbers for the maximal Sharpe ratio, this alternative strategy can also account for a significant part of cross-sectional dispersion in firm marginal products of capital. For lower values of the Sharpe ratio, the implied amount of MPK dispersion falls. However, taking the estimated distribution of Betas and the Sharpe ratio, along with standard parameters, implies very low cross-sectional dispersion in expected returns as well. Reconciling expected return dispersion with such a model may require multiple aggregate shocks which affect firms heterogeneously, which would also imply greater MPK dispersion.

4.4 Numerical Results

In this section, we turn to numerical methods to solve the model and estimate the underlying distribution of β's. Specifically, given the link uncovered in the previous section, we parameterize the dispersion in β's to match the cross-section of expected stock market returns.Further, we work with the more general case featuring a time-varying price of risk in order to address the second main prediction of our theory regarding expected MPK dispersion, i.e. its countercyclicality.

Table 7 summarizes our empirical approach. The calibration of the aggregate shock process is standard in the macro literature and follows Cooley and Prescott (1995). In particular, we set the quarterly conditional volatility, σ_ε to 0.007 and the persistence, ρ_x to 0.95, which give annual analogs of 0.014 and 0.81. The calibration of the stochastic discount factor follows Zhang (2005) - as in that paper, we set $\gamma_0 = 50$ and $\gamma_1 = -1000$. Additionally, we set $\rho = 0.99$. These values imply (1) an annualized real interest rate of 2.68% with an annualized volatility of 2.07% and (2) an average quarterly conditional Sharpe Ratio of 0.38.

We set returns to scale to 0.62 and depreciation to 0.025, which are standard values in the literature.[16] We set the mean volatility and persistence of firm-level shocks, $\sigma_{\tilde{\varepsilon}}$ and ρ_z to 0.05 and 0.9 respectively, in order to generate realistic cross-sectional dispersion in investment rates and profitability, similar to Gomes (2001).

[15] LO reports even high values after adjusting for serial correlation in returns.
[16] CITES

Table 7: Parameterization - Summary

Parameter	Description	Target/Value
Aggregate shocks/SDF		
σ_ε	Std. dev. of agg. shock	0.007
ρ_x	Persistence of agg. shock	0.95
γ_0	SDF - constant component	50
γ_1	SDF - time-varying component	-1000
Preferences/production		
ρ	Discount rate	0.99
θ	Returns to scale	0.62
$\sigma_{\tilde{\varepsilon}}$	Std. dev. of idiosyncratic shock	0.05
ρ_z	Persistence of idiosyncratic shock	0.9
σ_β	Std. dev of risk exposures	

Finally, and critically, we need to pin down the ex-ante dispersion in firms' exposure to systematic risk, σ^2_β. We estimate this dispersion so as to generate the average empirical ex-post dispersion of expected stock returns within an industry. We choose the within-industry dispersion as our target as other than risk exposures, our model does not feature any industry-specific elements. This value is, to some extent, dependent on the asset pricing model chosen and the precise form of the calculation. For example, adopting the Fama-French three-factor model as our benchmark results in a cross-sectional variance of expected stock returns of between 0.0014 and 0.023, depending on the interpretation of firm-specific α's (e.g., whether we include variation in α's as capturing unmeasured exposures to aggregate risks, or we exclude them as simply noise). In our benchmark simulations reported below, we choose a dispersion of 0.01, so roughly the midpoint, and provide sensitivity with respect to these choice.

We solve the model via value function iteration and estimate the dispersion in firm-level β's via a simulated method of moments approach, i.e., we choose σ^2_β so that the ex-post dispersion in expected stock returns from a simulated panel of firms matches the value from the data.

Results Our objective is to gauge the amount of MPK dispersion that a frictionless dynamic investment model with ex-ante heterogeneity in risk exposure can generate, *once calibrated to salient asset market data*. Also, we explore the dynamics of the cross-sectional dispersion in MPK over the business cycle, when we allow for realistically countercyclical movements in risk prices. We do so by examining the statistical properties of simulated panels of firms, from our benchmark calibration as well as relevant variations.

Our first set of results are in table 8. A first account of the results is as follows. For various parameter choices for the basic risk price parameters γ_0 and γ_1, we compute the cross-sectional standard deviation of within industry MPK, our main measure of variation in the marginal

Table 8: MPK and Risk

Model	(i)	(ii)	(iii)	(iv)	(v)	(vi)
A. Parameters						
γ_0	40	50	60	40	50	60
γ_1	-1000	-1000	-1000	0	0	0
B. Moments						
Average Sharpe Ratio	0.29	0.38	0.44	0.28	0.37	0.44
$\sigma^2_{\mathbb{E}[mpk]}$	0.22	0.36	0.47	0.24	0.28	0.48
$\text{Corr}(\sigma_{\mathbb{E}[mpk]}, y_t)$	-0.16	-0.22	-0.27	-0.06	0.03	0.02

This table reports average moments of simulated panels across various model specifications. We focus on varying the parameters underlying the stochastic discount factor, γ_0 and γ_1 to assess the effects of risk premia on the levels and dynamics of average *mpk*. To compute $\sigma_{\mathbb{E}[mpk]}$, we compute average *mpk* per firm, and determine its cross-sectional standard deviation.

product of capital. Our baseline case considered above, with $\gamma_0 = 50$ and $\gamma_1 = -1000$, gives a model analog of about 0.36. To gauge the magnitude of these results, we have computed the total amount of misallocation, i.e., σ^2_{mpk} in our data. The overall level of dispersion is equal to 0.67. Because our theory primarily speaks to persistent deviations from *mpk* equalization at the firm-level, we have also separated the *mpk* of each firm into a transitory and permanent component and computed the cross-sectional variance of the latter - this gives a value of 0.48. Putting these together, the results imply that ex-ante heterogeneity in risk exposures account for about 54% of overall *mpk* dispersion among Compustat firms and about three-quarters of dispersion in the permanent component of *mpk*.

Naturally, the amount of *mpk* dispersion is sensitive to the implied Sharpe ratios implicit in the stochastic discount factor (i.e., the degree of aggregate risk). There is some debate in the literature as to what Sharpe ratios can be obtained in financial markets, with value-growth portfolios reaching Sharpe ratios of closer to 0.8, and alternative investment strategies involving hedge fund and private equity exposure reporting Sharpe ratios above one, but likely subject to substantial measurement error. The baseline calibration of the stochastic discount factor matches well the Sharpe ratio on the market portfolio, but likely only gives a lower bound to Sharpe ratio attainable in financial markets by means of alternative investment strategies. We can trace out the implications of various indications of attainable Sharpe ratios in financial markets for MPK dispersion by suitable re-calibrations of the stochastic discount factor. The table shows that varying Sharpe ratios in a relatively confined range results in some variation in the implied *mpk* dispersion, although reasonable values of the Sharpe Ratio all suggest significant effects.

The table also shows that MPK dispersion displays significant variation over the business cycle provided the market price of risk is countercyclical, that is, when $\gamma_1 < 0$. In particular, the dispersion becomes notably countercyclical, in line with the empirical results in Eisfeldt and Rampini (2006). In the light of our beta pricing equation, (5), it is not surprising that countercyclical MPK dispersion obtains even in a completely frictionless model. The table confirms that these effects are empirically relevant.

5 Conclusion

In this paper, we have revisited the notion of 'misallocation' from the perspective of a risk-sensitive, or risk-adjusted, version of the stochastic growth model. The standard first order condition for investment in that framework suggests that expected firm-level marginal products should reflect exposure to factor risks, and their pricing. To the extent that firms are differentially exposed to these risks, as the literature on cross-sectional asset pricing suggests, the implication is that cross-sectional dispersion in *mpk* may not only reflect true misallocation, but also risk-adjusted capital allocation.

We empirically evaluate this proposition, and find strong support for it. Indeed, from an asset pricing perspective, we show that a long short portfolio of high minus low MPK stocks earns a significant premium, so that high MPK firms are effectively riskier, and that this premium is predictable and countercyclical, so that their cost of capital rises disproportionately in bad times. A calibrated dynamic model suggests that, indeed, risk-adjusted capital allocation accounts for a substantial fraction of observed *mpk* dispersion.

References

Asker, J., A. Collard-Wexler, and J. De Loecker (2014): "Dynamic inputs and resource (mis) allocation," *Journal of Political Economy*, 122, 1013–1063.

Balvers, R. J., L. Gu, D. Huang, M. Lee-Chin, et al. (2015): "Profitability, value and stock returns in production-based asset pricing without frictions," *Journal of Money, Credit, and Banking.*

Binsbergen, J. V. and C. Opp (2017): "Real Anomalies," Wharton Working Paper.

Buera, F. J., J. P. Kaboski, and Y. Shin (2011): "Finance and Development: A Tale of Two Sectors," *American Economic Review*, 101, 1964–2002.

Campbell, J. Y. (2003): "Consumption-based asset pricing," *Handbook of the Economics of Finance*, 1, 803–887.

Cochrane, J. (1991): "Production-Based Asset Pricing and the Link Between Stock Returns and Economic Fluctuations," *Journal of Finance*, 46, 207–234.

Cooley, T. F. and E. C. Prescott (1995): "Economic Growth and Business Cycles," *Frontiers of business cycle research*, 1.

David, J. M., H. A. Hopenhayn, and V. Venkateswaran (2016): "Information, Misallocation and Aggregate Productivity," *The Quarterly Journal of Economics*, 131, 943–1005.

David, J. M. and V. Venkateswaran (2016): "Capital Misallocation: Frictions or Distortions?" .

Eisfeldt, A. and A. Rampini (2006): "Capital reallocation and liquidity," *Journal of Monetary Economics*, 53, 369–399.

Fama, E. F. and K. R. French (1992): "Cross-Section of Expected Stock Returns," *The Journal of Finance*, 47, 3247–3265.

Gilchrist, S. and E. Zakrajsek (2012): "Credit Spreads and Business Cycle Fluctuations," *American Economic Review*, 102, 1692–1720.

Gomes, J. (2001): "Financing Investment," *American Economic Review.*

Gomes, J. and L. Schmid (2010): "Levered Returns," *Journal of Finance*, 65, 467–494.

Gomes, J., A. Yaron, and L. Zhang (2006): "Asset pricing implications of firms financing constraints," *Review of Financial Studies*, 19, 1321–1356.

Gopinath, G., S. Kalemli-Ozcan, L. Karabarbounis, and C. Villegas-Sanchez (2015): "Capital allocation and productivity in south europe," Tech. rep., National Bureau of Economic Research.

Hou, K., C. Xue, and L. Zhang (2015): "Digesting anomalies: An investment approach," *Review of Financial Studies*.

Hsieh, C. and P. Klenow (2009): "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 124, 1403–1448.

Lewellen, J. and S. Nagel (2006): "The conditional CAPM does not explain asset-pricing anomalies," *Journal of Financial Economics*, 82, 289–314.

Liu, L. X., T. M. Whited, and L. Zhang (2009): "Investment-based expected stock returns," *Journal of Political Economy*, 117, 1105–1139.

Midrigan, V. and D. Y. Xu (2014): "Finance and misallocation: Evidence from plant-level data," *The American Economic Review*, 104, 422–458.

Moll, B. (2014): "Productivity losses from financial frictions: can self-financing undo capital misallocation?" *The American Economic Review*, 104, 3186–3221.

Restoy, F. and M. Rockinger (1994): "On Stock Market Returns and Returns on Investment," *Journal of Finance*, 49, 543–556.

Restuccia, D. and R. Rogerson (2008): "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments," *Review of Economic Dynamics*, 11, 707–720.

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

——— (2017): "The Investment CAPM," *European Financial Management*.

Appendix

A Derivations and Proofs

This appendix provides detailed derivations for the expressions in the text.

A.1 Motivation

Derivation of equation (2).

$$\begin{aligned} 1 &= \mathbb{E}_t\left[M_{t+1}\left(MPK_{it+1}+1-\delta\right)\right] \\ &= \mathbb{E}_t\left[M_{t+1}\right]\mathbb{E}_t\left[MPK_{it+1}+1-\delta\right]+\text{cov}\left(M_{t+1},MPK_{it+1}\right) \\ &= \mathbb{E}_t\left[M_{t+1}\right]\left(MPK_{ft+1}+1-\delta\right) \\ \Rightarrow \mathbb{E}_t\left[MPK_{t+1}\right] &= MPK_{ft+1}-\frac{\text{cov}\left(M_{t+1},MPK_{t+1}\right)}{\mathbb{E}_t\left[M_{t+1}\right]} \\ &= \alpha_t+\beta_{it}\lambda_t \end{aligned}$$

where α_t, β_{it} and λ_t are as defined in the text, and MPK_{ft+1} is the MPK of the 'risk-free' firm defined by $\text{cov}\left(M_{t+1},MPK_{ft+1}\right)=0$. By a no-arbitrage condition, it must be the case that $\frac{1}{\mathbb{E}_t[M_{t+1}]}=MPK_{ft+1}+1-\delta=R_{ft}$ where R_{ft} is the gross risk-free interest rate.

No aggregate risk. With no aggregate risk, $M_{t+1}=\rho\,\forall\,t$ where ρ is the rate of time discount. The Euler equation gives

$$1=\rho\left(\mathbb{E}_t\left[MPK_{it+1}\right]+1-\delta\right)\ \forall\ i,t \quad \Rightarrow \quad \mathbb{E}_t\left[MPK_{it+1}\right]=\frac{1}{\rho}-\left(1-\delta\right)=r_f+\delta$$

CAPM. Clearly, $-\text{cov}\left(M_{t+1},MPK_{it+1}\right)=b\text{cov}\left(R_{mt+1},MPK_{it+1}\right)$ and $\text{var}\left(M_{t+1}\right)=b^2\text{var}\left(R_{mt+1}\right)$. Since the market return is an asset, it must satisfy $\mathbb{E}_t\left[R_{mt+1}\right]=R_{ft}+\frac{\lambda_t}{b}$ so that $\lambda_t=b\left(\mathbb{E}_t\left[R_{mt+1}\right]-R_{ft}\right)$. Substituting into expression (2) gives the CAPM expression in the text.

CCAPM. A log-linear approximation to the SDF around its unconditional mean gives $M_{t+1}\approx\mathbb{E}\left[M_{t+1}\right]\left(1+m_{t+1}-\mathbb{E}\left[m_{t+1}\right]\right)$ and in the case of $CRRA$ utility, $m_{t+1}=-\gamma\Delta c_{t+1}$ where Δc_{t+1} is log consumption growth. Substituting for M_{t+1} into expression (2) gives the CCAPM expression in the text.

A.2 MPK Dispersion

The Euler equation gives

$$\begin{aligned} 1 &= \mathbb{E}_t\left[M_{t+1}\left(\theta e^{z_{it+1}+\beta_i x_{t+1}}K_{it+1}^{\theta-1}+1-\delta\right)\right] \\ &= (1-\delta)\,\mathbb{E}_t\left[M_{t+1}\right]+\theta K_{it+1}^{\theta-1}\mathbb{E}_t\left[e^{m_{t+1}+z_{it+1}+\beta_i x_{t+1}}\right] \end{aligned}$$

Substituting for m_{t+1} and rearranging,

$$\begin{aligned} \mathbb{E}_t\left[e^{m_{t+1}+z_{it+1}+\beta_i x_{t+1}}\right] &= \mathbb{E}_t\left[e^{\log\rho+z_{it+1}+(\beta_i-\gamma_0-\gamma_1 x_t)x_{t+1}+(\gamma_0+\gamma_1 x_t)x_t}\right] \\ &= \mathbb{E}_t\left[e^{\log\rho+\rho_z z_{it}+\varepsilon_{it+1}+(\beta_i-\gamma_0-\gamma_1 x_t)(\rho_x x_t+\varepsilon_{t+1})+(\gamma_0+\gamma_1 x_t)x_t}\right] \\ &= \mathbb{E}_t\left[e^{\log\rho+\rho_z z_{it}+\varepsilon_{it+1}+((\beta_i-\gamma_0-\gamma_1 x_t)\rho_x+\gamma_0+\gamma_1 x_t)x_t+(\beta_i-\gamma_0-\gamma_1 x_t)\varepsilon_{t+1}}\right] \\ &= e^{\log\rho+\rho_z z_{it}+((\beta_i-\gamma_0-\gamma_1 x_t)\rho_x+\gamma_0+\gamma_1 x_t)x_t+\frac{1}{2}(\beta_i-\gamma_0-\gamma_1 x_t)^2\sigma_\varepsilon^2+\frac{1}{2}\left(\sigma_{\tilde{\varepsilon}}^2-\beta_i^2\sigma_\varepsilon^2\right)} \\ &= e^{\log\rho+\rho_z z_{it}+((\beta_i-\gamma_0-\gamma_1 x_t)\rho_x+\gamma_0+\gamma_1 x_t)x_t-\beta_i(\gamma_0+\gamma_1 x_t)\sigma_\varepsilon^2+\gamma_1 x_t\left(\gamma_0+\frac{1}{2}\gamma_1 x_t\right)\sigma_\varepsilon^2+\frac{1}{2}\gamma_0^2\sigma_\varepsilon^2+\frac{1}{2}\sigma_{\tilde{\varepsilon}}^2} \end{aligned}$$

so that

$$\theta K_{it+1}^{\theta-1}=\frac{1-(1-\delta)\,\mathbb{E}_t\left[M_{t+1}\right]}{e^{\log\rho+\rho_z z_{it}+((\beta_i-\gamma_0-\gamma_1 x_t)\rho_x+\gamma_0+\gamma_1 x_t)x_t-\beta_i(\gamma_0+\gamma_1 x_t)\sigma_\varepsilon^2+\gamma_1 x_t\left(\gamma_0+\frac{1}{2}\gamma_1 x_t\right)\sigma_\varepsilon^2+\frac{1}{2}\gamma_0^2\sigma_\varepsilon^2+\frac{1}{2}\sigma_{\tilde{\varepsilon}}^2}}$$

and in logs,

$$\begin{aligned} k_{it+1} &= \frac{1}{1-\theta}\left[\tilde{\alpha}_t+\rho_z z_{it}+\beta_i\rho_x x_t-\beta_i\left(\gamma_0+\gamma_1 x_t\right)\sigma_\varepsilon^2\right] \\ &= \frac{1}{1-\theta}\left[\tilde{\alpha}_t+\rho_z z_{it}+\beta_i\rho_x x_t-\beta_i\gamma_t\sigma_\varepsilon^2\right] \end{aligned}$$

where

$$\begin{aligned} \tilde{\alpha}_t &= \log\theta-\log\left(1-(1-\delta)\,\mathbb{E}_t\left[M_{t+1}\right]\right)+\log\rho-\left(\left(\gamma_0+\gamma_1 x_t\right)\rho_x-\gamma_0-\gamma_1 x_t\right)x_t \\ &+ \gamma_1 x_t\left(\gamma_0+\frac{1}{2}\gamma_1 x_t\right)\sigma_\varepsilon^2+\frac{1}{2}\gamma_0^2\sigma_\varepsilon^2+\frac{1}{2}\sigma_{\tilde{\varepsilon}}^2 \\ &= \log\theta-\log\left(1-(1-\delta)\,\mathbb{E}_t\left[M_{t+1}\right]\right)+\log\rho+\gamma_t\left(1-\rho_x\right)x_t+\frac{1}{2}\gamma_t^2\sigma_\varepsilon^2+\frac{1}{2}\sigma_{\tilde{\varepsilon}}^2 \end{aligned}$$

is a time-varying term that is constant across firms. Next,

$$\begin{aligned} mpk_{it+1} &= \log\theta + \pi_{it+1} - k_{it+1} \\ &= \log\theta + z_{it+1} + \beta_i x_{t+1} - (1-\theta)\, k_{it+1} \\ &= \log\theta + z_{it+1} + \beta_i x_{t+1} - \tilde{\alpha}_t - \rho_z z_{it} - \beta_i \rho_x x_t + \beta_i \gamma_t \sigma_\varepsilon^2 \\ &= \alpha_t + \varepsilon_{it+1} + \beta_i \varepsilon_{t+1} + \beta_i \gamma_t \sigma_\varepsilon^2 \end{aligned}$$

where $\alpha_t = \log\theta - \tilde{\alpha}_t$. The time-$t$ conditional expected mpk is

$$\mathbb{E}_t\left[mpk_{it+1}\right] = \alpha_t + \beta_i \gamma_t \sigma_\varepsilon^2$$

A.3 Stock Market Returns

We consider a log-linear approximation around the non-stochastic steady state, where $z_{it} = x_t = 1$ and so $M_t = \rho$. In the steady state, we have

$$\begin{aligned} MPK &= \frac{1}{\rho} + \delta - 1 \quad \Rightarrow \quad K = \left[\frac{1}{\theta}\left(\frac{1}{\rho} + \delta - 1\right)\right]^{\frac{1}{\theta-1}} \\ \Pi &= K^\theta \quad \Rightarrow \quad D = K^\theta - \delta K \\ P &= \frac{\rho}{1-\rho} D \\ R &= 1 + \frac{D}{P} \quad \Rightarrow r_f = -\log\rho \end{aligned}$$

where P denotes the (ex-dividend) stock price and R and r_f the gross and net interest rates, respectively.

The firm's dividend is given by

$$D_{it+1} = \Pi_{it+1} - K_{it+2} + (1-\delta) K_{it+1}$$

and log-linearizing,

$$\hat{d}_{it+1} = \frac{\Pi}{D}(\hat{z}_{it+1} + \beta_i \hat{x}_{t+1}) + \left[\theta \frac{\Pi}{D} + (1-\delta)\frac{K}{D}\right]\hat{k}_{it+1} - \frac{K}{D}\hat{k}_{it+2}$$

Adding back in the steady state values to convert into levels,

$$\begin{aligned} d_{it+1} &= \log D + \frac{\Pi}{D}(z_{it+1} + \beta_i x_{t+1}) + \left[\theta\frac{\Pi}{D} + (1-\delta)\frac{K}{D}\right](k_{it+1} - \log K) - \frac{K}{D}(k_{it+2} - \log K) \\ &= \frac{\Pi}{D}(z_{it+1} + \beta_i x_{t+1}) + \left[\theta\frac{\Pi}{D} + (1-\delta)\frac{K}{D}\right]k_{it+1} - \frac{K}{D}k_{it+2} + \log D - \left(\theta\frac{\Pi}{D} - \delta\frac{K}{D}\right)\log K \end{aligned}$$

From above, we have

$$k_{it+1} = \frac{1}{1-\theta}[\log\theta - \Phi_t + \rho_z z_{it} + \beta_i \rho_x x_t - \beta_i \gamma_0 \sigma_\varepsilon^2]$$

where

$$\Phi_t = \log(1 - (1-\delta)e^{-r_{ft}}) + r_{ft} - \frac{1}{2}\sigma_{\tilde{\varepsilon}}^2$$

(since $r_{ft} = -\log \mathbb{E}_t[\exp m_{t+1}] = -\log\rho - \gamma_0(1-\rho_x)x_t - \frac{1}{2}\gamma_0^2\sigma_\varepsilon^2$). Log-linearizing gives

$$\begin{aligned} \Phi_t &\approx \left[\log(1-(1-\delta)e^{-r_f}) - \frac{1}{2}\sigma_{\tilde{\varepsilon}}^2 - \frac{(1-\delta)e^{-r_f}}{1-(1-\delta)e^{-r_f}}r_f\right] + \frac{1}{1-(1-\delta)e^{-r_f}}r_{ft} \\ &= \left[\log(1-(1-\delta)e^{-r_f}) - \frac{1}{2}\sigma_{\tilde{\varepsilon}}^2 - \frac{(1-\delta)e^{-r_f}}{1-(1-\delta)e^{-r_f}}r_f\right] - \frac{\log\rho + \frac{1}{2}\gamma_0^2\sigma_\varepsilon^2}{1-(1-\delta)e^{-r_f}} - \frac{\gamma_0(1-\rho_x)}{1-(1-\delta)e^{-r_f}}x_t \end{aligned}$$

Defining the constant:

$$\Phi = \left[\log(1-(1-\delta)e^{-r_f}) - \frac{1}{2}\sigma_{\tilde{\varepsilon}}^2 - \frac{(1-\delta)e^{-r_f}}{1-(1-\delta)e^{-r_f}}r_f\right] - \frac{\log\rho + \frac{1}{2}\gamma_0^2\sigma_e^2}{1-(1-\delta)e^{-r_f}}$$

Φ_t can be written as

$$\Phi_t = \Phi - \frac{\gamma_0(1-\rho_x)}{1-(1-\delta)e^{-r_f}}x_t$$

Using this, along with the equation for k_{it+1}, we can write

$$\begin{aligned} d_{it+1} &= \left[\frac{\Pi}{D} + \frac{1}{1-\theta}\left(\theta\frac{\Pi}{D} + (1-\delta)\frac{K}{D}\right) - \frac{1}{1-\theta}\frac{K}{D}\rho_z\right]\rho_z z_{it} \\ &+ \left[\frac{\Pi}{D} + \frac{1}{1-\theta}\left(\theta\frac{\Pi}{D} + (1-\delta)\frac{K}{D}\right) - \frac{1}{1-\theta}\frac{K}{D}\rho_x\right]\beta_i\rho_x x_t \\ &+ \left[\frac{\Pi}{D} - \frac{K}{D}\rho_z\frac{1}{1-\theta}\right]\varepsilon_{it+1} + \left[\frac{\Pi}{D} - \frac{K}{D}\rho_x\frac{1}{1-\theta}\right]\beta_i\varepsilon_{t+1} \\ &+ \left[\frac{\Pi}{D} - \delta\frac{K}{D}\right]\frac{1}{1-\theta}\left(\log\theta - \beta_i\gamma_0\sigma_\varepsilon^2\right) + \left[\frac{K}{D}\Phi_{t+1} - \left(\theta\frac{\Pi}{D} + (1-\delta)\frac{K}{D}\right)\Phi_t\right]\frac{1}{1-\theta} \\ &+ \log D - \left(\theta\frac{\Pi}{D} - \delta\frac{K}{D}\right)\log K \end{aligned}$$

and since

$$\left[\frac{K}{D}\Phi_{t+1} - \left(\theta\frac{\Pi}{D} + (1-\delta)\frac{K}{D}\right)\Phi_t\right]\frac{1}{1-\theta} = -\frac{\Phi}{1-\theta}\left(\theta\frac{\Pi}{D} - \delta\frac{K}{D}\right) + \frac{1}{1-\theta}\frac{\gamma_0(1-\rho_x)}{1-(1-\delta)e^{-r_f}}\left(\theta\frac{\Pi}{D} + (1-\delta)\frac{K}{D} - \rho_x\frac{K}{D}\right)x_t - \frac{1}{1-\theta}\frac{\gamma_0(1-\rho_x)}{1-(1-\delta)e^{-r_f}}\frac{K}{D}\varepsilon_{t+1}$$

we can write

$$d_{it+1} = (\tilde{A}_1 + A_2)\rho_z z_{it} + [(A_1 + A_2)\beta_i\rho_x + A_4]x_t + \tilde{A}_1\varepsilon_{it+1} + (\beta_i A_1 - A_3)\varepsilon_{t+1} + A_0$$

where

$$\begin{aligned}
A_0 &= \frac{1}{1-\theta}\left(\alpha\frac{\Pi}{D} - \delta\frac{K}{D}\right)\left[\log\theta - \beta_i\gamma_0\sigma_\varepsilon^2 - \Phi\right] + \log D - \left(\theta\frac{\Pi}{D} - \delta\frac{K}{D}\right)\log K \\
A_1 &= \frac{\Pi}{D} - \frac{K}{D}\frac{\rho_x}{1-\theta} \\
\tilde{A}_1 &= \frac{\Pi}{D} - \frac{K}{D}\frac{\rho_z}{1-\theta} \\
A_2 &= \frac{1}{1-\theta}\left[\theta\frac{\Pi}{D} + (1-\delta)\frac{K}{D}\right] \\
A_3 &= \frac{1}{1-\theta}\frac{K}{D}\frac{\gamma_0(1-\rho_x)}{1-(1-\delta)e^{-r_f}} \\
A_4 &= \frac{1}{1-\theta}\left(\theta\frac{\Pi}{D} + (1-\delta)\frac{K}{D} - \rho_x\frac{K}{D}\right)\frac{\gamma_0(1-\rho_x)}{1-(1-\delta)e^{-r_f}} \\
A_5 &= \frac{1}{1-\theta}\frac{1}{\rho_x}\frac{\gamma_0(1-\rho_x)}{1-(1-\delta)e^{-r_f}}\left(\theta\frac{\Pi}{D} + (1-\delta)\frac{K}{D}\right)
\end{aligned}$$

with $A_4 = \rho_x(A_5 - A_3)$.

Next, log-linearizing the return equation $R_{it+1} = \frac{D_{it+1}+P_{it+1}}{P_{it}}$ gives

$$r_{it+1} = \frac{1}{R}p_{it+1} - p_{it} + \frac{R-1}{R}d_{it+1} + \frac{R-1}{R}\log\frac{P}{D} + \log R$$

Conjecture that p_{it} is linear in the states, i.e., $p_{it} = c_0 + c_1\beta_i\varepsilon_t + c_2\varepsilon_{it} + c_3\beta_i x_t + c_4 z_{it}$. Then,

$$
\begin{aligned}
r_{it+1} &= \left(1-\frac{1}{R}\right)(A_0 - c_0) + \frac{R-1}{R}\log\frac{P}{D} + \log R \\
&+ \left[\frac{\beta_i}{R}(c_1+c_3) + \frac{R-1}{R}(\beta_i A_1 - A_3)\right]\varepsilon_{t+1} + \left[\frac{c_2+c_4}{R} + \frac{R-1}{R}\tilde{A}_1\right]\varepsilon_{it+1} \\
&+ \left[\left(\frac{\rho_x}{R}-1\right)c_3\beta_i + \frac{R-1}{R}[(A_1+A_2)\beta_i\rho_x + A_4]\right]x_t + \left[\left(\frac{\rho_z}{R}-1\right)c_4 + \frac{R-1}{R}(\tilde{A}_1+A_2)\rho_z\right]z_{it} \\
&- c_1\beta_i\varepsilon_t - c_2\varepsilon_{it}
\end{aligned}
$$

The Euler equation requires

$$\log \mathbb{E}_t(e^{r_{it+1}+m_{t+1}}) = 0$$

so that

$$
\begin{aligned}
0 &= \left(1-\frac{1}{R}\right)(A_0 - c_0) + \frac{1}{2}\left[\frac{\beta_i}{R}(c_1+c_3) + \frac{R-1}{R}(\beta_i A_1 - A_3) - \gamma_0\right]^2\sigma_e^2 \\
&+ \frac{1}{2}\left[\frac{c_2+c_4}{R} + \frac{R-1}{R}\tilde{A}_1\right]^2(\sigma_{\tilde{\varepsilon}}^2 - \beta_i^2\sigma_\varepsilon^2) \\
&+ \left[\left(\frac{\rho_x}{R}-1\right)c_3\beta_i + \frac{R-1}{R}[(A_1+A_2)\beta_i\rho_x + A_4] + \gamma_0(1-\rho_x)\right]x_t \\
&+ \left[\left(\frac{\rho_z}{R}-1\right)c_4 + \frac{R-1}{R}(\tilde{A}_1+A_2)\rho_z\right]z_{it} \\
&- c_1\beta_i\varepsilon_t - c_2\varepsilon_{it} + \log\rho + \frac{R-1}{R}\log\frac{P}{D} + \log R
\end{aligned}
$$

This must hold for all ε_{it}, ε_t, z_{it} and x_t. Hence, the coefficients satisfy:

$$
\begin{aligned}
c_1 &= 0 \\
c_2 &= 0 \\
c_3\beta_i\left(\frac{\rho_x}{R}-1\right) &= -\frac{R-1}{R}[(A_1+A_2)\beta_i\rho_x + A_4] - \gamma_0(1-\rho_x) \\
c_4\left(\frac{\rho_x}{R}-1\right) &= -\frac{R-1}{R}(\tilde{A}_1+A_2)\rho_z \\
\left(1-\frac{1}{R}\right)(A_0-c_0) &= -\frac{1}{2}\left[\frac{\beta_i}{R}(c_1+c_3) + \frac{R-1}{R}(\beta_i A_1 - A_3) - \gamma_0\right]^2\sigma_\varepsilon^2 \\
&- \frac{1}{2}\left[\frac{c_2+c_4}{R} + \frac{R-1}{R}\tilde{A}_1\right]^2(\sigma_{\tilde{\varepsilon}}^2 - \beta_i^2\sigma_\varepsilon^2) - \log\rho
\end{aligned}
$$

and

$$r_{it+1} = \left(1 - \frac{1}{R}\right)(A_0 - c_0) + \left[\frac{\beta_i c_3}{R} + \frac{R-1}{R}(\beta_i A_1 - A_3)\right]\varepsilon_{t+1} + \left[\frac{c_4}{R} + \frac{R-1}{R}\tilde{A}_1\right]\varepsilon_{it+1} - \gamma_0(1-\rho_x)x_t$$

Then,

$$\begin{aligned}
\mathbb{E}_t[r_{it+1}] &= r_{ft} + \gamma_0\left[\frac{\beta_i c_3}{R} + \frac{R-1}{R}(\beta_i A_1 - A_3)\right]\sigma_\varepsilon^2 - \frac{1}{2}Var_t[r_{it+1}] \\
Var_t[r_{it+1}] &= \left[\frac{\beta_i c_3}{R} + \frac{R-1}{R}(\beta_i A_1 - A_3)\right]^2 \sigma_\varepsilon^2 + \left[\frac{c_4}{R} + \frac{R-1}{R}\tilde{A}_1\right]^2 (\sigma_{\tilde{\varepsilon}}^2 - \beta_i^2\sigma_\varepsilon^2) \\
\mathbb{E}_t[R_{it+1}] &= e^{r_{ft} + \gamma_0\left[\frac{\beta_i c_3}{R} + \frac{R-1}{R}(\beta_i A_1 - A_3)\right]\sigma_\varepsilon^2} \\
&= e^{r_{ft} + \gamma\left[-\frac{R-1}{R}A_3 - \frac{(1-\rho_x)\gamma_0}{\rho_x - R} - \frac{R-1}{R}\frac{1}{\rho_x - R}A_4\right] + \left(\frac{R-1}{R}\right)\left[A_1 + \frac{\rho_x}{R-\varphi}(A_1 + A_2)\right]\gamma_0\beta_i\sigma_\varepsilon^2} \\
&= e^{r_{ft} + \frac{1}{\psi_0} + \frac{1}{\psi_1}\gamma_0\sigma_\varepsilon^2\beta_i}
\end{aligned}$$

where $\frac{1}{\psi_0} = -\gamma_0\left[\frac{R-1}{R}A_3 + \frac{(1-\rho_x)\gamma_0}{\rho_x - R} + \frac{R-1}{R}\frac{1}{\rho_x - R}A_4\right]\sigma_\varepsilon^2$ and $\frac{1}{\psi_1} = \left(\frac{R-1}{R}\right)\left[\frac{R}{R-\rho_x}A_1 + \frac{\rho_x}{R-\rho_x}A_2\right]$.

Then,

$$\log \mathbb{E}_t\left[R_{it+1}\right] = r_{ft} + \frac{1}{\psi_0} + \frac{1}{\psi_1}\gamma_0\beta_i\sigma_\varepsilon^2$$

and the expected excess return (relative to the firm with no risk, i.e., β of zero) is

$$\mathbb{E}_t\left[r_{it+1}^e\right] \equiv \log \mathbb{E}_t\left[R_{it+1}^e\right] = \frac{1}{\psi_1}\gamma_0\beta_i\sigma_\varepsilon^2$$

with cross-sectional variance

$$\sigma_{\mathbb{E}_t[r]}^2 = \left(\frac{1}{\psi_1}\right)^2 \sigma_\beta^2\left(\gamma_0\sigma_\varepsilon^2\right)^2$$

Properties of the multiplier. The multiplier that translated expected stock return dispersion into expected *mpk* dispersion is given by

$$\psi_1 = \left(\frac{R}{R-1}\right)\left[\frac{R}{R-\rho_x}A_1 + \frac{\rho_x}{R-\rho_x}A_2\right]^{-1} \tag{16}$$

Plugging in the definitions of objects yields this as a function of only four parameters, $(\rho, \rho_x, \theta, \delta)$:

$$\psi_1 = \left(\frac{1-\rho\rho_x}{1-\rho}\right)\frac{\frac{1}{\rho} + (1-\theta)\delta - 1}{\frac{1}{\rho} + \delta - 1} \tag{17}$$

The derivatives of the multiplier with respect to these parameters can be derived as:

$$
\begin{aligned}
\frac{\partial \psi_1}{\partial \rho_x} &= -\frac{\rho}{1-\rho} \frac{\frac{1}{\rho}+(1-\theta)\delta-1}{\frac{1}{\rho}+\delta-1} = -\frac{1-\rho\rho_x}{\rho}\psi_1 < 0 \\
\frac{\partial \psi_1}{\partial \delta} &= \left(\frac{1-\rho\rho_x}{1-\rho}\right) \frac{-\theta\left(\frac{1}{\rho}-1\right)}{\left(\frac{1}{\rho}+\delta-1\right)^2} = -\psi_1 \frac{\theta\left(\frac{1}{\rho}-1\right)}{\left(\frac{1}{\rho}+\delta-1\right)\left(\frac{1}{\rho}+(1-\theta)\delta-1\right)} < 0 \\
\frac{\partial \psi_1}{\partial \theta} &= \left(\frac{1-\rho\rho_x}{1-\rho}\right) \frac{-\delta}{\frac{1}{\rho}+\delta-1} = -\psi_1 \frac{\delta}{\frac{1}{\rho}+(1-\theta)\delta-1} \leq 0 \\
\frac{\partial \psi_1}{\partial \rho} &= \frac{\left(\delta^2(1-\theta)\right)(1-\rho_x) + \frac{1-\rho}{\rho^2}\left((1-\rho)(1-\delta)+(1+\rho)\delta(1-\theta)+\delta\theta\rho\rho_x\right)}{\left((1-\rho)\left(\frac{1}{\rho}+\delta-1\right)\right)^2} > 0
\end{aligned}
$$

Given our assumptions about the domain of parameters ($\rho, \theta \in (0,1)$, $\delta, \rho_x \in [0,1]$), we know that $\frac{\partial \psi_1}{\partial \rho_x}, \frac{\partial \psi_1}{\partial \delta}, \frac{\partial \psi_1}{\partial \theta}$ are all negative and $\frac{\partial \psi_1}{\partial \rho}$ is positive.[17] Since ψ_1 is continuous in these parameters on this domain, we can conclude ψ_1 is increasing in ρ and decreasing in ρ_x, δ, θ.

A.4 Other Distortions

With distortions, the firm's value function becomes

$$
V\left(X_t, Z_{it}, K_{it}, \xi_{it+1}\right) = \max_{K_{it+1}} e^{z_{it}+\beta_i x_t+\tau_{it}} K_{it}^{\theta} - K_{it+1} + (1-\delta) K_{it} + \mathbb{E}_t\left[M_{t+1} V\left(X_{t+1}, Z_{it+1}, K_{it+1}, \xi_{it+1}\right)\right]
$$

and the Euler equation

$$
1 = (1-\delta) E_t\left[M_{t+1}\right] + \alpha K_{it+1}^{\alpha-1} E_t\left[e^{m_{t+1}+z_{it+1}+\beta_i x_{t+1}+\tau_{it+1}}\right]
$$

Substituting the form of the distortion from (14) and following similar steps as the baseline case,

$$
\begin{aligned}
E_t\left[e^{m_{t+1}+z_{it+1}+\beta_i x_{t+1}+\tau_{it+1}}\right] &= \mathbb{E}_t\left[e^{m_{t+1}+z_{it+1}+\beta_i x_{t+1}+\nu(z_{it+1}+\beta_i x_{t+1})-\xi_{it+1}}\right] \\
&= \mathbb{E}_t\left[e^{\log\rho+(1+\nu)\rho_z z_{it}+((\beta_i(1+\nu)-\gamma_0)\rho_x+\gamma_0)x_t+((1+\nu)\beta_i-\gamma_0)\varepsilon_{t+1}+(1+\nu)\varepsilon_{it+1}-\xi_{it+1}}\right] \\
&= e^{\log\rho+(1+\nu)\rho_z z_{it}+((\beta_i(1+\nu)-\gamma_0)\rho_x+\gamma_0)x_t+\frac{1}{2}((1+\nu)\beta_i-\gamma_0)^2\sigma_{\varepsilon}^2+\frac{1}{2}(1+\nu)^2\left(\sigma_{\tilde{\varepsilon}}^2-\beta_i^2\sigma_{\varepsilon}^2\right)-\xi_{it+1}} \\
&= e^{\log\rho+(1+\nu)\rho_z z_{it}+((\beta_i(1+\nu)-\gamma_0)\rho_x+\gamma_0)x_t-(1+\nu)\gamma_0\beta_i\sigma_{\varepsilon}^2+\frac{1}{2}\gamma_0^2\sigma_{\varepsilon}^2+\frac{1}{2}(1+\nu)^2\sigma_{\tilde{\varepsilon}}^2-\xi_{it+1}}
\end{aligned}
$$

[17] $\frac{\partial \psi_1}{\partial \theta} < 0$ if $\delta \in (0,1)$, $\frac{\partial \psi_1}{\partial \theta} \leq 0$ if $\delta \in [0,1)$

so that

$$\theta K_{it+1}^{\theta-1} = \frac{1-(1-\delta)\,\mathbb{E}_t\left[M_{t+1}\right]}{e^{\log\rho+(1+\nu)\rho_z z_{it}+((\beta_i(1+\nu)-\gamma_0)\rho_x+\gamma_0)x_t-(1+\nu)\gamma_0\beta_i\sigma_\varepsilon^2+\frac{1}{2}\gamma_0^2\sigma_\varepsilon^2+\frac{1}{2}(1+\nu)^2\sigma_{\tilde{\varepsilon}}^2-\xi_{it+1}}}$$

and in logs,

$$k_{it+1} = \frac{1}{1-\theta}\left[\tilde{\alpha}_t + (1+\nu)\,\rho_z z_{it} + (1+\nu)\,\beta_i\rho_x x_t - (1+\nu)\,\beta_i\gamma_0\sigma_\varepsilon^2 - \xi_{it+1}\right]$$

where

$$\tilde{\alpha}_t = \log\theta - \log\left(1-(1-\delta)\,\mathbb{E}_t\left[M_{it+1}\right]\right) + \log\rho + (1-\rho_x)\,\gamma_0 x_t + \frac{1}{2}\gamma_0^2\sigma_\varepsilon^2 + \frac{1}{2}(1+\nu)^2\,\sigma_{\tilde{\varepsilon}}^2$$

Then,

$$\begin{aligned}
mpk_{it+1} &= \log\theta + y_{it+1} - k_{it+1} \\
&= \log\theta + z_{it+1} + \beta_i x_{t+1} - (1-\theta)\,k_{it+1} \\
&= \log\theta + z_{it+1} + \beta_i x_{t+1} - \tilde{\alpha}_t - (1+\nu)\,\rho_z z_{it} - (1+\nu)\,\beta_i\rho_x x_t + (1+\nu)\,\beta_i\gamma_0\sigma_\varepsilon^2 + \xi_{it+1} \\
&= \alpha_t + \varepsilon_{it+1} + \beta_i\varepsilon_{t+1} - \nu\rho_z z_{it} - \nu\beta_i\rho_x x_{it} + (1+\nu)\,\beta_i\gamma_0\sigma_\varepsilon^2 + \xi_{it+1}
\end{aligned}$$

where $\alpha_t = \log\theta - \tilde{\alpha}_t$.

Expected mpk is

$$\mathbb{E}_t\left[mpk_{it+1}\right] = \alpha_t - \nu\rho_z z_{it} - \nu\beta_i\rho_x x_{it} + (1+\nu)\,\beta_i\gamma_0\sigma_\varepsilon^2 + \xi_{it+1}$$

and the cross-sectional variance

$$\sigma_{\mathbb{E}_t[mpk]}^2 = \nu^2\rho_z^2\sigma_z^2 + \nu^2\rho_x^2 x_t^2\sigma_\beta^2 + (1+\nu)^2\,\sigma_\beta^2\left(\gamma_0\sigma_\varepsilon^2\right)^2 + \sigma_\xi^2$$

B Empirical Predictions

Computation of Betas and Expected Returns We compute stock market betas for equity returns by running time series regressions of excess equity returns on factor portfolio returns for each firm. We compute stock market betas using monthly returns and a two-year rolling window horizon. We also compute "MPK Betas" as an alternative measure of firm exposure to the aggregate shock, regressing firm $log(Y/K)$ on factor portfolio returns. As firm $log(Y/K)$ is only observed at the quarterly frequency, we use quarterly returns and a five-year rolling window horizon.

To compute expected returns, we first run Fama-Macbeth regressions to estimate market prices of risk for each factor. We then compute two measures of expected returns. We then compute $E_t[r_{i,t}()] = \sum_j \beta_{i,j}\lambda_j + \bar{\epsilon}_i$, where j denotes the factor, i the firm, λ_j is the market price of risk for factor j, and $\bar{\epsilon}_i$ is the average residual for firm i in the Fama-Macbeth regressions. We then compute $E_t[r_{i,t}(\beta)] = \sum_j \beta_{i,j}\lambda_j$, a measure of expected returns coming just from the factors and the market prices of risk (and not mis-pricing). The results reported in the body of the text consider the Fama-French model, but we have replicated qualitatively similar results using a number of other factor models.

Supplemental Tables Table 9 displays a cross-sectional regression of the standard deviation of *mpk* (within each industry) on the standard deviation of betas and expected returns (within each industry).

Table 9: Cross-sectional Industry Regression of *mpk* Dispersion

	(1) $\sigma(LMPK)$	(2) $\sigma(LMPK)$	(3) $\sigma(LMPK)$	(4) $\sigma(LMPK)$
$\sigma(E_{FF}[ret])$	2.994^{***}			
	(11.38)			
$\sigma(E_{FF,\beta}[ret])$		16.64^{***}		
		(11.26)		
$\sigma(\beta_{MKT})$			0.263^{***}	
			(2.68)	
$\sigma(\beta_{HML})$			0.170^{***}	
			(3.11)	
$\sigma(\beta_{SMB})$			0.217^{***}	
			(3.41)	
$\sigma(\beta_{MKT,MPK})$				0.252^{***}
				(4.41)
$\sigma(\beta_{HML,MPK})$				-0.0157
				(-0.39)
$\sigma(\beta_{SMB,MPK})$				0.145^{***}
				(5.84)
Constant	0.377^{***}	0.198^{***}	0.0500	0.251^{***}
	(9.26)	(3.55)	(0.79)	(4.62)
Observations	206	206	206	113
R^2	0.388	0.383	0.450	0.521

Notes: $\mathbb{E}_t[r_t]$ is the expected return computed from a Fama-Macbeth regression. $\mathbb{E}_t[r_t(\beta)]$ is the expected return predicted from the β's of that regression alone. β' denotes the stock return β on the FF factors and β_{MPK} the *mpk* β on the same factors. *t*-statistics are in parentheses. Significance levels are denoted by: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 10: Panel Industry Regression of *mpk* Dispersion, Year FE

	(1)	(2)	(3)	(4)
$\sigma(E[ret])$	2.244***			
	(27.35)			
$\sigma(E_{\beta}[ret])$		12.66***		
		(27.69)		
$\sigma(\beta_{MKT})$			0.267***	
			(10.36)	
$\sigma(\beta_{HML})$			0.107***	
			(7.33)	
$\sigma(\beta_{SMB})$			0.129***	
			(7.85)	
$\sigma(\beta_{CAPM,MPK})$				0.138***
				(6.32)
$\sigma(\beta_{HML,MPK})$				0.0961***
				(6.16)
$\sigma(\beta_{SMB,MPK})$				0.0703***
				(6.33)
Observations	2721	2746	2734	1427
R^2	0.219	0.221	0.275	0.291

Notes: $\mathbb{E}_t[r_t]$ is the expected return computed from a Fama-Macbeth regression. $\mathbb{E}_t[r_t(\beta)]$ is the expected return predicted from the β's of that regression alone. β' denotes the stock return β on the FF factors and β_{MPK} the *mpk* β on the same factors. t-statistics are in parentheses. Significance levels are denoted by: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$