Risk-Adjusted Capital Allocation and Misallocation

Joel M. David, Lukas Schmid, and David Zeke

December 18, 2020

Risk-Adjusted Capital Allocation and Misallocation[*]

Joel M. David[†]   Lukas Schmid[‡]   David Zeke[§]

FRB Chicago   USC Marshall   USC

December 18, 2020

Abstract

We develop a theory linking "misallocation," i.e., dispersion in marginal products of capital (MPK), to macroeconomic risk. Dispersion in MPK depends on (i) heterogeneity in firm-level risk premia and (ii) the price of risk, and thus is countercyclical. We document strong empirical support for these predictions. Stock market-based measures of risk premia imply that risk considerations explain about 30% of observed MPK dispersion among US firms and rationalize a large persistent component in firm-level MPK. Risk-based MPK dispersion, although not *prima facie* inefficient, lowers long-run aggregate productivity by as much as 6%, suggesting large "productivity costs" of business cycles.

JEL Classifications: D24, D25, E22, E32, G12, O47

Keywords: misallocation, productivity, costs of business cycles, risk premia

1 Introduction

A large and growing body of work has documented the "misallocation" of resources across firms, i.e., dispersion in the marginal product of inputs into production, and the resulting adverse effects on aggregate outcomes, such as productivity and output. Recent studies have found that even after accounting for a host of leading candidates – for example, adjustment costs, financial frictions or imperfect information – a substantial portion of observed misallocation seems to stem from other firm-specific factors, specifically, of a type that are orthogonal to firm fundamentals and are extremely persistent (if not permanent) to the firm.[1] Identifying exactly what – if any – underlying economic forces lead to this type of distortion has proven puzzling.

In this paper, we propose and quantitatively evaluate just such a theory, linking capital misallocation to macroeconomic risks. To the best of our knowledge, we are the first to make the connection between standard notions of the risk-return tradeoff and the resulting dispersion in the marginal product of capital (MPK). Indeed, our framework provides a natural way to translate firm-level financial market outcomes into the implications for the allocation of capital across firms. Further, we are able to quantify the effects of risk considerations – i.e., dispersion in risk premia and the extent of aggregate volatility (and so aggregate risk) – on macroeconomic variables, such as aggregate total factor productivity (TFP). Through the marginal product dispersion they induce, risk premium effects – though not *prima facie* inefficient – depress the achieved level of TFP, leading to a previously unexplored "productivity cost" of business cycles in the spirit of Lucas (1987).[2]

Our point of departure is a standard neoclassical model of firm investment in the face of both aggregate and idiosyncratic uncertainty. Firms discount future payoffs using a stochastic discount factor that is also a function of aggregate conditions. Critically, this setup implies that firms optimally equalize not necessarily MPK, but rather *expected, appropriately discounted*, MPK. With little more structure than this, the framework gives rise to a sharp condition governing the firm's expected MPK – firms with higher exposure to aggregate risk require a higher risk premium on investments, which translates into a higher expected MPK. In fact, the model implies an asset pricing equation of exactly the same form that is often used to price the cross-section of stock market returns. The equation simply states that a firm's expected MPK should be linked to the exposure of its MPK to aggregate risk (i.e., the firm's "beta"), and the "price" of that risk. This firm-specific risk premium appears exactly as what would otherwise be labeled a persistent distortion or "wedge" in the firm's investment decision.

We use a combination of firm-level production and stock market data to empirically in-

---

[1]See, e.g., David and Venkateswaran (2019). We discuss the literature in more detail below.

[2]Our analysis is also reminiscent of the approach in Alvarez and Jermann (2004), who use data on asset prices to measure the welfare costs of aggregate fluctuations.

vestigate and verify two key implications of this simple framework: (i) firm-level exposure to aggregate risk, measured using standard risk factors priced in financial markets, is an important determinant of expected MPK and (ii) MPK dispersion is increasing in the price of risk, measured using common proxies such as credit spreads and the aggregate price/dividend ratio. Because the price of risk is countercyclical, this link introduces a countercyclical element into MPK dispersion as observed in the data. We use the empirical results to perform a back-of-the-envelope calculation that points to an significant role of risk effects in generating MPK dispersion. Intuitively, the calculation relies on the facts that (i) dispersion in risk premia – readily measured using data on expected stock market returns – is large and (ii) regression estimates yield a sizable elasticity of firm-level MPK to expected returns.

After establishing these empirical results, we interpret them and gauge their magnitudes through the lens of a quantitative model. To that end, we enrich our theory by explicitly linking the sources of uncertainty to idiosyncratic and aggregate productivity risk.[3] We add two key elements to this framework: (i) a stochastic discount factor designed to match standard asset pricing facts and (ii) *ex-ante* cross-sectional heterogeneity in firm exposure, i.e., beta, with respect to the aggregate productivity shock. The profitability (e.g., productivity or demand) of high beta firms is highly sensitive to the realization of aggregate productivity (which captures the state of the business cycle), low beta firms have low sensitivity, and indeed, the profitability of firms with negative beta may move countercyclically. The investment side of the model is analytically tractable and yields sharp characterizations of firm investment decisions and MPK.

This setup is consistent with the key empirical results described above, namely, firm-level expected MPKs depend on exposures to the aggregate productivity shock (the systemic risk factor in the economy) and due to the countercyclical nature of the price of risk, the cross-sectional dispersion in expected MPK is also countercyclical. Further, we derive an expression for aggregate TFP, which is a strictly decreasing function of MPK dispersion. By inducing MPK dispersion, cross-sectional variation in exposure to aggregate risk and a higher price of risk (which depends on the degree of aggregate volatility) reduce the long-run (average) level of achieved TFP. Thus, the model provides a novel, quantifiable link between financial market conditions, i.e., the nature of aggregate risk, and longer-run macroeconomic performance.

The strength of these connections relies on three key parameters – the degree of heterogeneity in firm-level risk exposures and the magnitude and time-series variation in the price of risk. We devise an empirical strategy to identify these parameters using salient moments from firm-level and aggregate stock market data, specifically, (i) the cross-sectional dispersion in expected stock returns, (ii) the market equity premium and (iii) the market Sharpe ratio. We use a linearized

---

[3]These can also be interpreted as shocks to demand. Later, we show that the environment can be extended to incorporate multiple risk factors and financial shocks.

version of our model to derive analytical expressions for these moments and show that they are tightly linked to the structural parameters. The latter two pin down the level and volatility of the price of risk and the first identifies the cross-sectional dispersion in firm-level risk exposures. Indeed, in some simple cases of our model, the dispersion in expected MPK coming from risk premium effects is directly proportional to the dispersion in expected stock returns – intuitively, both of these moments are determined by cross-sectional variation in betas.

Before quantitatively evaluating this mechanism, we add other investment frictions to the environment, specifically, capital adjustment costs. Although they do not change the main insights from the simpler model, we uncover an important interaction between these costs and risk premia – namely, adjustment costs amplify the effects of beta variation on MPK dispersion. Intuitively, adjustment costs generate a second source of co-movement of firm outcomes with aggregate conditions – i.e., fluctuations in the value of installed capital – and hence add an additional component to the risk premium. Because the value of capital turns out to be more procyclical for high beta firms, the risk premium is in turn higher for these firms. On their own, adjustment costs have only transitory effects and do not lead to persistent dispersion in firm-level MPK. However, they can augment the effects of other factors that do, such as the heterogeneity in risk premia we analyze here.

We apply our methodology to data on US publicly traded firms from Compustat/CRSP. Our estimates reveal substantial variation in firm-level betas and a sizable price of risk – together, these imply a significant amount of risk-induced MPK dispersion. For example, our results suggest risk premium effects can explain as much as 30% of total observed MPK dispersion. Importantly, this dispersion is largely due to persistent MPK deviations at the firm-level, exactly of the type that compose a large portion of measured misallocation. Indeed, risk effects can account for as much as 47% of the permanent component in the data. The implications of these findings for the long-run level of aggregate TFP are significant – cross-sectional variation in risk reduces TFP by as much as 6%. Note that this represents a quantitative estimate of the impact of financial market outcomes on macroeconomic performance and further, a new connection between the nature of business cycle volatility and long-run outcomes in the spirit of Lucas (1987). Here, higher aggregate volatility leads to greater aggregate risk, increasing dispersion in required rates of return and MPK and thus reducing TFP. Our results suggest these "productivity costs" of business cycles may be substantial.

Our estimates also imply a significant countercyclical element in expected MPK dispersion. For example, the parameterized model produces a correlation between the cross-sectional variance in expected MPK and the state of the business cycle (measured by the aggregate productivity shock) of -0.31. To put this number in context, the correlation between MPK dispersion and aggregate productivity in the data is about -0.27. This result provides a risk-based

explanation for the puzzling observation, made forcefully by Eisfeldt and Rampini (2006), that capital reallocation is procyclical, despite the apparently countercyclical gains – due to the countercyclical nature of the price of risk and high beta of high MPK firms, such reallocation in downturns would require capital to flow to the riskiest of firms in the riskiest of times.

We pursue two important extensions of our baseline analysis. First, we add a flexible class of firm-specific "distortions" of the type that have been emphasized in the misallocation literature. These distortions can be fixed or time-varying and may be correlated or uncorrelated with firm-level characteristics (including betas) and with the state of the business cycle. We show that to a first-order approximation, our identification strategy and results are either unaffected by these distortions or are likely conservative (depending on the exact correlation structure of the distortions). These findings highlight an important feature of our empirical approach: although observed misallocation may stem from a wide variety of sources, our approach to measuring risk premium effects yields a robust estimate of the contribution of this one source alone. Second, we provide further, direct evidence on the extent of beta dispersion. Rather than relying on stock market data, we compute firm-level betas using production-side data by estimating time-series regressions of firm-level productivity on measures of aggregate productivity. The beta is the coefficient from this regression. This approach yields beta dispersion on par with the dispersion implied by the cross-section of stock market returns.

Why do firms (within an industry) have different exposure to the business cycle? Although our analysis does not require us to take a stand on this question, we explore a number of potential explanations. First, we investigate heterogeneity in production technologies (i.e., input elasticities) and markups. We show that these types of heterogeneity can indeed lead to variation in firm responsiveness/exposure to shocks but at most are likely to account for about 12% and 6% of the estimated standard deviation of betas, respectively. Although non-negligible, these results suggest that the majority of beta dispersion stems from other sources. Next, we show that theories of "trading down" over the business cycle as in, e.g., Jaimovich et al. (2019), may be a promising explanation. In times of economic expansion, when purchasing power is high, consumers tend to substitute towards higher quality goods while in downturns they substitute towards lower quality ones. Thus, higher quality products are more procyclical and lower quality ones less so. Although systematically quantifying this channel is challenging (such an analysis would require comprehensive product quality data), we provide evidence from a single industry – eating places – where we were able to obtain a proxy for quality, namely, average price. Low-price establishments tend to have lower betas than high-price ones and further, average price is positively related to betas, expected returns and MPK. This case study of a single industry helps illustrate the main relationships implied by our theory and suggests quality differentiation may be an important factor behind differences in firm cyclicality.

Related Literature. Our paper relates to several branches of literature. Foremost is the large body of work investigating resource misallocation, seminal examples of which include Hsieh and Klenow (2009) and Restuccia and Rogerson (2008). A number of recent papers have explored the role of financial frictions, for example, Midrigan and Xu (2014), Moll (2014) and Buera et al. (2011) study collateral constraints and Gilchrist et al. (2013) firm-specific borrowing costs. Gopinath et al. (2017) and Kehrig and Vincent (2017) study the interaction of financial frictions and adjustment costs in explaining recent dynamics of misallocation in Spain and within firms, respectively. 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 agents. The addition of aggregate risk is a key innovation of our analysis – existing work has typically abstracted from this channel (either by assuming no aggregate uncertainty or risk-neutral agents). We show that the link between aggregate risk and observed misallocation is quite tight in the presence of heterogeneous exposures to that risk.[4] Papers studying additional candidates behind marginal product dispersion include Peters (2016), Edmond et al. (2018) and Haltiwanger et al. (2018) who focus on markup dispersion, David et al. (2016) information frictions and Asker et al. (2014) capital adjustment costs.[5] David and Venkateswaran (2019) provide an empirical methodology to disentangle various sources of capital misallocation and establish a large role for highly persistent firm-specific factors. In our theory, firm-level risk premia manifest themselves as persistent firm-specific "wedges" of exactly this type.

Kehrig (2015) documents in detail the countercyclical nature of productivity dispersion. We build on this finding by relating fluctuations in MPK dispersion to time-series variation in the price of risk. A growing literature, starting with Eisfeldt and Rampini (2006), investigates the procyclical nature of capital reallocation, which is puzzling since higher cross-sectional dispersion in MPK in downturns should lead capital to flow to highly productive, high MPK firms in recessions. Our results bear on that observation by noting that the countercyclicality of the price of risk, in conjunction with heterogeneity in firm-level risk exposures, goes some way toward reconciling this puzzle.

Our work also relates to a large literature exploring the link between financial market re-

---

[4]Gilchrist et al. (2013) find a limited role for firm-specific borrowing costs. Where they focus mainly on costs of debt, we find a larger role for differences in costs of equity, which is an important source of financing for the firms in our data (for example, the average leverage ratio in our sample is 0.28). Indeed, in our simple theory, the Modigliani-Miller theorem holds, i.e., all firms are able to borrow at the common risk-free rate and thus there is no dispersion in borrowing costs. Yet equity costs – and so total costs of capital – may differ widely. One contribution of our work is extending the insights in Gilchrist et al. (2013) to a broader notion of financing costs and showing that the implications for misallocation can be quite different.

[5]Many papers study the role of firm-specific distortions, e.g., Bartelsman et al. (2013). Restuccia and Rogerson (2017), Hopenhayn (2014) and Eisfeldt and Shi (2018) provide excellent overviews of recent work on capital misallocation/reallocation.

turns and the return to capital ("investment returns"). Cochrane (1991), Restoy and Rockinger (1994) and Balvers et al. (2015) show that stock returns and investment returns are closely linked (indeed, exactly coincide under constant returns to scale). Recent work builds on this insight to examine the cross-section of stock returns from the perspective of investment returns, interpreting common risk factors through firms' investment policies and showing that investment-based factors are priced in the cross-section of returns, e.g., Zhang (2005), Gomes et al. (2006), Liu et al. (2009) and Zhang (2017). We examine investment returns and the marginal product of capital as a joint manifestation of risk premia, most readily measured through stock returns, and extend this connection to analyze the implications for the allocation of capital and macroeconomic outcomes, such as aggregate TFP. Binsbergen and Opp (2017) also investigate the implications of asset market considerations for the real economic decisions of firms. They propose a framework where distortions in agents' subjective beliefs lead to "alphas," i.e., cross-sectional mis-pricings, and real efficiency losses, whereas we focus on the MPK dispersion induced by heterogeneity in aggregate risk exposures. Our empirical work establishes a connection between MPK and financial market outcomes and our quantitative work uses a workhorse macroeconomic model of firm dynamics augmented with risk-sensitive agents and aggregate risk to evaluate the implications of this connection. One of our key messages shares a common theme with this line of work – financial market considerations can have sizable effects on real outcomes by affecting capital allocation decisions.[6]

2 Motivation

In this section, we lay out a simple version of the standard, frictionless neoclassical theory of investment to illustrate the main insight of our analysis, namely, the link between firm-level MPK and risk premia. We use this framework to motivate a number of empirical exercises exploring this connection and guide a simple back-of-the-envelope calculation that suggests a significant role for risk in generating MPK dispersion.[7] Section 3 enriches this environment along several dimensions for purposes of our quantitative work.

2.1 MPK and Risk Premia

Firms produce output using capital and labor according to a Cobb-Douglas technology and face constant (or infinite) elasticity demand curves. 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,

---

[6]Relatedly, David et al. (2014) find that risk considerations play an important role in determining the allocation of capital across countries, i.e., can explain some portion of the "Lucas Paradox."

[7]All derivations for this section are in Appendix A.

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 $\Pi_{it} = \Pi_{it}(X_t, Z_{it}, K_{it})$ denote the operating profits of the firm – revenues net of labor costs – where X_t and Z_{it} denote aggregate and idiosyncratic shocks to firm profitability, respectively, and K_{it} the firm's stock of capital. Both X_t and Z_{it} may be vectors, i.e., there may be multiple sources of both idiosyncratic and aggregate risk. The analysis can accommodate a number of interpretations of the fundamental shocks, for example, as productivity or demand shifters. With these assumptions, the profit function takes a Cobb-Douglas form, is homogeneous in K of degree $\theta < 1$ (due to curvature in production and/or demand) and is proportional to revenues. The marginal product of capital is equal to $MPK_{it} = \theta \frac{\Pi_{it}}{K_{it}}$. 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 is correlated with the aggregate shock(s), X_t. We can write the firm's problem recursively as

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

where $\mathbb{E}_t [\cdot]$ denotes the firm's conditional expectations. The Euler equation is given by

$$1 = \mathbb{E}_t [M_{t+1} (MPK_{it+1} + 1 - \delta)] \quad \forall\, i, t . \tag{2}$$

MPK dispersion. An immediate implication of expression (2) is that MPK (or even expected MPK) need not be equated across firms; rather, it is only expected, appropriately discounted MPK that is equalized. To the extent that firms' MPK co-move differently with the SDF, their expected MPK will differ. From here, we can derive the following equation for expected MPK:

$$\mathbb{E}_t [MPK_{it+1}] = \alpha_t + \beta_{it} \lambda_t . \tag{3}$$

Here, $\alpha_t = r_{ft} + \delta$ is the risk-free user cost of capital, where r_{ft} is the (net) risk-free interest rate, $\beta_{it} \equiv -\frac{\text{cov}_t(M_{t+1}, MPK_{it+1})}{\text{var}_t(M_{t+1})}$ captures the elasticity, or exposure, of the firm's MPK to movements in the SDF – i.e., the riskiness of the firm – and $\lambda_t \equiv \frac{\text{var}_t(M_{t+1})}{\mathbb{E}_t[M_{t+1}]}$ is the market price of that risk. Expression (3) illustrates the first main insight: expected MPK is not necessarily common across firms, but rather is a function of the (common) risk-free rate and a firm-specific risk premium, which depends on the firm's beta 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

$$\sigma^2_{\mathbb{E}_t[MPK_{it+1}]} = \sigma^2_{\beta_t} \lambda^2_t , \tag{4}$$

where $\sigma_{\beta_t}^2$ is the cross-sectional variance of time t betas. Expression (4) reveals the second main insight: the extent to which risk considerations lead to dispersion in expected MPK is increasing in the price of risk (and in the cross-sectional variation in risk exposures). A key observation underlying our analysis is that financial market data imply that risk prices are high (e.g., a large equity premium and observed Sharpe ratios on various investment strategies), suggesting a potentially important role for differences in risk exposure in leading to MPK dispersion. Further, given persistence in firm-level betas, the theory implies persistent differences in firm-level MPK – as observed in the data – driven by dispersion in required rates of return.[8]

Examples. It is useful to consider a few concrete illustrative examples:

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 (3) and (4) 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 additional frictions, expected MPK should be equalized across firms.[9] The same result 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: CRRA preferences. In the case of CRRA utility with coefficient of relative risk aversion γ, standard approximation techniques give

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

where Δc_{t+1} denotes log consumption growth. Expected MPK is determined by the covariance of the firm's MPK with consumption growth. The price of risk is the product of the coefficient of relative risk aversion and the conditional volatility of consumption growth.

Example 3: CAPM. In the CAPM, the SDF is linearly related to the return on the aggregate stock market, i.e., $M_{t+1} = a - b r_{mt+1}$ for some constants a and b. Because the market portfolio

---

[8]To see this more clearly, we can take the unconditional expectation of equation (3) to obtain an approximate expression for the variance of firms' mean MPKs as $\sigma_{\mathbb{E}[MPK]}^2 \approx \sigma_\beta^2 \lambda^2$, where $\sigma_\beta^2 \equiv \sigma_{\mathbb{E}[\beta_{it}]}^2$ denotes the variance of unconditional betas and $\lambda \equiv \mathbb{E}\left[\lambda_t\right]$ the unconditional expectation of the price of risk. The approximation is valid as long as $\text{cov}\left(\beta_i, \text{cov}\left(\beta_{it}, \lambda_t\right)\right)$ is small. In line with the results in Lewellen and Nagel (2006), we find the time-series variation in betas to be quite modest. Further, they are persistent (for example, we find that CAPM betas have an implied one-year autocorrelation of 0.87). In the case of constant betas or if β_{it} is orthogonal to λ_t, the expression is exact.

[9]With time-to build for capital and uncertainty over upcoming shocks, there may still be dispersion in *realized* MPK, but not in *expected* terms, and so these forces do not lead to persistent firm-level MPK deviations.

is itself an asset with $\beta = 1$, it is straightforward to derive

$$\mathbb{E}_t \left[MPK_{it+1} \right] = \alpha_t + \underbrace{\frac{\mathrm{cov}_t \left(r_{mt+1}, MPK_{it+1} \right)}{\mathrm{var}_t \left(r_{mt+1} \right)}}_{\beta_{it}} \underbrace{\mathbb{E}_t \left[r_{mt+1} - r_{ft} \right]}_{\lambda_t},$$

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

Extensions. As can be seen from this simple framework, the link between firm-level MPK and risk is quite general and does not depend on specific assumptions about the SDF. In Appendix A.2, we provide two additional examples to show that the connection holds under alternative assumptions on the sources of risk premia. Specifically, we study versions where firm-specific risk is due to (i) firm-specific distortions to the pricing of risk (which show up as "alphas" or "mis-pricing") and (ii) heterogeneity in firm exposures to cyclical fluctuations in the price of investment goods (in the spirit of, e.g., Kogan and Papanikolaou (2013)). In both cases, differences in firm-level risk premia leads to MPK dispersion exactly as above, i.e, an expression analogous to (3) continues to hold. The main message from these excercises is that the link between MPK and risk is quite robust and does not depend on the precise source of differences in risk premia (although the normative implications, e.g., whether the resulting MPK dispersion is efficient or not – and so represents a true "misallocation" – may).

In our quantitative model in Section 3, we add capital adjustment costs, which lead to endogenous fluctuations in the value of installed capital, i.e., Tobin's Q, and study the role of additional factors (e.g., firm-level distortions) that determine expected MPK as emphasized in the misallocation literature (e.g., Hsieh and Klenow (2009)). In these cases, equations (3) and (4) do not hold exactly, since there are now additional forces generating MPK dispersion, but we develop an empirical strategy that accurately captures the portion of this dispersion that comes from risk effects alone, even in the presence of these other factors.

2.2 Empirical Evidence

In this section, we explore the two key implications of the simple framework laid out thus far: (i) exposure to aggregate risk – and hence, the firm-specific risk premium – is a determinant of expected MPK and (ii) MPK dispersion is increasing in the price of risk.

Measuring risk premia. In addition to influencing firms' capital choices, exposure to aggregate risk affects firms' stock returns. Indeed, in our model, expected excess stock returns

over the risk-free rate (i.e., the predictable component of the excess return) reflects compensation for risk exposure and thus represents a direct measure of the firm-level risk premium. We can exploit this link to use well-studied measures of firm risk exposures and risk premia from stock market data to explore the connection with MPK. To motivate this approach, we derive the following approximate expressions for firm-level expected excess MPK (over the user cost), denoted MPK_{it+1}^e, and excess stock market returns, R_{it+1}^e:

$$Empk_{it+1}^e \equiv \log \mathbb{E}_t\left[MPK_{it+1}^e\right] = -\mathrm{cov}_t\left(mpk_{it+1}, m_{t+1}\right) \tag{5}$$

$$Er_{it+1}^e \equiv \log \mathbb{E}_t\left[R_{it+1}^e\right] = -\psi\mathrm{cov}_t\left(mpk_{it+1}, m_{t+1}\right) , \tag{6}$$

where ψ is a constant and lowercase denotes natural logs. To a first-order approximation, expected stock returns and MPK are proportional, since they are jointly determined by the underlying risk characteristics of the firm.[10] Combining, we have

$$Empk_{it+1}^e = \frac{1}{\psi}Er_{it+1}^e \quad \Rightarrow \quad \mathrm{var}\left(Empk_{it+1}^e\right) = \left(\frac{1}{\psi}\right)^2 \mathrm{var}\left(Er_{it+1}^e\right) , \tag{7}$$

which shows that (in logs) the cross-sectional dispersion of expected MPK is proportional to the cross-sectional dispersion of expected stock market returns. We can use expression (7) to (i) verify the link between firm-level MPK and risk premia, measured from expected stock returns and (ii) calculate a back-of-the-envelope estimate of the MPK dispersion that may stem from heterogeneous risk premia.

Firm-level MPK and risk. Expression (7) reveals a tight link between expected MPK and expected stock returns. However, empirically implementing this equation is challenging since neither expected MPK nor expected returns are directly observable. One approach is to proxy these variables with their realized values, but this may be problematic: since both variables respond to the same realizations of shocks, a positive relationship may be mechanical and not indicative of the relationship between expected values. To overcome these hurdles, we follow a two-stage instrument variables approach in which we instrument for returns using common measures of risk exposure. First, in a preliminary step, we estimate time-varying risk exposures from backwards-looking rolling window regressions of individual firm stock returns on aggregate risk factors. Then, in the first stage, we estimate period-by-period cross-sectional regressions of realized stock market returns on the estimated exposures. The predicted values from these regressions yield measures of expected excess returns, i.e., risk premia, that are driven only

---

[10]The proportionality is exact under a single source of aggregate risk. A slightly modified expression holds with multiple sources.

by exposure to the aggregate risk factors considered.[11] Next, we estimate the second stage regression implied by equation (7) using a panel specification with these predicted values as instruments. Because the focus in the misallocation literature is generally on within-industry dispersion in MPK (MPK may vary across industries due to heterogeneity on a number of additional dimensions), we include a full set of industry-by-year fixed effects. We perform this procedure using three different sets of aggregate risk factors, taken from three common asset pricing models: the CAPM, the Fama-French 3 Factor model, and the Hou et al. (2015) q^5 5 factor model. In the CAPM, the aggregate market return is the single source of aggregate risk. The latter two models add returns on additional portfolios of firms sorted by a number of different characteristics. We provide further details in Appendix B.1.

In principle, once we have a measure of expected returns, proxying expected MPK with realized MPK should be innocuous, since the forecast errors in future realizations of MPK – the difference between the two measures – should be uncorrelated with backwards-looking data on stock returns. However, for completeness, we also construct a measure of expected MPK as follows: under our assumptions, realized (log) MPK is given by $mpk_{it} = y_{it} - k_{it}$ (we suppress constants that play no role), where y_{it} denotes (log) revenue. We assume that firm-level productivity, equal to $a_{it} = y_{it} - \theta k_{it}$, follows an AR(1) process with persistence ρ_a. Then, expected MPK is given by $\mathbb{E}_t[mpk_{it+1}] = \mathbb{E}_t[y_{it+1}] - k_{it+1} = \rho_a a_{it} - (1 - \theta) k_{it+1}$. Consistent with our estimates in Section 4 below, we use $\rho_a = 0.94$ and $\theta = 0.62$.[12]

We implement this approach using firm-level data from the Center for Research in Security Prices (CRSP) and Compustat.[13] We measure firm capital stock, K_{it}, as the (net of depreciation) value of property, plant and equipment and firm revenue, Y_{it}, as reported sales.[14] Full details of the data sample are in Appendix B.1.

Table 1 reports results from regressions of MPK and expected MPK on lagged expected stock market returns. The left-hand panel shows that each of the asset pricing models we consider yields a statistically significant relationship between expected returns and future realized MPK.[15] The economic magnitudes are also significant: the estimated coefficients imply that that a 100 basis point increase in the expected return is associated with a 3-5% increase in

---

[11]Because the estimates of risk exposures use data only through time t, they forecast only the expected component of realized returns at time $t + 1$.

[12]A simplifying assumption here is that the aggregate and idiosyncratic components of a_{it} have the same persistence. Our quantitative work in Section (4) suggests this is a reasonable approximation, where we find the persistences of the two components to be quite close.

[13]By using Compustat data, our analysis focuses on large, publicly traded firms for which financial market data are available. Extending the analysis to private firms would be a valuable exercise, but faces a significant challenge in deriving accurate measures of risk premia (our alternative approach in Section 4.4 may be one way).

[14]Using book assets, a broader notion of firm capital, yields similar results.

[15]Standard errors are two-way clustered by firm and year, but do not account for estimation error in measured risk exposures. In Appendix C, we perform a bootstrapping procedure designed to address this issue.

Table 1: MPK and Risk Premia

	mpk			$\mathbb{E}[mpk]$		
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{E}_{capm}[ret]$	4.798***			5.388***		
	(4.08)			(4.62)		
$\mathbb{E}_{ff3}[ret]$		3.284***			4.159***	
		(6.80)			(8.70)	
$\mathbb{E}_{q^5}[ret]$			5.016***			6.299***
			(8.22)			(10.71)

Notes: This table reports results from a two-step procedure in which we estimate the elasticity of firm-level MPK to expected excess stock returns. In the first stage, we instrument for expected returns using measures of firm risk exposures from stock market data. In the second stage, we run regressions of the form (7) with industry-by-year fixed effects. Standard errors are two-way clustered by firm and year. *t*-statistics in parentheses. Significance levels are denoted by: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

MPK. The right-hand panel reports analogous results using measures of expected, rather than realized, MPK. The estimated coefficients are similar, and if anything, slightly larger, though the differences are not statistically significant. Thus, our two-stage approach reveals a significant relationship between firm-level risk premia, instrumented using common measures of risk exposure from stock market data, and MPK.

The results in Table 1 suggest a simple back-of-the-envelope calculation of the role of risk premia in driving MPK dispersion. As an example, consider the Fama-French 3 factor model, perhaps the most widely used framework to study the cross-section of stock returns. Using this model to estimate expected returns yields a within-industry standard deviation of about 0.127 (details in Appendix B.2). Multiplying this value by the estimated elasticity of MPK to expected returns in Table 1 (3.3) and squaring yields a predicted cross-sectional variance of *mpk* of about 0.17. The total within-industry variance of *mpk* in the Compustat sample is about 0.45, implying that variation in risk premia can account for almost 40% of the total. A similar calculation using expected MPK gives a slightly larger share.

Although suggestive, this calculation ignores a number of factors that may affect the relationship between MPK and returns, e.g., possible nonlinearities, adjustment costs and/or other investment distortions of the type that have been emphasized in the misallocation literature. In Sections 3 and 4, we develop a quantitative model and empirical strategy that incorporates these additional considerations and allows for counterfactual decompositions of the role of risk effects in driving observed MPK dispersion. The results from the more general model turn out to be broadly in line with (though slightly smaller than) the simple calculation performed here.

MPK dispersion and the price of risk. Expression (4) illustrates a second key implication of the simple framework: MPK dispersion is positively related to the price of risk. To explore

Table 2: MPK Dispersion and the Price of Risk

	$\sigma\left(mpk_{it+1}\right)$			$\sigma\left(\mathbb{E}_t\left[mpk_{it+1}\right]\right)$		
	(1)	(2)	(3)	(4)	(5)	(6)
Excess Bond Premia	0.357***			0.633***		
	(3.08)			(6.89)		
GZ Spread		0.191***			0.338***	
		(2.75)			(5.68)	
PD Ratio			-0.176***			-0.228***
			(-3.51)			(-4.09)
Observations	174	174	174	174	174	174
R-squared	0.0644	0.0653	0.0660	0.218	0.221	0.117

Notes: This table reports time-series regressions of four-quarter ahead *mpk* dispersion on measures of the price of risk. *t*-statistics are in parentheses, computed using Newey-West standard errors. Significance levels are denoted by: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

this prediction, we estimate time-series regressions of the form

$$\sigma\left(mpk_{t+1}\right) = \psi_0 + \psi_1\lambda_t + \zeta_{t+1},$$

where λ_t denotes three different proxies for the price of risk: the price/dividend (PD) ratio on the aggregate stock market and two measures of credit spreads – the Gilchrist and Zakrajsek (2012) (GZ) spread, a high-information and duration-adjusted measure of the mean credit spread and the excess bond premium, which measures the portion of the GZ spread not attributable to default risk.[16] Table 2 reports regressions of the within-industry standard deviations of MPK (left-hand panel) and expected MPK (right-hand panel) on the lagged values of these measures.[17] Each predicts MPK dispersion, and in the direction the theory suggests: the GZ Spread and excess bond premium predict greater MPK dispersion, while a higher PD ratio predicts lower dispersion. Because the measures of the price of risk are countercyclical, the results imply that time-series variation in risk premia induce a countercyclical component in MPK dispersion, in line with (and potentially in part accounting for) the well known evidence of countercyclicality documented in Eisfeldt and Rampini (2006) and Kehrig (2015).[18]

---

[16]We extract the cyclical component of the PD ratio and MPK dispersion using a one-sided Hodrick-Prescott filter. The credit spread measures do not exhibit significant longer-term trends.

[17]Within-industry standard deviations are calculated by first de-meaning *mpk* (expected *mpk*) by industry-year. To control for the changing composition of firms, for each quarter, we include only firms that were present in the previous quarter and calculate changes in the standard deviation for these firms. We then use those changes to construct a composition-adjusted series that is unaffected by new additions or deletions from the dataset. Details of this procedure are in Appendix B.1.

[18]We report the correlations of these measures with de-trended GDP and TFP in Table 6 in Appendix B.3.

Additional exercises. In Appendix C, we explore a number of additional implications of our framework. We use MPK-sorted portfolios of firms to show that high MPK firms tend to offer higher expected stock market returns. We verify that firm-level MPK is related to direct measures of the sensitivity of MPK (rather than stock returns) to aggregate risk, as suggested by (3). We show that industries with greater MPK dispersion also tend to be those with more dispersion in expected returns and heterogeneity in firm-level risk exposures.

3   Quantitative Model

In this section, we use a more detailed version of the investment model laid out above to quantitatively investigate the contribution of heterogeneous risk premia to observed MPK dispersion. The model is kept deliberately simple in order to isolate the role of our basic mechanism, namely dispersion in exposure to macroeconomic risk. The theory consists of two main building blocks: (i) a stochastic discount factor, which we directly parameterize to be consistent with salient patterns in financial markets, i.e., high and countercyclical prices of risk and (ii) 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 us to sidetrack challenges with generating empirically relevant risk prices in general equilibrium, and focus on gauging the quantitative strength of our mechanism. To hone in on the effects of risk premia, we begin with a simplified version in which we abstract from additional adjustment frictions. In this case, our framework yields exact closed form solutions for firm investment decisions. In Section 3.2, we extend the model to include capital adjustment costs. Our theoretical results there reveal an important amplification effect of these costs on the impact of risk premia.[19]

Heterogeneity in risk exposures. The setup is a fleshed-out version of that in Section 2. We consider a discrete time, infinite-horizon economy. A continuum of firms of fixed measure one, indexed by i, produce a homogeneous good using capital and labor according to:

$$Y_{it} = X_t^{\hat{\beta}_i} \hat{Z}_{it} K_{it}^{\theta_1} N_{it}^{\theta_2}, \quad \theta_1 + \theta_2 < 1 . \tag{8}$$

Firm productivity (in logs) is equal to $\hat{\beta}_i x_t + \hat{z}_{it}$, where x_t denotes an aggregate component that is common across firms and $\hat{\beta}_i$ captures the exposure of the productivity of firm i to aggregate conditions. We assume that $\hat{\beta}_i$ is distributed as $\hat{\beta}_i \sim \mathcal{N}\left(\bar{\hat{\beta}}, \sigma_{\hat{\beta}}^2\right)$ across firms. Heterogeneity in this exposure is a key ingredient of our framework – cross-sectional variation

---

[19]We also consider the effects of other investment frictions, e.g., "wedges," or distortions, in Section 4.3.

in $\hat{\beta}_i$ will lead directly to dispersion in expected MPK. The term $\hat{z}_{it}$ denotes a firm-specific, idiosyncratic component of productivity.[20]

The two productivity components follow AR(1) processes (in logs):

$$
\begin{aligned}
x_{t+1} &= \rho_x x_t + \varepsilon_{t+1}, \quad \varepsilon_{t+1} \sim \mathcal{N}\left(0, \sigma_\varepsilon^2\right) \\
\hat{z}_{it+1} &= \rho_z \hat{z}_{it} + \hat{\varepsilon}_{it+1}, \quad \hat{\varepsilon}_{it+1} \sim \mathcal{N}\left(0, \sigma_{\hat{\varepsilon}}^2\right) .
\end{aligned}
\tag{9}
$$

Thus, there are two sources of uncertainty at the firm level – aggregate uncertainty, with conditional variance σ_ε^2, and idiosyncratic uncertainty, with variance $\hat{\sigma}_{\hat{\varepsilon}}^2$.

Stochastic discount factor. In line with the large 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 specify the SDF as

$$
\begin{aligned}
\log M_{t+1} &\equiv m_{t+1} = \log \rho - \gamma_t \varepsilon_{t+1} - \frac{1}{2}\gamma_t^2 \sigma_\varepsilon^2 \\
\gamma_t &= \gamma_0 + \gamma_1 x_t ,
\end{aligned}
\tag{10}
$$

where ρ, $\gamma_0 > 0$ and $\gamma_1 \leq 0$ are constant parameters.[21] The SDF is determined by shocks to aggregate productivity. The conditional volatility of the SDF, $\sigma_m = \gamma_t \sigma_\varepsilon$, varies through time as determined by γ_t. This formulation allows us to capture in a simple manner a high, time-varying and countercyclical price of risk as observed in the data (since $\gamma_1 < 0$, γ_t is higher following economic contractions, i.e., when x_t is negative). Additionally, directly parameterizing γ_0 and γ_1 enables the model to be quantitatively consistent with key moments of asset returns, which are important for our analysis. The risk free rate is constant and equal to $-\log \rho$. Thus, γ_0 and γ_1 only affect the properties of equity returns., easing the interpretation of these parameters. The maximum attainable Sharpe ratio is equal to the conditional standard deviation of the SDF, i.e., $SR_t = \gamma_t \sigma_\varepsilon$, and the price of risk is equal to the square of the Sharpe ratio, $\gamma_t^2 \sigma_\varepsilon^2$.

For simplicity, the setup thus far features (i) a single source of aggregate risk and (ii) a tight link between financial market conditions (i.e., γ_t) and macroeconomic conditions (i.e., x_t). In Appendix F we extend this framework to (i) include multiple risk factors and (ii) decouple movements in financial and macroeconomic conditions by including pure financial shocks that affect the price of risk but otherwise do not impact firm profits/productivity. Similar insights from the simpler model go through under those extensions.

---

[20]More broadly, expression (8) should be thought of as a revenue-generating function and the "productivity" components as also capturing demand factors, see, e.g., Section 5.

[21]This specification builds closely on those in, for example, Zhang (2005), Gomes and Schmid (2010) and Jones and Tuzel (2013).

Input choices. Firms hire labor period-by-period at a competitive wage, W_t. To keep the labor market simple, we assume that the equilibrium wage is given by

$$W_t = X_t^\omega \ ,$$

i.e., the wage is a constant elasticity and increasing function of aggregate productivity, where $\omega \in [0,1]$ determines the sensitivity of wages to aggregate conditions.[22] Maximizing over the static labor decision gives operating profits – revenues less labor costs – as

$$\Pi_{it} = G X_t^{\beta_i} Z_{it} K_{it}^\theta \ , \tag{11}$$

where $G \equiv (1-\theta_2)\, \theta_2^{\frac{\theta_2}{1-\theta_2}}$, $\beta_i \equiv \frac{1}{1-\theta_2}\left(\hat{\beta}_i - \omega\theta_2\right)$, $Z_{it} \equiv \hat{Z}_{it}^{\frac{1}{1-\theta_2}}$ and $\theta \equiv \frac{\theta_1}{1-\theta_2}$. The exposure of firm profits to aggregate conditions is captured by β_i, which is a simple transformation of the underlying exposure of firm productivity to the aggregate component, $\hat{\beta}_i$, and the sensitivity of wages, ω.[23] The idiosyncratic component of productivity is similarly scaled, by $\frac{1}{1-\theta_2}$. The curvature of the profit function is equal to θ, which depends on the relative elasticities of capital and labor in production. These scalings reflect the leverage effects of labor liabilities on profits. From here on, we will primarily work with z_{it}, which has the same persistence as $\hat{z}_{it}$, i.e., ρ_z, and innovations $\varepsilon_{it+1} = \frac{1}{1-\theta_2}\hat{\varepsilon}_{t+1}$ with variance $\sigma_\varepsilon^2 = \left(\frac{1}{1-\theta_2}\right)^2 \sigma_{\hat{\varepsilon}}^2$. We will also use the fact that $\sigma_\beta^2 = \left(\frac{1}{1-\theta_2}\right)^2 \sigma_{\hat{\beta}}^2$. Notice that the profit function takes precisely the form assumed in Section 2. Thus, the firm's dynamic investment problem takes the form in expression (1).

Optimal investment. The simplicity of this setting leads to exact analytical expressions for the firm's investment decision. Specifically, we show in Appendix D.1 that the firm's optimal investment policy is given by:

$$k_{it+1} = \frac{1}{1-\theta}\left(\tilde{\alpha} + \beta_i\rho_x x_t + \rho_z z_{it} - \beta_i\gamma_t\sigma_\varepsilon^2\right) \ , \tag{12}$$

where $\tilde{\alpha} \equiv \log\theta + \log G - \alpha$, $\alpha \equiv \log(r_f + \delta)$.[24] The firm's choice of capital is increasing in x_t and z_{it} due to their direct effect on expected future productivity (i.e., $\beta_i\rho_x x_t + \rho_z z_{it} = \mathbb{E}_t\left[\beta_i x_{t+1} + z_{it+1}\right]$), but, *ceteris paribus*, firms with higher betas choose a lower level of capital. The magnitude of this effect is larger when γ_t is large, i.e., in economic downturns. Clearly,

---

with risk neutrality, i.e., $\gamma_0 = \gamma_1 = 0$, the last term is zero and investment is purely determined by expected productivity.

For a slightly different intuition, we substitute for γ_t and write the expression as

$$k_{it+1} = \frac{1}{1-\theta} \left(\tilde{\alpha} + \beta_i \left(\rho_x - \gamma_1 \sigma_\varepsilon^2 \right) x_t + \rho_z z_{it} - \beta_i \gamma_0 \sigma_\varepsilon^2 \right) . \tag{13}$$

The risk premium affects the capital choice through both the time-varying and constant components of the price of risk: first, a more negative γ_1 increases the responsiveness of firms to aggregate conditions. Intuitively, a high (low) realization of x_t has two effects – first, since x_t is persistent, it signals that productivity is likely to be high (low) in the future, increasing (decreasing) investment (this force is captured by the ρ_x term). Moreover, a high (low) realization of x_t implies a low (high) price of risk, which further increases (decreases) investment. Second, the constant component of the risk premium, γ_0, adds a firm-specific constant – i.e., a fixed effect – which leads to permanent dispersion in firm-level capital.

MPK dispersion. By definition, the realized mpk is given by $mpk_{it+1} = \log \theta + \pi_{it+1} - k_{it+1}$. Substituting for k_{it+1},

$$mpk_{it+1} = \alpha + \varepsilon_{it+1} + \beta_i \varepsilon_{t+1} + \beta_i \gamma_t \sigma_\varepsilon^2 , \tag{14}$$

and taking conditional expectations,

$$Empk_{it+1} \equiv \mathbb{E}_t \left[mpk_{it+1} \right] = \alpha + \beta_i \gamma_t \sigma_\varepsilon^2 , \tag{15}$$

where α is as defined in equation (12) and reflects the risk-free user cost of capital. Expression (14) shows that dispersion in the realized mpk can stem from uncertainty over the realization of shocks, as well as the risk premium term, which is persistent at the firm level and depends on (i) the firm's exposure to the aggregate shock, β_i (and is increasing in β_i), and (ii) the time t price of risk, which is reflected in the term $\gamma_t \sigma_\varepsilon^2$. Intuitively, firm-level mpk deviations are composed of both a transitory component due to uncertainty and a persistent component due to the risk premium. The transitory components are i.i.d. over time and lead to purely temporary deviations in mpk (even though the underlying productivity processes are autocorrelated); the risk premium, on the other hand, leads to persistent deviations – firms that are more exposed to aggregate shocks, and so are riskier, will have persistently high mpk.

Expression (15) hones in on this second force and shows the persistent effects of risk premia on the conditional expectation of time $t+1$ mpk, denoted $Empk$. Indeed, in this simple case, the ranking of firms' mpk will be constant in expectation as determined by the risk premium – high beta firms will have permanently high $Empk$ and low beta firms the opposite. Importantly,

the *value* of *Empk* will fluctuate with γ_t, but the ordering across firms will be preserved. This is the sense that we call this component persistent/permanent. Expression (14) shows that this ordering will not be preserved in realized *mpk* – due to the realization of shocks, the ranking of firms' *mpk* will fluctuate, but the firm-specific risk premium adds a persistent component.[25] Because the uncertainty portion of the realized *mpk* is always additively separable and is independent of our mechanism, from here on we primarily work with *Empk*.

Expression (16) presents the cross-sectional variance of *Empk*:

$$\sigma^2_{Empk_t} \equiv \sigma^2_{\mathbb{E}_t[mpk_{it+1}]} = \sigma^2_\beta \left(\gamma_t \sigma^2_\varepsilon \right)^2 . \tag{16}$$

Cross-sectional variation in *Empk* depends on the dispersion in beta and the price of risk. Dispersion will be greater when risk prices, reflected by $\gamma_t \sigma^2_\varepsilon$, are high and so will be countercyclical. The average long-run level of *Empk* dispersion is given by

$$E\sigma^2_{Empk} \equiv \mathbb{E}\left[\sigma^2_{Empk_t} \right] = \sigma^2_\beta \left(\gamma_0^2 + \gamma_1^2 \sigma_x^2 \right) \left(\sigma^2_\varepsilon \right)^2 \quad \text{where} \quad \sigma_x^2 = \frac{\sigma^2_\varepsilon}{1 - \rho_x^2} . \tag{17}$$

Aggregate outcomes. Appendix D.3 shows that aggregate output can be expressed as

$$\log Y_{t+1} \equiv y_{t+1} = a_{t+1} + \theta_1 k_{t+1} + \theta_2 n_{t+1} ,$$

where k_{t+1} denotes the aggregate capital stock, n_{t+1} aggregate labor and a_{t+1} the level of aggregate TFP, given by

$$a_{t+1} = a^*_{t+1} - \frac{1}{2} \frac{\theta_1 (1 - \theta_2)}{1 - \theta_1 - \theta_2} \sigma^2_{mpk,t+1} , \tag{18}$$

where $\sigma^2_{mpk,t+1}$ is realized *mpk* dispersion in period $t + 1$. The term a^*_{t+1} is the first-best level of TFP in the absence of any frictions (i.e., where marginal products are equalized). Thus, aggregate TFP monotonically decreases in the extent of capital "misallocation," captured by σ^2_{mpk}. The effect of misallocation on aggregate TFP depends on the overall curvature in the production function, $\theta_1 + \theta_2$ and the relative shares of capital and labor. The higher is $\theta_1 + \theta_2$, that is, the closer to constant returns to scale, the more severe the losses from mis-allocated resources. Similarly, fixing the degree of overall returns to scale, for a larger capital share, θ_1, a given degree of misallocation has larger effects on aggregate outcomes.

Using equation (16), the conditional expectation of one-period ahead TFP is given by

$$\mathbb{E}_t[a_{t+1}] = \mathbb{E}_t[a^*_{t+1}] - \frac{1}{2} \frac{\theta_1 (1 - \theta_2)}{1 - \theta_1 + \theta_2} \sigma^2_\beta \left(\gamma_t \sigma^2_\varepsilon \right)^2 . \tag{19}$$

---

[25]With additional adjustment frictions, there will be other factors confounding the relationship between beta and the realized and expected *mpk*.

The expression shows that risk premium effects unambiguously reduce aggregate TFP and disproportionately more so in business cycle downturns, since γ_t is countercyclical. Taking unconditional expectations gives the effects on the average long-run level of TFP in the economy:

$$\overline{a} \equiv \mathbb{E}\left[\mathbb{E}_t\left[a_{t+1}\right]\right] = a^* - \frac{1}{2}\frac{\theta_1\left(1-\theta_2\right)}{1-\theta_1+\theta_2}\sigma_\beta^2\left(\gamma_0^2+\gamma_1^2\sigma_x^2\right)\left(\sigma_\varepsilon^2\right)^2 \ . \tag{20}$$

The expression directly links the extent of cross-sectional dispersion in required rates of return (which are in turn determined by the prices of risk and volatility of aggregate shocks) to the long-run level of aggregate productivity and gives a natural way to quantify the implications of these effects. Further, and perhaps more importantly, it uncovers a new connection between aggregate volatility and long-run economic outcomes, i.e., a "productivity cost" of business cycles – *ceteris paribus*, the higher is aggregate volatility (σ_ε^2 and σ_x^2 in the expression), the more depressed will be the average long-run level of TFP (relative to an environment with no aggregate shocks and/or risk premia).

In Appendix F, we show that our model can be extended to include multiple sources of aggregate risk and to allow γ_t to depend on additional factors beyond the state of technology and so expressions (19) and (20) provide a more general connection between financial conditions (that may be less than perfectly correlated with the real economy), the cross-sectional allocation of resources and aggregate TFP.[26] Thus, more broadly, these expressions provide one way to link the rich findings of the literature on cross-sectional asset pricing to real allocations and macroeconomic outcomes.

3.1 The Cross-Section of Expected Stock Returns and MPK

In this section, we derive a sharp link between a firm's beta – and so expected *mpk* – and its expected stock market return, along the lines developed in Section 2.2. This connection suggests an empirical strategy to measure the dispersion in beta and so quantify the *mpk* dispersion that arises from risk considerations using stock market data. Our key finding is that, to a first-order approximation, the firm's expected stock return is a linear (and increasing) function of its beta.[27] The implication is that, in the simple model outlined thus far, expected *mpk* is proportional to expected stock returns, and thus, the dispersion in expected stock returns puts tight empirical discipline on the dispersion in betas and so expected *mpk* arising from risk channels. We use this connection to provide transparent intuition for our numerical approach in Section 4.

---

[26]Further, we can verify that the additional extensions discussed in Section 2.1, i.e., firm-specific "alphas" and heterogeneous exposures to investment goods prices have similar implications.

[27]It is well known that a first-order approximation may not be sufficient to capture risk premia. In our quantitative work in Section 4, we work with numerical higher order approximations.

We obtain an analytic approximation for expected stock market returns by log-linearizing around the non-stochastic steady state where $X_t = Z_t = 1$. To a first-order, the (log of the) expected excess stock return is equal to (derivations in Appendix D.4)

$$Er^e_{it+1} \equiv \log \mathbb{E}_t \left[R^e_{it+1} \right] = \psi \beta_i \gamma_t \sigma^2_\varepsilon , \tag{21}$$

where

$$\psi = \frac{\frac{1}{\rho} + \delta - 1}{\frac{1}{\rho} + \delta (1 - \theta) - 1} \frac{1 - \rho}{1 - \rho \rho_x + \rho \gamma_1 \sigma^2_\varepsilon} .$$

The expected excess return depends on the firm's beta (indeed, is linear and increasing in beta) and is increasing in the price of risk. Because the price of risk is countercyclical, risk premia increase during downturns for all firms and fall during expansions.[28] The time t cross-sectional dispersion in expected excess returns is given by

$$\sigma^2_{Er^e_t} \equiv \sigma^2_{\log \mathbb{E}_t \left[R^e_{it+1} \right]} = \psi^2 \sigma^2_\beta \left(\gamma_t \sigma^2_\varepsilon \right)^2 . \tag{22}$$

Similar to our findings for expected mpk, the expression reveals a tight link between beta dispersion and expected stock return dispersion. Indeed, if firms had identical betas, dispersion in expected returns would be zero. Moreover, as with expected mpk dispersion, expected stock return dispersion is increasing in the price of risk and so is countercyclical.

Comparing equations (15) and (21) shows that expected excess returns are proportional to expected mpk and equations (16) and (22) show that $\sigma^2_{Er^e_t}$ is proportional to $\sigma^2_{Empk_t}$. Thus, the expressions reveal a tight connection between cross-sectional dispersion in expected stock returns and expected mpk – both are dependent on the variation in betas. Although the exact proportionality will not hold exactly in the full non-linear solution – or in the presence of other frictions/distortions to firm investment decisions – we will use this intuition to quantify the role of risk considerations in generating dispersion in expected mpk.

Specifically, these results suggest an empirical strategy to estimate the three key structural parameters – γ_0, γ_1 and σ^2_β – using readily available stock market data. First, it is straightforward to verify that the market index – i.e., a perfectly diversified portfolio with no idiosyncratic

---

[28]Strictly speaking, these results hold in the approximation so long as $1 - \rho \rho_x + \rho \gamma_1 \sigma^2_\varepsilon > 0$. This condition does not play a role in the numerical solution.

risk – achieves the maximal Sharpe ratio:[29]

$$SR_{mt} = \gamma_t \sigma_\varepsilon, \qquad ESR_m \equiv \mathbb{E}\left[SR_{mt}\right] = \gamma_0 \sigma_\varepsilon . \tag{23}$$

The expression links the market Sharpe ratio to γ_0. Indeed, in this linearized environment, the mapping is one-to-one (given σ_ε^2). Next, deriving equation (21) for the market index gives

$$Er_{mt+1} = \psi \bar{\beta} \gamma_t \sigma_\varepsilon^2, \qquad Er_m \equiv \mathbb{E}\left[Er_{mt+1}\right] = \psi \bar{\beta} \gamma_0 \sigma_\varepsilon^2 . \tag{24}$$

For a given value of γ_0, the equity premium is increasing as γ_1 becomes more negative through its effects on ψ ($\bar{\beta}$ denotes the mean beta across firms). Lastly, equation (22) connects dispersion in beta, σ_β^2, to dispersion in expected returns. Together, equations (22), (23) and (24) tightly link three observable moments of asset market data to the three parameters, γ_0, γ_1 and σ_β^2.

3.2 Adjustment Costs

In this section, we extend the framework to include capital adjustment costs. Although the main insights from the previous sections go through, we illustrate an important interaction between these costs and the effects of risk premia, namely, adjustment costs amplify the impact of these systematic risk exposures on mpk dispersion.

We assume that capital investment is subject to quadratic adjustment costs, given by

$$\Phi\left(I_{it}, K_{it}\right) = \frac{\xi}{2}\left(\frac{I_{it}}{K_{it}} - \delta\right)^2 K_{it} .$$

With these costs, the return on capital is no longer equal to the MPK plus the undepreciated capital stock, but depends on endogenous fluctuations in the value of installed capital, i.e., Tobin's Q. Although exact analytic solutions are no longer available as in the simpler case, a first-order approximation yields the return on capital (the investment return) to be[30]

$$r_{it+1}^I = \left(1 - \rho\left(1 - \delta\right)\right) mpk_{it+1} + \rho q_{it+1} - q_{it} . \tag{25}$$

The investment return depends on mpk_{it+1}, but additionally on q_{it} and q_{it+1}, where $q_{it} \equiv$

---

[29]The Sharpe ratio for an individual firm is $SR_{it} = \dfrac{\beta_i \gamma_t \sigma_\varepsilon^2}{\sqrt{\left(\frac{1 - \rho \rho_x + \rho \gamma_1 \sigma_\varepsilon^2}{1 - \rho \rho_z}\right)^2 \sigma_{\bar{\varepsilon}}^2 + \beta_i^2 \sigma_\varepsilon^2}}$, which shows that, due to the

presence of idiosyncratic risk, individual firms do not attain the maximum Sharpe ratio. However, in this linear environment, the diversified index faces no risk from $\sigma_{\bar{\varepsilon}}^2$, so that the expression collapses to (23). Although in the full numerical solution the market may not exactly attain this value due to the nonlinear effects of idiosyncratic shocks, the expression highlights that the market Sharpe ratio is informative about γ_0.

[30]Throughout this section, we suppress constant terms that play no role.

$\xi\left(k_{it+1}-k_{it}\right)$. As above, the risk premium on capital is equal to (the negative of) the covariance of the return with the SDF, i.e.,

$$\log \mathbb{E}_t\left[r_{it+1}^I\right] = -\operatorname{cov}_t\left((1-\rho(1-\delta)) mpk_{it+1} + \rho q_{it+1}, m_{t+1}\right) , \qquad (26)$$

which shows that adjustment costs add an additional element to the risk premium through endogenous changes in q_{it} that are correlated with the SDF.

Appendix D.2 derives the log-linearized version of the firm's optimal investment policy:[31]

$$k_{it+1} = \phi_1\beta_i x_t + \phi_2 z_{it} + \phi_3 k_{it} - \phi_4\beta_i\gamma_0\sigma_\varepsilon^2 , \qquad (27)$$

where

$$0 = \left((\theta-1)-\hat{\xi}(1+\rho)\right)\phi_3 + \rho\hat{\xi}\phi_3^2 + \hat{\xi}$$

$$\phi_1 = \frac{\left(\rho_x-\gamma_1\sigma_\varepsilon^2\right)\phi_3}{\hat{\xi}\left(1-\rho\phi_3\left(\rho_x-\gamma_1\sigma_\varepsilon^2\right)\right)}, \qquad \phi_2 = \frac{\rho_z\phi_3}{\hat{\xi}\left(1-\rho\rho_z\phi_3\right)}$$

$$\phi_4 = \frac{\phi_3}{\hat{\xi}\left(1-\rho\phi_3\right)}\frac{1}{1-\rho\phi_3\left(\rho_x-\gamma_1\sigma_\varepsilon^2\right)} ,$$

and $\hat{\xi} \equiv \frac{\xi}{1-\rho(1-\delta)}$ is a composite parameter that captures the severity of adjustment costs.

Now, the past level of capital affects the new chosen level. The coefficient ϕ_3 captures the strength of this relationship. It lies between zero and one and is increasing in the adjustment cost, $\hat{\xi}$. It is independent of the risk premium. The other coefficients each have a counterpart in equation (13), but are modified to reflect the influence of adjustment costs. The coefficients ϕ_1 and ϕ_2 are both decreasing in these costs – intuitively, adjustment costs reduce the firm's responsiveness to transitory shocks. Importantly, ϕ_4 is increasing in these costs, showing that they increase the importance of the firm's beta in determining its choice of capital.[32] The expression for ϕ_4 also reveals an interaction between adjustment costs and time-varying risk – the denominator contains the product of ϕ_3 and γ_1, which implies that a more negative γ_1 leads to higher values of ϕ_4 as long as adjustment costs are non-zero. By increasing the value of ϕ_4, this interaction effect strengthens the impact of beta dispersion on *Empk* dispersion.

From here, we can derive the following expression for conditional expected *mpk*:

$$\mathbb{E}_t\left[mpk_{it+1}\right] = \frac{1}{1-\rho\phi_3\left(\rho_x-\gamma_1\sigma_\varepsilon^2\right)}\beta_i\gamma_t\sigma_\varepsilon^2 + \rho\hat{\xi}\mathbb{E}_t\left[k_{it+2}-k_{it+1}\right] - \hat{\xi}\left(k_{it+1}-k_{it}\right) . \qquad (28)$$

---

[31]As above, we ignore terms reflecting variance adjustments that are close to zero.

[32]Strictly speaking, this is true so long as $1-\rho\phi_3\left(\rho_x-\gamma_1\sigma_\varepsilon^2\right) > 0$. This condition holds for any reasonable level of adjustment costs, for example, given our estimates of the other parameters, ξ must be less than approximately 2180.

Expected *mpk* depends on both the risk premium and adjustment costs (realized *mpk* also depends on the realization of shocks, as above). The last two terms capture the effects of adjustment costs alone and, conditional on current and expected future capital stocks, are independent of aggregate risk. The first term captures the risk premium. Without adjustment costs, $\phi_3 = \hat{\xi} = 0$, and the risk premium is identical to expression (13)). The risk premium is increasing in those costs (i.e., as ϕ_3 gets larger), showing that adjustment costs amplify risk premium effects. Intuitively, as shown in (26), adjustment costs add an additional source of co-movement of capital returns with the SDF through fluctuations in firm-level Q. High beta firms invest more in expansions (both because their expected productivity is high and the cost of capital is low due to a low price of risk) and so the Q of these firms is negatively correlated with the SDF (e.g., is more procyclical), inducing a higher risk premium.[33]

The extended model continues to give rise to *mpk* deviations that are persistent at the firm-level. In particular, taking the unconditional expectation of (28) yields expressions for the persistent component of *Empk* and its cross-sectional dispersion:

$$\mathbb{E}\left[Empk_{it+1}\right] = \frac{1}{1 - \rho\phi_3\left(\rho_x - \gamma_1\sigma_\varepsilon^2\right)}\beta_i\gamma_0\sigma_\varepsilon^2 \tag{29}$$

$$\Rightarrow$$

$$\sigma_{\mathbb{E}[Empk_{it+1}]}^2 = \left(\frac{1}{1 - \rho\phi_3\left(\rho_x - \gamma_1\sigma_\varepsilon^2\right)}\right)^2 \left(\gamma_0\sigma_\varepsilon^2\right)^2 \sigma_\beta^2. \tag{30}$$

The risk premium is an essential ingredient for the model to generate persistence in firm-level *mpk*; adjustment costs are not sufficient. When $\gamma_0 = 0$ and hence risk effects are absent, there is no persistent *Empk* dispersion, even with adjustment costs (beta dispersion is also necessary). Thus, on their own, adjustment costs generate dispersion only in the transitory component of *mpk* (through their role in expression (28)). However, (29) and (30) show that when the persistent risk premium component is present, adjustment costs have a further amplification effect on that component, equal to the fraction in those expressions. How large might this amplification effect be? Using the parameter values from the next section, which include a relatively modest level of adjustment costs, the scaling factor from these costs is about 1.75, which implies a cross-sectional variance in *Empk* that is scaled up by a factor of three relative to the case with no adjustment costs. Thus, although they do not change the qualitative predictions of the model, adjustment costs can have an important quantitative effect on the results. In contrast, the transitory effects coming from these costs will turn out to be small.

Finally, how do adjustment costs change the relationship between expected *mpk*, beta and

---

[33]For a related, but slightly different intuition, adjustment costs cause capital to be a long-lived asset and thus increase the length of the relevant time horizon when considering a capital investment. Because the amount of risk is increasing in the length of the horizon, the risk premium is naturally larger.

expected stock returns? Appendix D.4 shows that to a first-order, expected returns are not affected by adjustment costs and so the results from Section 3.1 continue to hold.[34] Thus, the arguments made in that section linking the key parameters of the model to moments of asset returns go through unchanged.

4 Quantitative Analysis

In this section, we use the analytical insights laid out above to numerically quantify the extent of mpk dispersion arising from risk premia effects.

4.1 Parameterization

We begin by assigning values to the more standard production parameters of our model. Following Atkeson and Kehoe (2005), we set the overall returns to scale in production $\theta_1 + \theta_2$ to 0.85. We assume standard shares for capital and labor of 0.33 and 0.67, respectively, which gives $\theta_1 = 0.28$ and $\theta_2 = 0.57$. These values imply $\theta = 0.65$.[35] We assume a period length of one year and accordingly set the rate of depreciation to $\delta = 0.08$. We estimate the adjustment cost parameter, ξ, in order to match the autocorrelation of investment, denoted $\text{corr}(\Delta k_t, \Delta k_{t-1})$, which is 0.38 in our data. Equation (39) in Appendix D.5 provides a closed-form expression for this moment, which reveals a tight connection with the severity of adjustment frictions.[36]

To estimate the parameters governing the aggregate shock process, we build a long sample of Solow residuals for the US economy using data from the Bureau of Economic Analysis on real GDP and aggregate labor and capital. The construction of this series is standard (details in Appendix B.4). With these data, we use a standard autoregression to estimate the parameters ρ_x and σ_ε^2. This procedure gives values of 0.94 and 0.0247 for the two parameters, respectively.[37]

---

[34]Although this is only exactly true under our first-order approximation, Table 4 verifies numerically that at their estimated level, adjustment costs have relatively modest effects on moments of returns.

[35]This is close to the values generally used in the literature. For example, Cooper and Haltiwanger (2006) estimate a value of 0.59 for US manufacturing firms. David and Venkateswaran (2019) use a value of 0.62.

[36]The expression also reveals that for ρ_x close to ρ_z, which we find in the data, described next, the autocorrelation of within-firm investment is almost invariant to the firm's beta (indeed, the invariance is exact if $\rho_x = \rho_z$). Thus, even with dispersion in betas, we may not see large variation in this moment across firms.

[37]The autoregression does not reject the presence of a unit root at standard confidence levels. We have also worked with the annual TFP series developed by John Fernald, available at:
https://www.frbsf.org/economic-research/indicators-data/total-factor-productivity-tfp/.
These data are only available for the more recent post-war period, but also show that the series is close to a random walk (i.e., the autocorrelation of growth rates is essentially zero). A potential concern with this approach is that these series reflect not only the process on exogenous technology, but also the effects of mpk dispersion itself (since dispersion affects measured aggregate productivity). However, at our estimates, these effects are small – mpk dispersion primarily impacts the level of aggregate productivity (which does not affect our estimates of persistence or volatility) but has only a small impact on its time-series properties (we

Under our assumptions, firm-level productivity (including the aggregate component) can be measured directly (up to an additive constant) as $y_{it} - \theta k_{it}$. After controlling for the level of aggregate productivity, a similar autoregression on the residual (firm-specific) component yields values for ρ_z and $\sigma_{\tilde{\varepsilon}}$ of 0.93 and 0.28, respectively.

Turning to the parameters of the SDF, we set $\rho = 0.988$ to match an average annual risk-free rate of 1.2%. Following the arguments in Section 3.1, we estimate the values of γ_0 and γ_1 to match the post-war (1947-2017) average annual excess return on the market index of 7.7% and Sharpe ratio of 0.53.[38] This strategy is equivalent to matching both the mean and volatility of market excess returns (the standard deviation is 14.6%). To be comparable to the data, stock returns in the model need to be adjusted for financial leverage. To do so, we scale the mean and standard deviation of the model-implied returns by a factor of $1 + \frac{D}{E}$ where $\frac{D}{E}$ is the debt-to-equity ratio. We follow, e.g., Barro (2006) and assume an average debt-to-equity ratio of 0.5. Because both the numerator and denominator are scaled by the same constant, the Sharpe ratio is unaffected. For ease of interpretation, in what follows, we report the properties of levered returns. To compute the model-implied market return, we must also take a stand on the mean beta across firms. Assuming that the mean of $\hat{\beta}_i$ (the underlying productivity beta) is one, and using the value of ω (the sensitivity of wages to aggregate shocks) suggested by İmrohoroğlu and Tüzel (2014) of 0.20, we can compute the mean beta to be 1.99.[39] This is simply the mean productivity beta adjusted for the leverage effects of labor liabilities. This procedure yields values of $\gamma_0 = 32$ and $\gamma_1 = -140$.

Finally, again following the insights in Section 3.1, we estimate the dispersion in betas to match the cross-sectional dispersion in expected stock returns. To be consistent with the broad literature, we use the expected returns predicted from the Fama-French model as computed in Section 2.2. We de-lever firm-level expected returns following the approach in Bharath and Shumway (2008) and Gilchrist and Zakrajsek (2012) (details in Appendix B.2). This procedure yields an estimated average within-industry standard deviation of un-levered expected returns of 0.127 (we report details and plot the full histogram of the expected return distribution in Appendix B.2: for example, the mean is about 9%, and the interquartile range is just under 12%; the standard deviation of raw expected returns, i.e., not de-levered or controlling for industry, is about 0.156).[40] Feeding this value into our quantitative model yields an estimate

Table 3: Parameterization - Summary

Parameter	Description	Value
Production		
θ_1	Capital share	0.28
θ_2	Labor share	0.57
δ	Depreciation rate	0.08
ξ	Adjustment cost	0.04
$\sigma_{\hat{\beta}}$	Std. dev. of risk exposures	4.80
Stochastic Processes		
ρ_x	Persistence of agg. shock	0.94
σ_ε	Std. dev. of agg. shock	0.0247
ρ_z	Persistence of idiosyncratic shock	0.93
$\sigma_{\tilde{\varepsilon}}$	Std. dev. of idiosyncratic shock	0.28
ω	Wage elasticity	0.20
Stochastic Discount Factor		
ρ	Time discount rate	0.988
γ_0	SDF – constant component	32
γ_1	SDF – time-varying component	-140

for σ_β of 12, and adjusting for the scaling $1 - \theta_2$ gives the dispersion in underlying productivity betas, $\sigma_{\hat{\beta}}$, equal to 4.80.[41]

We parameterize the model using simulated method of moments (details in Appendix E). Table 3 summarizes our empirical approach/results.

4.2   Risk-Based Dispersion in MPK

Table 4 presents our main quantitative results. We report four variants of the framework. Column (1) ("Baseline") corresponds to the full model with time-varying risk and adjustment costs. Column (2) ("No Risk") shows results from a version where risk effects are completely absent (specifically, we set $\gamma_0 = \gamma_1 = 0$). Column (3) ("Only Risk") reports the effects of risk premia alone, without adjustment costs (i.e., $\xi = 0$). Column (4) ("Constant Risk") examines a version with adjustment costs but a constant price of risk (i.e., $\gamma_1 = 0$). Column (5) ("Only

---

distribution from a number of predictive models. For example, using monthly data, he finds an annualized cross-sectional standard deviation of up to 17.5% (Model 3, Panel A, Table 5 of that paper).

[41]Although this a significant amount of dispersion, it composes only a modest fraction of overall dispersion in firm-level productivity. To see this, note that the cross-sectional variance of productivity at time t is $\left(\frac{\sigma_{\hat{\beta}}}{1-\theta_2}\right)^2 x_t^2 + \sigma_z^2$, where $\sigma_z^2 = \frac{\sigma_{\tilde{\varepsilon}}^2}{1-\rho_z^2}$. Plugging in our estimates and assuming, for example, that the economy is 2% above or below trend, gives the first term to be about 8% of the total. It remains relatively modest for reasonable deviations from trend. Thus, despite firms' diverse sensitivities to business cycle shocks, our estimates still point to firm-level idiosyncratic conditions as the dominate factor driving cross-sectional heterogeneity.

Table 4: Risk Premia and Misallocation

	Baseline	No Risk	Only Risk	Constant Risk	Only Constant Risk
	(1)	(2)	(3)	(4)	(5)
MPK Implications					
$E\sigma^2_{Empk}$	0.17	0.03	0.05	0.16	0.05
% of total σ^2_{mpk}	37.9%	7.4%	11.5%	35.9%	10.4%
$\sigma^2_{\overline{Empk}}$	0.14	0.00	0.05	0.13	0.05
% of total $\sigma^2_{\overline{mpk}}$	47.3%	0.1%	15.7%	41.9%	15.7%
$\Delta\overline{a}$	0.07	0.01	0.02	0.07	0.02
corr $\left(\sigma^2_{Empk_t}, x_t\right)$	-0.31	0.01	-0.97	0.45	0.00
Moments					
Er^e_m	0.08	0.00	0.10	0.05	0.06
ESR_m	0.53	0.00	0.61	0.52	0.65
corr $(\Delta k_t, \Delta k_{t-1})$	0.38	0.38	-0.02	0.38	-0.03

Constant Risk") has a constant price of risk and no adjustment costs. Our goal in showing these different permutations is to understand the role that each element of the model plays in leading to various patterns in mpk dispersion. In order to interpret the results as decomposition, each variant holds all the parameters fixed at their estimated values except the one under study.

Long-run effects. The first row of the table shows the average level of mpk dispersion that arises in each variant of the model.[42] The second row shows the percentage of total observed misallocation that this value accounts for. In our sample, overall σ^2_{mpk} is 0.45. This is the denominator in that row. Next, we calculate the dispersion stemming from only the permanent component of firm-level MPK deviations (given by equation (28)), which we report in the third row of the table. To compute this value in the data, for each firm, we regress the time-series of its mpk on a firm-level fixed effect. The fixed effect is the permanent component of firm-level mpk and the residuals transitory components. We then compute the variance of the permanent component, which yields a value of $\sigma^2_{\overline{mpk}} = 0.30$, about two-thirds of the total.[43] This is the denominator in the fourth row of the table, which displays the model-implied permanent dispersion as a percentage of the observed permanent component in the data. The next row quantifies the implications of the estimated dispersion for the long-run level of aggregate TFP. It reports the gains in the average level of TFP from eliminating the predicted mpk dispersion,

---

[42]With adjustment costs, we do not have analytic expressions for period-by-period $Empk$ dispersion. We compute these values using simulation and then average over them. Without adjustment costs, we can use expression (17) directly.

[43]Other approaches give a similar breakdown, see, e.g., David and Venkateswaran (2019).

denoted $\Delta\bar{a}$.[44] This is essentially an application of expression (18).

Column (1) shows that the full model generates mpk dispersion of about 0.17. This accounts for about 38% of overall mpk dispersion in the data.[45] Of the model-implied dispersion, about 0.14 is permanent in nature, which explains about 47% of the permanent component in the data. The costs of this dispersion represent a loss in long-run TFP of about 7%. In the full model, both risk effects and adjustment costs lead to dispersion in $Empk$. To hone in on the role of risk alone, column (2) shows the same statistics when we eliminate risk effects and only adjustment costs are present. Adjustment costs on their own generate relatively modest dispersion in $Empk$ ($E\sigma^2_{Empk} = 0.03$) and as proved in Section 3.2, do not lead to any dispersion in the permanent component, ($\sigma^2_{\overline{Empk}} = 0$). Thus, risk premia effects are crucial to generating the substantial and persistent dispersion in column (1). Subtracting column (2) from column (1) captures the contribution of risk effects alone: $Empk$ dispersion of 0.14 (about 30% of the total in the data), dispersion in the permanent component of 0.14 (47% of the data) and long-run TFP losses of 6%. These results suggest (i) heterogeneity in risk premia can generate significant MPK dispersion, particularly when compared to the permanent component in the data, and (ii) the consequences for measures of aggregate performance such as TFP – i.e., the "productivity costs" of business cycles – can be substantial.

Column (3) removes adjustment costs to illustrate their amplification of existing risk premia effects. On their own (i.e., without adjustment costs), risk premia generate mpk dispersion of 0.05, which accounts for 11.5% of total σ^2_{mpk} in the data and explain about 16% of the permanent component. Thus, although the impact of risk premia remains significant in isolation, they are less than half of those in column (1), where the amplification from adjustments costs is taken into account. TFP losses are also smaller, but remain significant, at approximately 2%.

Columns (4) and (5) show that the majority of these effects stems from the presence of a high persistent component in the price of risk, i.e., γ_0, rather than from the time-variation from γ_1. Setting $\gamma_1 = 0$ only modestly reduces the size of these effects in the presence of adjustment costs (compare columns (1) and (4)) and has a negligible effect on the results without them (columns (3) vs. (5)). The implication is that time-variation in the price of risk does not add much to the long-run level of mpk dispersion.

Countercyclical dispersion. The last row in the top panel examines the second main implication of the theory, namely, the countercyclicality of mpk dispersion, which we measure as the correlation of $\sigma^2_{Empk_t}$ with the state of the business cycle, x_t. Column (1) shows that the full

---

[44]This calculation does not mean that policies eliminating this source of mpk dispersion would necessarily be desirable. We merely see this as a useful way to quantify the implications of our findings.

[45]The model-implied ratio of the standard deviation of MPK to that in expected stock returns is about 3.25 ($\frac{.17^{\frac{1}{2}}}{.127}$), which is close to the empirical estimates in Section 2.2.

model generates significantly countercyclical dispersion in $Empk$ – the correlation of $\sigma^2_{Empk_t}$ with the state of the cycle is -0.31. To put this figure in context, Table 6 in Appendix B.3 shows that the correlation between σ^2_{mpk} and the cyclical component of aggregate productivity in the data is -0.27. Thus, the model predicts countercyclical dispersion on par with this value. Column (2) shows that adjustment costs alone do not generate any cyclicality in $Empk$ dispersion. Column (3) shows that as the only factor behind $Empk$ dispersion, the time-varying risk premium would lead to an almost perfectly negative correlation with the business cycle. This is a clear implication of equation (16). The additional presence of adjustment costs in the first column confounds this relationship and leads to a smaller correlation (in absolute value) that is more in line with the data. Finally, the last two columns illustrate that time-varying risk is key to generating countercyclical dispersion. Without this element, $Empk$ dispersion is significantly positive with adjustment costs and without them, is exactly acyclical. Thus, our findings suggest that the interaction of a countercyclical price of risk with adjustment frictions is crucial in yielding a negative (though far from negative one) correlation between $Empk$ dispersion and the state of the business cycle.

To highlight the potential implications of the countercyclical $Empk$ dispersion produced by our model, consider the connection with the empirical findings in Eisfeldt and Rampini (2006), who show that firm-level dispersion measures tend to be countercyclical, yet most capital reallocation is procyclical. Our theory can – at least in part – reconcile this observation due to the countercyclical nature of factor risk prices and the high beta of high MPK firms: countercyclical reallocation would entail moving capital to the riskiest of firms in the riskiest of times. Thus, in light of our results, it may not be as surprising that countercyclical dispersion obtains, even in a completely frictionless environment.[46]

Moments. In the bottom panel of Table 4, we investigate the role of each element in matching the target moments. Our full model in column (1) is directly parameterized to match the three moments, i.e., the equity premium, Sharpe ratio and autocorrelation of investment. Column (2) shows that without risk aversion, risk premia are essentially zero. Column (3) shows that, as implied by the approximation in Section 3.2, adjustment costs have a modest effect on the properties of returns (eliminating them somewhat raises the equity premium and Sharpe ratio). However, the autocorrelation of investment falls dramatically without them, indeed, becoming slightly negative (due to the mean-reverting nature of shocks). Thus, some degree of adjustment costs is crucial for matching this moment. Comparing columns (1) and (4) shows that without

---

[46]The main measure of reallocation in Eisfeldt and Rampini (2006) includes both mergers and acquisitions (M&A) as well as sales of disassembled capital (sales of property, plant and equipment). Even excluding M&A, they find the latter is significantly procyclical (correlation with GDP of about 0.4; data from `https://sites.google.com/site/andrealeisfeldt/home/capital-reallocation-and-liquidity`).

time-varying risk, the model struggles to match the equity premium, which falls almost by half, from about 8% to 5%. As implied by expressions (24), (23) and (39), time-varying risk is tightly linked to average excess returns, but has only modest effects on the average Sharpe ratio and the autocorrelation of investment. A similar pattern emerges from columns (3) and (5) – in the absence of adjustment costs, removing time-varying risk significantly reduces the equity premium but has smaller effects on the other two moments.

In sum, the results in Table 4 show first, heterogeneity in firm-level risk premia leads to quantitatively important dispersion in mpk, with significant adverse effects on aggregate TFP; moreover, much of this dispersion is persistent and can account for a significant portion of what seems to be a puzzling pattern in the data, namely, persistent mpk deviations at the firm-level. Second, these risk premium effects add a notably countercyclical element to mpk dispersion, going some way towards reconciling the countercyclical nature of firm-level dispersion measures.

4.3   Other Distortions

Recent work has pointed to a number of additional factors (beyond fundamentals and adjustment frictions) that may affect firms' investment decisions and lead to mpk dispersion, for example, financial frictions or policy-induced distortions. Moreover, it has been pointed out that attempts to identify one of these forces – while abstracting from others – may yield misleading conclusions. This section demonstrates that our strategy of using asset market data is robust to this critique. In other words, our approach yields accurate estimates of risk premium effects, even in the presence of other, un-modeled, distortions.

Rather than take a stand on the exact nature of these factors, we follow the broad literature, e.g., Hsieh and Klenow (2009) and Restuccia and Rogerson (2008), and model these distortions using a flexible class of "taxes" or "wedges," which can have a rich correlation structure over time and with both firm-level characteristics and aggregate conditions (in Section 5 we analyze two additional sources of measured mpk dispersion, namely, heterogeneity in markups and production function parameters). Specifically, we introduce a proportional "tax" on firm-level operating profits, $1 - e^{\tau_{it+1}}$ (so that the firm keeps a portion $e^{\tau_{it+1}}$), of the form:

$$\tau_{it+1} = -\nu_1 z_{it+1} - \nu_2 x_{t+1} - \nu_3 \beta_i x_{t+1} - \eta_{it+1} . \tag{31}$$

The first term captures a component correlated with the firm's idiosyncratic productivity, where the strength of the relationship is governed by ν_1. If $\nu_1 > 0$, the wedge discourages (encourages) investment by firms with high (low) idiosyncratic productivity. If $\nu_1 < 0$, the opposite is true. The next two terms capture the correlation of the wedge with the state of the business cycle, x_t. We allow for a component through which all firms are equally distorted by the cyclical

31

portion of the wedge, captured by ν_2, and a component through which high beta firms are disproportionately affected by the cyclical portion, captured by ν_3. Through this piece, the wedge can be correlated with firm-level betas. The last term, η_{it+1}, captures factors that are uncorrelated with firm or aggregate conditions. It can be either time-varying or fixed and is normally distributed with mean zero and variance σ_η^2. Low (high) values of η spur (reduce) investment by firms irrespective of their underlying characteristics or the state of the business cycle.[47] David and Venkateswaran (2019) show that a related formulation describes observed MPK dispersion well (although they do not have beta dispersion or aggregate shocks). 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 allocation process.

To gain intuition, we analyze each component of the distortion in turn. First, we focus only on the first and last terms, i.e., we set $\nu_2 = \nu_3 = 0$. In this case, the wedge is purely idiosyncratic in the cross-section, i.e., it is always mean zero and has no aggregate component. This formulation is closest to the ones typically used in the literature, which has typically focused on idiosyncratic distortions with no aggregate shocks. Appendix G derives the following expressions for expected mpk and its cross-sectional variance:

$$Empk_{it+1} = \alpha + \nu_1\rho_z z_{it} + \eta_{it+1} + \beta_i\gamma_t\sigma_\varepsilon^2, \quad \Rightarrow \quad \sigma_{Empk_t}^2 = (\nu_1\rho_z)^2\sigma_z^2 + \sigma_\eta^2 + \sigma_\beta^2\left(\gamma_t\sigma_\varepsilon^2\right)^2 . \quad (32)$$

In this case, $Empk$ includes (i) a component that reflects the correlated distortion, ν_1, and depends on the firm's expectations of its idiosyncratic productivity ($\rho_z z_{it}$), leading to mpk deviations that are correlated with idiosyncratic productivity, and (ii) a term that reflects the uncorrelated distortion, η, which leads to mpk deviations that are uncorrelated with productivity. The last term reflects the risk premium. All of these components lead to dispersion in $Empk$ (dispersion in realized mpk also reflects uncertainty over shocks).

Crucially, expression (32) reveals that the risk premium (and resulting risk-based dispersion) are unaffected by the presence of these additional distortions. Further, Appendix G proves that expected stock returns are also unaffected, i.e., equation (21) still holds. The result implies that the mapping from expected returns to beta is, to a first-order, unaffected by the distortions, as is the mapping from beta dispersion to its effects on $Empk$. This leads to an important finding: even in the richer environment here featuring a common class of mis-allocative distortions, using stock market data continues to yield accurate estimates of the effects of heterogeneous risk exposures alone. Clearly, a strategy using mpk dispersion directly does not share this feature: measuring risk effects alone would be complicated by the presence of other distortions.

---

[47]We have also studied a version where the distortion is size-dependent, i.e., $\tau_{it+1} = -\nu_k k_{it+1}$. This turns out to be equivalent to the specification in (31) where $\nu_1 = \nu_3 = \frac{\nu_k}{1-\theta+\nu_k}$ and $\nu_2 = 0$.

Next, we add the components that are correlated with aggregate conditions. First, consider the case with a common cyclical component, i.e., $\nu_2 \neq 0$. We can prove a similar result as with only idiosyncratic wedges – the distortion does not affect the cross-sectional dispersion in expected stock returns or the risk-related dispersion in $Empk$.

Finally, consider the case where high beta firms are disproportionately affected by the aggregate distortion, i.e., $\nu_3 \neq 0$. If $\nu_3 > 0$, the distortion discourages (encourages) investment by high (low) beta firms in good times and the reverse in bad times (in this sense, it works like a cyclical productivity-dependent component, since high beta firms are relatively more productive in good times). There is also an aggregate implication of the wedge: averaging across firms gives $\bar{\tau}_{t+1} = -\nu_3 \bar{\beta} x_{t+1}$. If $\nu_3 > 0$ (< 0), the tax is pro- (counter) cyclical. $Empk$ is given by:

$$Empk_{it+1} = \alpha + \nu_1 \rho_z z_{it} + \nu_3 \beta_i \rho_x x_t + (1 - \nu_3) \beta_i \gamma_t \sigma_\varepsilon^2 + \eta_{it+1} \ . \tag{33}$$

The second and third terms captures the effects of idiosyncratic and aggregate distortions, respectively, and are independent of risk. The second to last term captures the risk premium, which is now scaled by a factor $1 - \nu_3$. Further, we can show that expected stock returns are scaled by exactly the same factor. The key implication is that all results from the baseline version go through, with a reinterpretation of the beta we recover from stock market data: rather than picking up the true beta alone, stock market returns yield a measure of the distorted beta, $(1 - \nu_3) \beta_i$. Since this is the object that also determines the risk premia in MPK, the remainder of our results stay largely unaffected.[48]

4.4 Directly Measured Productivity Betas

Our baseline approach to measuring firm-level risk exposures used the link between beta and expected stock returns laid out in Section 3.1. Here, we use an alternative strategy to estimate the dispersion in these exposures using only production-side data. In one sense, this approach is more direct – there is no need to employ firm-level stock market data to measure risk exposures. On the other hand, computing betas 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 may be difficult to apply this method to firms in developing countries (where measured misallocation tends to be larger), since most firm-level datasets there have

---

[48]One caveat is that when taking the cross-sectional variance of (33), an additional term arises from the covariance of the risk premium term with the aggregate distortion term. If the wedge worsens in downturns, i.e., $\nu_3 < 0$, which may be a plausible conjecture, we can prove that our baseline calculations yield a lower bound on risk premium effects on mpk dispersion. If the wedge is procyclical, i.e., $\nu_3 > 0$, we could be at risk of overstating these effects. However, Appendix G derives an upper bound on this bias at the estimated parameters and shows that it is quantitatively small.

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.

For each firm, we regress measured productivity growth, i.e., $\Delta z_{it} + \beta_i \Delta x_t$, on aggregate productivity growth Δx_t. It is straightforward to verify that the coefficient from this regression is exactly equal to β_i. Using these estimates, we can compute the firm's underlying productivity beta, $\hat{\beta}_i$, and calculate the cross-sectional dispersion in these estimates, $\sigma_{\hat{\beta}}^2$. We have applied this procedure using three different measures of the aggregate shock: (i) our long sample of Solow residuals, (ii) the series we construct from firm-level data (both of these are described in Appendix B.4) and (iii) the Fernald annual TFP series. The results yield values of $\sigma_{\hat{\beta}}$ of 6.4, 4.3 and 5.9, respectively. Recall that our estimate for this value using stock return data was 4.8, which is in line with – and towards the lower end of – the range found here.

4.5 Measurement Concerns

In this section, we address a number of potential measurement-related issues. First, following the recent literature, e.g., Hsieh and Klenow (2009) and Gopinath et al. (2017), we measure firm-level capital stocks using reported book values. An alternative approach is to use the perpetual inventory method along with detailed data on investment flows and investment good price deflators to construct capital stocks. Although this is in general an important issue for the firm dynamics/misallocation literatures, our empirical approach allows us to largely avoid this concern. To see this, notice that our estimation relies on measures of firm-level capital in only two places: first, to calculate the properties of idiosyncratic shocks, i.e., ρ_z and $\sigma_{\hat{\varepsilon}}^2$, and second, to calculate the autocorrelation of investment, which largely identifies the extent of adjustment costs. As shown, for example, in equations (16) and (22), idiosyncratic shocks have no effect on dispersion in expected mpk or on expected stock returns (the latter to a first-order). In our framework, idiosyncratic risk, though crucial in explaining firm dynamics, is not priced, and thus does not affect risk premia.[49] David and Venkateswaran (2019) measure firm-level capital using both approaches and find a larger serial correlation of investment using the perpetual inventory method. Since our adjustment cost estimate is increasing in the serial correlation, this approach would likely lead to larger estimates, and, as shown above, further amplify the risk premium effects we uncover.[50] Largely avoiding the use of firm-level capital measures is an

---

[49]We have also verified that idiosyncratic shocks have little effect on our estimates in the full non-linear model (see Appendix I).

[50]If the serial correlation is lower, we can generally think of the results in column 2 of Table 4, where we set adjustment costs to zero, as a lower bound. Although not part of our estimation, we also use firm-level capital to calculate total mpk dispersion, e.g., the denominator in the second row of Table 4. David and Venkateswaran (2019) show that this statistic is very similar under the two measurement approaches (see Tables 2 and 18 in that paper).

important feature of our use of stock market data.[51]

How about the effects of measurement error? Our use of stock market data is also useful in this regard – in general, stock market data should be quite precisely measured and so largely free of this concern. Measurement error in capital may affect our estimate of adjustment costs, but we can show that this error would likely lead us to a conservative estimate for these costs. To see this, consider first the case of (classical) measurement error that is iid over time. This unambiguously reduces the observed serial correlation (i.e., the true one is higher), which would yield higher adjustment cost estimates. Alternatively, consider the opposite case where the measurement error is permanent. Then, since we work with the growth rate of capital, our results would be unaffected. Of course, similarly to mis-measured capital, measurement error may affect the observed amount of mpk dispersion.

These issues may also be concerns for our estimates of productivity betas in the previous section, where we used measures of capital to calculate firm-level productivity. However, in Appendix H we show that any potential bias is likely quite small. Loosely speaking, mis-measured capital introduces error into the dependent variable of the regression, which, under certain conditions, will not affect our estimates (specifically, so long as changes in the measurement error are uncorrelated with changes in aggregate productivity). In that appendix we also investigate the potential bias in those estimates coming from unobserved heterogeneity in parameters across firms, i.e., θ, and show that it is quite small.

5 The Sources of Betas

Cross-firm variation in exposure to aggregate shocks, i.e., beta, is an essential ingredient in our theory. In this section, we investigate some potential sources of this type of heterogeneity – namely, dispersion in technological parameters (input elasticities in production) and markups as well as in the sensitivity of demand to business cycle fluctuations. Importantly, we show that each of these forms of heterogeneity is reflected in our measured betas, so that our main results on risk premia and mpk dispersion go through unchanged. Our goal here is simply to gain some further insight into why firms exhibit different sensitivities to aggregate shocks.

Heterogeneous technologies/markups. Firm-level heterogeneity in production function parameters or markups are potential sources of beta dispersion. Intuitively, both of these forces lead firms to have different responsiveness and so exposure to aggregate shocks. In Appendix I, we explore each of these in detail (to allow for markup dispersion, we extend our baseline setup

---

[51]Of course, some of the measured mpk dispersion in the data – i.e., the denominators in rows 2 and 4 of Table 4 may be coming from mis-measurement of capital.

to an environment where firms produce differentiated goods, are monopolistically competitive and face constant, but potentially heterogeneous, elasticities of demand). First, we show that a version of our analysis in Section 3 continues to hold in both cases, where the firm's beta now also reflects these additional sources of heterogeneity. Second, we calculate how much of the observed beta dispersion can be attributed to each of these forces. Using dispersion in labor's share of revenue as a likely upper bound for technology dispersion, we find it can potentially account for about 12% of the overall standard deviation of betas from Section 4. Similarly, using recent estimates of markup dispersion among Compustat firms, we find it can account for about 6%. Thus, in total, heterogeneity in input elasticities and markups are likely to explain at most about 18% of measured beta dispersion. Although this is a significant fraction, these findings also suggest that the majority of beta dispersion seems to arise from other sources.[52]

Heterogeneous demand sensitivities. A recent literature has pointed out variation in the response of firm-level demand to the business cycle. For example, Jaimovich et al. (2019) document a "trading down" phenomenon – during expansions, when purchasing power is high, households tend to consume higher quality goods and in downturns substitute towards lower quality ones – Nevo and Wong (2015) show that during the Great Recession, consumers substituted towards cheaper generic products and discount stores and Coibion et al. (2015) show that during downturns, consumers substitute towards low-price retailers.[53] This pattern makes high quality products more procyclical and lower quality ones less so (or even countercyclical). To see the implications of those findings for our analysis, consider the following system of demand and production functions:

$$Q_{it} = P_{it}^{-\mu} \left(X_t^{\hat{\beta}_i} \hat{Z}_{it} \right)^{\mu}, \quad Y_{it} = K_{it}^{\hat{\theta}_1} N_{it}^{\hat{\theta}_2} .$$

Here, X_t is interpreted as an aggregate component of demand rather than technology (it is straightforward to include aggregate technology shocks as well) and $\hat{Z}_{it}$ as idiosyncratic demand. The firm-specific sensitivity to X_t, $\hat{\beta}_i$, captures the idea that in expansions, when demand for all goods is high, consumers substitute towards some goods and away from others. In downturns, when X_t is low, the opposite pattern holds: consumers substitute away from those same goods. This is a simple way to capture the "trading down" phenomenon. Firm revenues are given by

$$P_{it} Y_{it} = X_t^{\hat{\beta}_i} \hat{Z}_{it} K_{it}^{\theta_1} N_{it}^{\theta_2} ,$$

<hr/>

[52]Appendix I also investigates potential heterogeneity in the depreciation rate, δ, and the parameters governing idiosyncratic shocks, ρ_z and $\sigma_{\hat{\varepsilon}}^2$, as well as the effects of adjustment costs alone. We find that these forces are unlikely to account for much of the dispersion in risk premia.

[53]A related literature documents a similar "flight from quality" in response to contractionary exchange rate devaluations. e.g., Burstein et al. (2005), Bems and Di Giovanni (2016) and Chen and Juvenal (2018).

where $\theta_j = \left(1 - \frac{1}{\mu}\right) \hat{\theta}_j$, $j = 1, 2$. With this reinterpretation, the expression is exactly equivalent to (8). In other words, differences in the responsiveness of firm-level demand to the business cycle may be behind our beta estimates.

Since direct data on quality are hard to come by, systematically quantifying the dispersion in these "demand betas" is challenging. However, in Appendix I, we examine one industry where we were able to obtain a proxy for quality, namely average check per person (price) in SIC 5812, Eating Places (i.e., restaurants). Pricing data are from a number of publicly available sources, including company SEC filings and investment bank reports. The appendix shows that higher quality establishments, as proxied by price, have greater exposure to aggregate shocks, and higher expected stock returns and MPK. Thus, the main message of that study of a single industry is that differences in the cyclicality of firm-level demand due to quality differences and "trading down" seems a promising explanation for beta dispersion.

6 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 with heterogeneous firms. The standard optimality condition for investment in this framework suggests that expected firm-level marginal products should reflect exposure to macroeconomic risks, and their pricing. To the extent that firms are differentially exposed to these risks, cross-sectional dispersion in MPK may not only reflect true misallocation, but also risk-adjusted capital allocation. We provide empirical support for this proposition and demonstrate that a suitably parameterized model of firm-level investment suggests that, indeed, risk-adjusted capital allocation accounts for a significant fraction of observed MPK dispersion among US firms. Importantly, much of this dispersion is persistent in nature, which speaks to the large portion of observed MPK dispersion that arises from seemingly persistent/permanent sources. Further, our setup leads to a novel link between aggregate volatility, risk premia and long-run productivity – our results suggest that there can be substantial "productivity costs" of business cycles.

There are several promising directions for future research. Our framework points to a new connection between business cycle dynamics and the cross-sectional allocation of inputs. Investigation of this link, for example, a further exploration of the sources of beta variation across firms, would lead to a better understanding of the underlying causes of observed marginal product dispersion. Much of the misallocation literature examines differences in marginal product dispersion across countries. A natural next step would be to implement a similar analysis in a set of developing countries – because those countries typically have high business cycle volatility, it may be that dispersion in risk premia is larger there. The tractability of our setup allowed

us to quantify the effects of financial market considerations, e.g., cross-sectional variation in required rates of return, on measures of macroeconomic performance, i.e., aggregate TFP. This link provides a new way to evaluate the implications of the rich set of empirical findings in cross-sectional asset pricing. For example, pursuing multifactor/financial shock extensions of our analysis (e.g., along the lines laid out in Appendix F) to incorporate the many risk factors pointed out in that literature would be fruitful to measure the implications of those factors for allocative efficiency. Of particular interest would be whether those factors are efficient or not, e.g., to what extent do capital allocations reflect the "mis-pricing" of assets.

References

ALVAREZ, F. AND U. J. JERMANN (2004): "Using asset prices to measure the cost of business cycles," *Journal of Political economy*, 112, 1223–1256.

ASKER, J., A. COLLARD-WEXLER, AND J. DE LOECKER (2014): "Dynamic inputs and resource (mis) allocation," *Journal of Political Economy*, 122, 1013–1063.

ATKESON, A. AND P. J. KEHOE (2005): "Modeling and measuring organization capital," *Journal of Political Economy*, 113, 1026–1053.

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*.

BARRO, R. J. (2006): "Rare disasters and asset markets in the twentieth century," *The Quarterly Journal of Economics*, 121, 823–866.

BARTELSMAN, E., J. HALTIWANGER, AND S. SCARPETTA (2013): "Cross Country Differences in Productivity: The Role of Allocative Efficiency," *American Economic Review*, 103, 305–334.

BELO, F., X. LIN, AND S. BAZDRESCH (2014): "Labor hiring, investment, and stock return predictability in the cross section," *Journal of Political Economy*, 122, 129–177.

BEMS, R. AND J. DI GIOVANNI (2016): "Income-induced expenditure switching," *American Economic Review*, 106, 3898–3931.

BHARATH, S. T. AND T. SHUMWAY (2008): "Forecasting Default with the Merton Distance to Default Model," *Review of Financial Studies*, 21, 1339–1369.

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.

BURSTEIN, A., M. EICHENBAUM, AND S. REBELO (2005): "Large devaluations and the real exchange rate," *Journal of political Economy*, 113, 742–784.

CHEN, N. AND L. JUVENAL (2018): "Quality and the great trade collapse," *Journal of Development Economics*, 135, 59–76.

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

COIBION, O., Y. GORODNICHENKO, AND G. H. HONG (2015): "The cyclicality of sales, regular and effective prices: Business cycle and policy implications," *American Economic Review*, 105, 993–1029.

COOPER, R. W. AND J. C. HALTIWANGER (2006): "On the nature of capital adjustment costs," *The Review of Economic Studies*, 73, 611–633.

DAVID, J. M., E. HENRIKSEN, AND I. SIMONOVSKA (2014): "The risky capital of emerging markets," Tech. rep., National Bureau of Economic Research.

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 (2019): "The Sources of Capital Misallocation," *American Economic Review*, 109, 2531–67.

DONANGELO, A., F. GOURIO, M. KEHRIG, AND M. PALACIOS (2018): "The cross-section of labor leverage and equity returns," *Journal of Financial Economics*.

EDMOND, C., V. MIDRIGAN, AND D. Y. XU (2018): "How costly are markups?" Tech. rep., National Bureau of Economic Research.

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

EISFELDT, A. L. AND Y. SHI (2018): "Capital Reallocation," Tech. rep., University of California, Los Angeles.

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

FAMA, E. F. AND J. D. MACBETH (1973): "Risk, Return, and Equilibrium: Empirical Tests," *Journal of Political Economy*, 81, 607–636.

GILCHRIST, S., J. W. SIM, AND E. ZAKRAJŠEK (2013): "Misallocation and financial market frictions: Some direct evidence from the dispersion in borrowing costs," *Review of Economic Dynamics*, 16, 159–176.

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

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., Ş. KALEMLI-ÖZCAN, L. KARABARBOUNIS, AND C. VILLEGAS-SANCHEZ (2017): "Capital Allocation and Productivity in South Europe," *The Quarterly Journal of Economics*, 132, 1915–1967.

GUREN, A. M., A. MCKAY, E. NAKAMURA, AND J. STEINSSON (2018): "Housing Wealth Effects: The Long View," Tech. rep., Working Paper.

HALTIWANGER, J., R. KULICK, AND C. SYVERSON (2018): "Misallocation measures: The distortion that ate the residual," Tech. rep., National Bureau of Economic Research.

HOPENHAYN, H. A. (2014): "Firms, misallocation, and aggregate productivity: A review," *Annu. Rev. Econ.*, 6, 735–770.

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.

İMROHOROĞLU, A. AND Ş. TÜZEL (2014): "Firm-level productivity, risk, and return," *Management Science*, 60, 2073–2090.

JAIMOVICH, N., S. REBELO, AND A. WONG (2019): "Trading down and the business cycle," *Journal of Monetary Economics*.

Jones, C. S. and S. Tuzel (2013): "Inventory Investment and The Cost of Capital," *Journal of Financial Economics*, 107, 557–579.

Kehrig, M. (2015): "The Cyclical Nature of the Productivity Distribution," *Working paper*.

Kehrig, M. and N. Vincent (2017): "Do Firms Mitigate or Magnify Capital Misallocation? Evidence from Planet-Level Data," *Working Paper*.

Kennan, J. (2006): "A note on discrete approximations of continuous distributions," *University of*.

Kogan, L. and D. Papanikolaou (2013): "Firm characteristics and stock returns: The role of investment-specific shocks," *The Review of Financial Studies*, 26, 2718–2759.

Lewellen, J. (2015): "The Cross-section of Expected Stock Returns," *Critical Finance Review*, 4, 1–44.

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.

Lucas, R. E. (1987): *Models of Business Cycles*, vol. 26, Basil Blackwell Oxford.

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.

Nevo, A. and A. Wong (2015): "The elasticity of substitution between time and market goods: Evidence from the Great Recession," Tech. rep., National Bureau of Economic Research.

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

Peters, M. (2016): "Heterogeneous Mark-Ups, Growth and Endogenous Misallocation," *Working Paper*.

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.

——— (2017): "The causes and costs of misallocation," *Journal of Economic Perspectives*, 31, 151–74.

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

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

Appendix: For Online Publication

A   Motivation

A.1   Derivation of equation (3) and examples.

$$
\begin{aligned}
1 &= \mathbb{E}_t \left[M_{t+1} \left(MPK_{it+1} + 1 - \delta \right) \right] \\
&= \mathbb{E}_t \left[M_{t+1} \right] \mathbb{E}_t \left[MPK_{it+1} + 1 - \delta \right] + \mathrm{cov}_t \left(M_{t+1}, MPK_{it+1} \right)
\end{aligned}
$$

The (gross) risk-free rate satisfies $R_{ft} = \frac{1}{\mathbb{E}_t [M_{t+1}]}$. Combining and rearranging yields

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

where α_t, β_{it} and λ_t are as defined in the text.

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

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

CAPM. Clearly, $-\mathrm{cov} \left(M_{t+1}, MPK_{it+1} \right) = b\, \mathrm{cov} \left(r_{mt+1}, MPK_{it+1} \right)$ and $\mathrm{var} \left(M_{t+1} \right) = b^2 \mathrm{var} \left(r_{mt+1} \right)$. 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 (3) gives the CAPM expression in the text.

CRRA preferences. 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 (3) gives the CCAPM expression in the text.

A.2   Extensions

Alphas.   We model alphas as firm-level distortions in discount rates. Specifically, we assume the payoffs of firm i are discounted using

$$
\tilde{M}_{t+1} = M_{t+1} T_{it+1}
$$

where T_{it+1} denotes a firm-specific distortion to the discount factor.

The Euler equation then takes the form

$$1 = (1 - \delta) \mathbb{E}_t \left[\tilde{M}_{t+1} \right] + \mathbb{E}_t \left[\tilde{M}_{it+1} MPK_{it+1} \right]$$

Applying a similar approach as above and, for simplicity, assuming that *expected* discount factors are undistorted, we obtain

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

where α_t and λ_t are as defined in expression (3) and $\beta_{it} = -\frac{\text{cov}_t \left(MPK_{it+1}, \tilde{M}_{it+1} \right)}{\text{var}_t (M_{t+1})}$. Thus, even if all firms have the same dynamic process for MPK, dispersion in expected MPK can arise from differences in the stochastic processes of the discount factors, $\tilde{M}_{it+1}$.

Firm-specific investment prices. Here, we allow firms to face different prices of capital, denoted Q_{it}, so that the cost of new investment in period t is equal to $Q_{it} \left(K_{it+1} - (1 - \delta) K_{it} \right)$. The Euler equation is given by

$$Q_{it} = \mathbb{E}_t \left[M_{t+1} \left(MPK_{it+1} + Q_{it+1} (1 - \delta) \right) \right] \tag{34}$$

and rearranging,

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

where $\alpha_{it} = Q_{it} R_{ft} - (1 - \delta) \mathbb{E}_t \left[Q_{it+1} \right]$ is the (now firm-specific) risk-free cost of capital, $\beta_{it} = -\frac{\text{cov}_t (MPK_{it+1} + (1-\delta) Q_{it+1}, M_{t+1})}{\text{var}_t (M_{t+1})}$ and $\lambda_t = \frac{\text{var}_t (M_{t+1})}{\mathbb{E}_t [M_{t+1}]}$. Thus, even if all firms have the same co-movement of MPK with the SDF, i.e., $\text{cov}_t \left(MPK_{it+1}, M_{t+1} \right)$ is constant across firms, heterogeneity in the co-movement of Q_{it} with the SDF will lead to differences in risk premia across firms and hence in expected MPK.

A.3 MPK and Stock Returns

To derive equation (5), use the Euler equation

$$
\begin{aligned}
1 &= \mathbb{E}_t \left[M_{t+1} \left(MPK_{it+1} + 1 - \delta \right) \right] \\
&= (1 - \delta) \mathbb{E}_t \left[M_{t+1} \right] + \mathbb{E}_t \left[M_{t+1} MPK_{it+1} \right]
\end{aligned}
$$

Assume that MPK and M_{t+1} are jointly log-normal and use the fact that the risk-free rate satisfies $R_{ft} = \frac{1}{\mathbb{E}_t[M_{t+1}]}$ to obtain

$$R_{ft} = (1 - \delta) + e^{\mathbb{E}_t[mpk_{it+1}] + \frac{1}{2}\text{var}_t(mpk_{it+1}) + \text{cov}_t(mpk_{it+1}, m_{t+1})}$$

or, rearranging and suppressing variance terms for simplicity,

$$Empk_{it+1}^e \equiv \mathbb{E}_t[mpk_{it+1}] - \log(r_{ft} + \delta) \approx -\text{cov}_t(mpk_{it+1}, m_{t+1})$$

To derive equation (6), standard techniques give the risk premium on stocks as

$$Er_{it+1}^e \equiv \mathbb{E}_t[r_{it+1}] - r_{ft} \approx -\text{cov}_t(r_{it+1}, m_{t+1})$$

With a single source of aggregate risk, a first order approximation gives the return as its expected value plus terms that are linear in the unexpected shocks, i.e.,

$$r_{it+1} = \mathbb{E}_t[r_{it+1}] + \psi_\varepsilon \varepsilon_{it+1} + \psi \beta_{it} \varepsilon_{t+1}$$

where ψ_ε and ψ are constants of linearization and β_i captures the firm-specific exposure to the aggregate shock (ψ_ε and ε_{it} can either be scalars, or vectors of exposures and realizations of idiosyncratic shocks). Similarly, mpk satisfies

$$mpk_{it+1} = \mathbb{E}_t[mpk_{it+1}] + \varepsilon_{it+1} + \beta_{it} \varepsilon_{t+1}$$

Then,

$$\text{cov}_t(r_{it+1}, m_{t+1}) = \psi \text{cov}_t(mpk_{it+1}, m_{t+1})$$

Substituting yields

$$Er_{it+1}^e \approx -\psi \text{cov}_t(mpk_{it+1}, m_{t+1})$$

For a detailed derivation of these results, see the approach in Appendix D.1 and D.4. For a multifactor version, see Appendix F.1.

B Data

In this appendix, we describe the various data sources used throughout our analysis.

B.1 Sources and Series Construction

We obtain firm-level data from COMPUSTAT and CRSP.[54] We include firms coded as industrial firms from 1965-2015. Our time-series regressions and portfolio sorts use data from 1973-2015, since data on the GZ spread and excess bond (EB) premium begin in 1973 and because there are relatively few industries with at least 10 firms in a given year pre-1973.[55] We further exclude financial firms by dropping those with COMPUSTAT SIC codes that correspond to finance, insurance, and real estate (FIRE, SIC codes 6000-6999). We also exclude firms with missing SIC codes or coded as non-classifiable, as much of our analysis examines within-industry variables. We measure firm revenue using sales from Compustat (series SALE), and capital using the depreciated value of plant, property, and equipment (series PPENT). We measure firm marginal product of capital in logs (up to an additive constant) as the difference between log revenue and capital, $mpk_{it} = y_{it} - k_{it}$. Market capitalization is measured as the price times shares outstanding from CRSP and profitability as the ratio of earnings before interest, taxes, depreciation, and amortization (EBITDA) divided by book assets (AT). We measure market leverage as the ratio of book debt to the sum of market capitalization plus book debt, where book debt is measured as current liabilities (LCT) + 1/2 long term debt (DLTT), following Gilchrist and Zakrajsek (2012).

We obtain data on aggregate risk factors from the following sources. Data on the Fama-French factors are from Kenneth French's website, `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/`, while the Hou et al. (2015) q^5 factors are from `http://global-q.org/factors.html`. Updated data on the price/dividend ratio are from Robert J. Shiller's website, `http://www.econ.yale.edu/~shiller/` and updated measures of the GZ spread and excess bond premium are from Simon Gilchrist's website, `http://people.bu.edu/sgilchri/`.

Computation of betas and expected returns. Here, we describe our procedure to compute betas and expected returns.

We estimate stock market betas by performing time-series regressions of firm-level excess returns (realized returns from CRSP in excess of the risk-free rate), r_{it}^e, on aggregate factors, denoted by the $N \times 1$ vector $\mathbf{F_t}$. For each firm, the specification takes the form

$$r_{it}^e = \alpha_{i\tau} + \beta_{i\tau}\mathbf{F}_t + \epsilon_{it} \tag{35}$$

We estimate these regressions at the quarterly frequency using backwards-looking five-year rolling windows, i.e., for $t \in \{\tau - N_\tau + 1, \tau - \tau_T + 2, ..., \tau\}$, where $\beta_{i\tau}$ denotes the $1 \times N$ vector

---

[54]Source: CRSP®, Center for Research in Security Prices, Booth School of Business, The University of Chicago. Used with permission. All rights reserved.

[55]The results are qualitatively similar if we use data from the full 1965-2015 sample.

of factor loadings and N_τ the length of the window.[56] Under the CAPM, the single risk factor is the aggregate market return. Under the Fama-French 3 factor model, the risk factors are the market return (MKT), the return on a portfolio that is long in small firms and short in large ones (SMB) and the return on a portfolio that is long in high book-to-market firms and short in low ones (HML). Under the Hou et al. (2015) q^5 5 factor model, the risk factors are the market return, the return on a portfolio that is long in small firms and short in large ones, the return on a portfolio that is long in low investment firms and short in high investment ones, the return on a portfolio that is long in high profitability (return on equity) firms and short in low profitability ones and the return on a portfolio that is long in firms with high expected 1-year ahead investment-to-assets changes and short in firms with low ones.

Next, we estimate the following cross-sectional regression in each period:

$$r_{it}^e = \alpha_t + \lambda_t \beta_{it} + \epsilon_{it} \tag{36}$$

where λ_t denotes the $1 \times N$ vector of period t factor risk prices and β_{it} the $N \times 1$ vector of exposures, estimated as just described. We calculate expected stock returns as $\alpha_i + \lambda \beta_{it}$, where β_{it} is as estimated from equation (35), λ is calculated using the estimates from (36), and α_i is calculated as $\alpha_i = \frac{1}{T} \sum_{t=1}^{T} (\alpha_{it} + \epsilon_{it})$ also using the estimates from (36).

Composition-adjusted measures of dispersion. For Table 2, we compute time-series of the cross-sectional dispersion in MPK. Because Compustat is an unbalanced panel with significant changes in the composition of firms over time, it is important to ensure that we measure the variation in dispersion due to changes in firm MPK, rather than additions or deletions from the dataset (especially since many additions and deletions to the Compustat data may not be true firm entry or exit). We therefore compute composition-adjusted measures of the cross-sectional standard deviation in MPK that are only affected by firms who continue on in the dataset. We use the following procedure:

For each set of adjacent periods, e.g., t and $t+1$, we compute the cross-sectional standard deviation in each time period only for those firms that are present in the data in both periods. Taking the difference yields the change from time t to $t+1$ that is due only to changes in the common set of firms. Completing this procedure yields a time-series of changes in the cross-sectional standard deviation of MPK. We then combine this time-series of changes with the initial value of the standard deviation (across all firms in the initial period) to construct a synthetic series for the standard deviation that is not affected by the changing composition of firms in the data.

---

[56]We have also estimated the stock market betas using higher frequency monthly data (and two-year rolling windows) and obtained similar results.

B.2 Expected Return Distribution

Table 5 reports statistics from the cross-sectional distribution of expected returns ($E[r^e]$) and unlevered expected returns ($E[r^a]$), which is a measure of expected asset returns, estimated from the Fama-French model. We de-lever expected returns using an adjustment factor computed from Black-Scholes following the approach in, e.g., Bharath and Shumway (2008) and Gilchrist and Zakrajsek (2012). Specifically, we implement an iterative procedure using data on realized equity volatility, firm debt, and firm market capitalization to compute the implied value of assets and asset volatility. The Black-Scholes equations imply $\mathbb{E}[r^a] \approx \frac{Mkt. \ cap.}{V_A}\Phi(\delta_1)\mathbb{E}[r^e]$, where V_A is the total firm asset value implied by Black-Scholes as a function of the market capitalization of equity, book debt, and realized backwards-looking equity volatility and $\Phi(\delta_1)$ is the Black-Scholes "delta" of equity, as defined in, e.g., Gilchrist and Zakrajsek (2012). We compute the adjustment factor $\frac{Mkt. \ cap.}{V^A}\Phi(\delta_1)$ for each firm using daily data and a 21 day backwards-looking window for equity volatility and then calculate a firm-year adjustment factor by averaging this adjustment factor for each firm-year. Finally, we compute un-levered expected returns for each firm as the product of its expected equity return multiplied by this factor. To find the cross-sectional distribution of within-industry expected returns, we de-mean expected returns by industry-year, keeping industry-years with at least 10 observations. We then add back the means and report the resulting distribution.[57] Figure 1 plots the full cross-sectional distribution of within-industry expected excess asset and equity returns.

Table 5: The Distribution of Expected Excess Returns

Percentile	10th	25th	Mean	75th	90th	Std. Dev.
Panel A: Not Industry-Adjusted						
$E[r^a]$	-3.6%	4.0%	9.8%	17.1%	24.7%	13.2%
$E[r^e]$	-5.3%	6.6%	12.1%	20.6%	28.6%	15.6%
Panel B: Industry-Adjusted						
$E[r^a]$	-3.6%	4.7%	9.8%	16.6%	23.6%	12.7%
$E[r^e]$	-4.6%	6.6%	12.1%	20.3%	28.1%	15.0%

Notes: This table reports the cross-sectional distributions of un-levered expected excess equity returns, $E[r^a]$, and expected excess equity returns, $E[r^e]$. Industry adjustment is done by demeaning each measure of expected returns by industry-year. We then add back the mean returns to these distributions.

B.3 Time-Series Correlations

Table 6 reports contemporaneous correlations between (within-industry) MPK dispersion and indicators of the price of risk and the business cycle.

---

[57]The results are similar if we compute our cross-sectional statistics within each year or industry-year and average over the years/industry-years
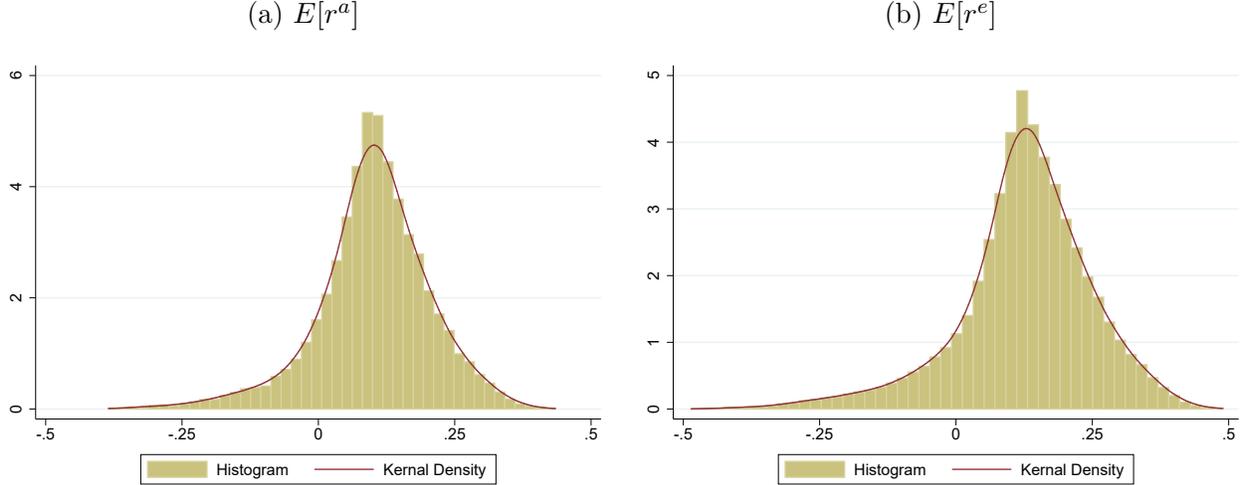
Figure 1: Cross-Sectional Distribution of Expected Excess Returns

Notes: This figure displays the cross-sectional distributions of un-levered expected excess equity returns, $E[r^a]$, and expected excess equity returns, $E[r^e]$. Industry adjustment is done by demeaning each measure of expected returns by industry-year. We then add back the mean returns to these distributions. The vertical bars denote the histograms of these distributions, while the solid lines are the results of kernel smoothing regressions with a bandwidth of 0.25.

B.4 Aggregate Productivity Series

Solow residuals. To build a series of Solow residuals, we obtain data on real GDP and aggregate labor and capital from the Bureau of Economic Analysis. Data on real GDP are from BEA Table 1.1.3 ("Real Gross Domestic Product"), data on labor are from BEA Table 6.4 ("Full-Time and Part-Time Employees") and data on the capital stock are from BEA Table 1.2 ("Net Stock of Fixed Assets"). The data are available annually from 1929-2016. With these data we compute $x_t = y_t - \theta_1 k_t - \theta_2 n_t$. We extract a linear time-trend and then estimate the autoregression in equation (9).

Firm-level series. To construct the alternative series for aggregate productivity from the firm-level data, we use the following procedure. First, we compute firm-level productivity as $z_{it} + \beta_i x_t = y_{it} - \theta k_{it}$. We then average these values across all firms in each year. Because z_{it} is mean-zero and independent across firms, this yields a scaled measure of aggregate productivity, $\bar{\beta} x_t$, where $\bar{\beta}$ is the mean beta across firms, which under our assumptions, is approximately two. We extract a linear time-trend from this series and then estimate the autoregression. The coefficient from this regression gives ρ_x. The standard deviations of the residuals gives $\bar{\beta}\sigma_\varepsilon$ and after dividing by $\bar{\beta}$ gives the true volatility of shocks. Applying this procedure to the set of Compustat firms over the period 1962-2016 yields values of $\rho_x = 0.92$ and $\sigma_\varepsilon = .0245$.

49

Table 6: Correlations of MPK Dispersion, the Price of Risk and the Business Cycle

	MPK Dispersion	PD Ratio	GZ Spread	EB Premium	GDP	TFP
MPK Dispersion	1.00					
PD Ratio	-0.42	1.00				
GZ Spread	0.39	-0.51	1.00			
EB Premium	0.51	-0.57	0.68	1.00		
GDP	-0.53	0.46	-0.59	-0.66	1.00	
TFP	-0.27	0.43	-0.32	-0.44	0.70	1.00

Notes: This table reports time-series correlations of MPK dispersion, measures of the price of risk and the business cycle. MPK dispersion is measured as the within-industry standard deviation in mpk. The PD ratio is the aggregate stock market price/dividend ratio. The GZ spread and EB (excess bond) premium are measures of credit spreads. GDP is log GDP and TFP is log TFP. We extract the cyclical components of GDP, TFP and the PD ratio using a one-sided Hodrick-Prescott filter. All series are described in more detail in the main text and Appendix B.1. All data are quarterly and are from 1973-2015.

C   Additional Empirical Results

MPK and stock returns.   We perform two additional exercises examining the link between MPK, stock market returns and exposure to aggregate risk. First, we verify that high MPK firms tend to offer higher expected stock market returns. To do so, we group firms into five bins, or portfolios, based on their MPK and assess whether the high MPK groups tend exhibit higher stock market returns than the low MPK groups. We sort firms into five portfolios based on their year t MPK, where portfolio 1 contains low MPK firms and portfolio 5 high MPK ones. The portfolios are rebalanced annually. We then compute four versions of the equal-weighted stock excess stock return to each portfolio: the contemporaneous return, denoted r_t^e, the one-period ahead return, r_{t+1}^e, the three-period ahead return, r_{t+3}^e, and the one-period ahead unlevered return, r_{t+1}^a, which we calculate using an unlimited liability model, $r_{t+1}^a = \frac{Mktcap}{Mktcap+Debt} r_{t+1}^e$.[58] We also compute the excess return on a high-minus-low MPK 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.

Examining firms grouped by MPK helps eliminate firm-specific factors unrelated to MPK that may affect returns and so allows us to hone in on the predictability of excess returns by MPK and follows recent practice in empirical finance, which has generally moved from addressing variation in individual firm returns to returns on portfolios of firms, sorted by factors that are known to predict returns. In our context, however, this is likely to provide only a noisy measure of the true relationship between MPK and risk premia. For example, the main text studies a number of additional reasons why MPK may differ across firms, e.g., the

---

[58]When computing one-period ahead returns, we follow Fama and French (1992) and associate the MPK for fiscal year t with returns from July of year $t+1$ to June of year $t+2$. Similar timing holds for three-period ahead returns. Value-weighted portfolios yield similar magnitudes, though the standard errors are greater in some specifications since value-weighting the smaller within-industry samples can increase the variance of portfolio returns. In our log-normal model, equal-weighted dispersion is the key object of interest.

realization of unanticipated shocks (equation (14), capital adjustment costs (Section 3.2), other frictions/distortions (Section 4.3), etc. Each of these forces influence MPK and thus affect the sorting variable (e.g., the high MPK bin includes firms with a high risk premium, but also firms with a low risk premium but high realization of the idiosyncratic shock or a large distortion). The effects of this additional noise is analogous to that of measurement error in the right-hand side variable of a regression, attenuating the true relationship between the variables. Our two-step strategy in Table 1 is designed in part to address this concern. However, we think it useful to examine whether MPK is associated with stock returns using this simpler approach, with the caveat that the quantitative magnitudes should be interpreted with caution and likely represent a lower bound.

The focus of our analysis (and the misallocation literature more broadly) is on within-industry variation in MPK and so to control for industry effects, we demean firm-level mpk by industry-year and sort firms based on this de-meaned measure.[59] For completeness we also present results for total, non industry-adjusted sorts, since the non-adjusted results may be interesting in their own right (discussed more below) and confirm that the link between MPK and stock returns holds at various levels of aggregation.

We report within-industry results in Panel A of Table 7. The table reveals a strong relationship between MPK and stock returns – high MPK portfolios tend to earn high excess returns. The first row shows that the difference in contemporaneous returns between high and low MPK firms, i.e., the excess return on the MPK-HML portfolio, is over 8% annually. The second row confirms that this finding does not simply result from the simultaneous response of stock returns and MPK to the realization of unexpected shocks – one-period ahead excess returns are in fact predictable by MPK. The predictable spread on the MPK-HML portfolio is over 2.5% annually. Both the contemporaneous and future MPK-HML spreads are statistically different from zero at the 99% level. The last two rows of Panel A confirm that the results continues to hold when examining returns further in the future and thus exhibits persistence and after de-levering equity returns. Thus, high MPK firms tend to offer high stock returns, both in a realized and an expected sense, suggesting that MPK differences reflect exposure to risk factors for which investors demand compensation in the form of a higher rate of return.[60]

Panel B of Table 7 reports the "total" results not controlling for industry. Comparing the two panels shows that the relationship between MPK and returns is even stronger when taken

---

[59]There may be heterogeneity across industries on a number of dimensions, for example, in production function coefficients or industry-level exposure to aggregate shocks.

[60]Although there are several measurement differences, the results in Table 7 are related to the "profitability premium" documented in Novy-Marx (2013) and others, i.e., high profit-to-capital firms earn high excess returns (both industry and non industry-adjusted). Further, Novy-Marx (2013) finds that the sales-to-assets component of profitability is the most directly related to higher returns (Appendix A.2 in that paper).

unconditionally across industries, suggesting that there is indeed an industry-level component of excess returns that is predictable by an industry-level component of MPK. Although we do not explore this finding in more detail, it is reassuring confirmation of the link we are after – firms in industries with high average MPK tend to offer higher returns (in a predictable sense) than firms in low MPK industries, suggesting that industry-level exposures to aggregate risk factors may be important as well. Note that across all the variations reported in Table 7, the within-industry effects are well over half of the total, implying a key role for the within-industry component.[61]

Table 7: Excess Returns on MPK-Sorted Portfolios

Portfolio	Low	2	3	4	High	MPK-HML
			Panel A: Within-Industry			
r_t^e	6.98	8.91**	10.59***	12.28***	15.78***	8.80***
	(1.63)	(2.52)	(3.05)	(3.30)	(3.73)	(9.54)
r_{t+1}^e	11.10***	11.55***	12.71***	12.70***	13.69***	2.59***
	(2.61)	(3.35)	(3.75)	(3.50)	(3.36)	(2.98)
r_{t+3}^e	11.95***	12.27***	12.04***	12.60***	13.82***	1.87**
	(2.99)	(3.71)	(3.75)	(3.60)	(3.58)	(2.22)
r_{t+1}^a	6.86**	7.16***	8.04***	8.17***	8.84***	1.97***
	(2.13)	(2.94)	(3.37)	(3.15)	(3.04)	(2.66)
			Panel B: Total			
r_t^e	7.00**	9.08**	10.67***	12.00***	15.25***	8.25***
	(2.01)	(2.53)	(2.93)	(3.09)	(3.71)	(4.54)
r_{t+1}^e	8.60**	12.27***	13.48***	13.73***	13.48***	4.87***
	(2.48)	(3.47)	(3.80)	(3.62)	(3.36)	(2.81)
r_{t+3}^e	9.63***	12.43***	12.69***	13.90***	12.99***	3.36*
	(2.96)	(3.69)	(3.71)	(3.81)	(3.38)	(1.96)
r_{t+1}^a	4.64*	7.53***	8.69***	8.66***	8.22***	3.58***
	(1.88)	(3.07)	(3.53)	(3.26)	(3.02)	(3.05)

Notes: This table reports stock market returns for portfolios sorted by mpk. r_t^e denotes equal-weighted contemporaneous annualized monthly excess stock returns (over the risk-free rate) measured in the year of the portfolio formation from January to December of year t. r_{t+1}^e denotes the analogous future returns, measured from July of year $t+1$ to June of year $t+2$ and r_{t+3}^e from July of year $t+3$ to June of year $t+4$. r_{t+1}^a denotes equal-weighted unlevered ("asset") returns from from July of year $t+1$ to June of year $t+2$, where we use an unlimited liability model to unlever equity returns. Industry adjustment is done by de-meaning mpk by industry-year and sorting portfolios on de-meaned mpk, where industries are defined at the 4-digit SIC code level. t-statistics in parentheses, computed using Newey-West standard errors. Significance levels are denoted by: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

MPK and measures of risk exposure. Next, we directly relate firm MPK to measures of risk exposure using the sensitivity of MPK to aggregate risk factors. To do so, we estimate

---

[61]In unreported results, we have verified that the relationship between MPK and returns continues to hold when we expand the number of portfolios (to 10) and control for size and book-to-market (though they are both correlated with MPK).

regressions of the form

$$mpk_{it+1} = \psi_0 + \psi_\beta \beta_{it} + \zeta_{it+1} , \qquad (37)$$

where β_{it} is a measure of firm i's MPK exposure to aggregate risk at time t. Although there are several important measurement concerns in calculating these exposures compared to stock market-based measures – e.g., they are lower frequency, may be more prone to issues of measurement/sampling error (long enough samples of MPK are only available for a subset of firms) and require assumptions about the factor structure of aggregate risk in MPK that is less well-explored than that in stock returns – it can still be useful to examine the relationship of these exposures to firm-level MPK, keeping in mind these important caveats.[62]

We calculate measures of MPK exposure using the CAPM and Fama-French models. Specifically, we follow an analogous procedure to (35) and (36), replacing excess stock market returns on the left-hand side of (35) and (36) with mpk_{it}. The first regression yields measures of β_{MPK}, i.e., the exposure of each firm's MPK to the aggregate risk factors. The second regression combines these exposures into a single value in the multi-factor Fama-French model using the coefficients from cross-sectional Fama and MacBeth (1973) regressions, i.e., as

$$\beta_{it,FF} = \lambda \beta_{it} = \sum_x \lambda_x \beta_{it,x}, \ x \in MKT, HML, SMB$$

where $\lambda_{\mathbf{x}} = \frac{1}{T} \sum_{t=1}^{T} \lambda_{\mathbf{x}t}$.

We estimate (37) at an annual frequency and lag the right-hand side variable to control for the simultaneous effect of unexpected shocks on contemporaneous measures of beta and MPK. We report the results in columns (1)-(2) in Table 8. The estimates imply that firm-level MPK is significantly related to the sensitivity of MPK to measures of aggregate risk, i.e., the aggregate market return and the three Fama-French factors.[63] In columns (3)-(4), we estimate analogous regressions with the addition of industry-year fixed effects and a set of standard firm-level controls, namely, market capitalization, book-to-market ratio, profitability, and market leverage.[64] All of the coefficients remain positive and statistically significant. Thus, the results help confirm a key implication of expression (3): firm-level risk exposures – measured using "MPK betas" – are significantly related to firm-level expected MPK.

---

[62]Our two-stage approach in Section 2.2 helps deal with some of these issues.

[63]We report two-way clustered standard errors by firm and industry-year to allow for arbitrary time-series correlations for a given firm and for correlations across firms within an industry at a particular time. These standard errors do not account for the error associated with the generated regressors (betas). As in Guren, McKay, Nakamura, and Steinsson (2018), this requires a bootstrap procedure that clusters only on time but precludes clustering on other dimensions. In unreported results, we follow Guren, McKay, Nakamura, and Steinsson (2018) and perform such a bootstrap. The estimates remain significant across almost all specifications.

[64]We describe these series in Appendix B.1.

Table 8: Regressions of MPK on MPK Risk Exposures

	(1)	(2)	(3)	(4)
$\beta_{CAPM,MPK}$	0.065***		0.024***	
	(5.46)		(3.80)	
$\beta_{FF,MPK}$		4.005***		1.097***
		(9.49)		(4.68)
Observations	79404	78920	72477	71990
F.E.	No	No	Yes	Yes
Controls	No	No	Yes	Yes

Notes: This table reports the results of a panel regression of year-ahead *mpk* regressed on measures of firm *mpk* exposure to aggregate risk. Each observation is a firm-year. The dataset contains approximately 10,000 unique firms. F.E. denotes the presence of industry-year fixed effects. Standard errors are two-way clustered by firm and industry-year. *t*-statistics in parentheses. Significance levels are denoted by: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

MPK dispersion and risk premia dispersion. An additional implication of expression (4) is that across groups of firms or segments of the economy, dispersion in expected MPK should be positively related to dispersion in risk permia. We investigate this implication using variation in the dispersion of expected stock market returns and measured risk exposures across industries. Specifically, for each industry in each year, we compute the standard deviation of MPK, $\sigma(mpk)$, expected stock returns, $\sigma(\mathbb{E}[r])$, and various measures of beta, $\sigma(\beta)$. We then estimate regressions of industry-level MPK dispersion on the dispersion in expected returns and betas, i.e.,

$$\sigma(mpk_{jt+1}) = \psi_0 + \psi_1\sigma(x_{jt}) + \zeta_{jt+1} \quad x_{jt} = \mathbb{E}[r_{jt}], \beta_{jt} ,$$

where j denotes industry. To avoid potential simultaneity biases from the realization of shocks, we lag the independent variables (dispersion in expected returns and betas).

Table 9 reports the results of these regressions and verifies that industries with higher dispersion in expected stock returns and risk exposures exhibit greater dispersion in MPK. Column (1) reveals this fact using expected returns calculated from the Fama-French model. Variation in expected return dispersion predicted by the Fama-French model explains over 20% of the variation in MPK dispersion across industry-years. Column (2) regresses MPK dispersion on dispersion in each of the three individual factors – variation in the beta on each factor is significantly related to MPK dispersion. Next, we repeat the exercise using dispersion in MPK betas (described above) as the right-hand side variables. The results in column (3) show that industries with greater dispersion in MPK betas (on each of the Fama-French factors) exhibit greater dispersion in MPK. Columns (4)-(6) add year fixed effects and a number of controls capturing additional measures of firm heterogeneity within industries – the standard deviations of profitability, size, book-to-market, and market leverage. Across these specifications, measures of within-industry heterogeneity in expected returns and aggregate risk

exposures remain positive and significant predictors of within-industry dispersion in MPK.[65]

Table 9: Industry-Level Dispersion in MPK, Expected Stock Returns and Beta

	(1)	(2)	(3)	(4)	(5)	(6)
$\sigma(\mathbb{E}[r])$	2.71***			1.20***		
	(30.11)			(9.82)		
$\sigma(\beta_{MKT})$		0.11***			0.08***	
		(6.48)			(3.31)	
$\sigma(\beta_{HML})$		0.14***			0.10***	
		(11.18)			(5.61)	
$\sigma(\beta_{SMB})$		0.14***			0.07***	
		(13.72)			(5.77)	
$\sigma(\beta_{CAPM,MPK})$			0.01***			0.09***
			(8.58)			(4.08)
$\sigma(\beta_{HML,MPK})$			0.06***			0.06***
			(7.96)			(4.80)
$\sigma(\beta_{SMB,MPK})$			0.06***			0.06***
			(10.38)			(5.70)
Observations	3203	3210	2398	3188	3194	2380
R^2	0.221	0.265	0.200	0.261	0.285	0.348
Industries	157	161	142	153	156	138
Year F.E.	No	No	No	Yes	Yes	Yes
Controls	No	No	No	Yes	Yes	Yes

Notes: This table reports a panel regression of the dispersion in *mpk* within industries on lagged measures of dispersion in risk exposure within those industries. An observation is an industry-year. $\mathbb{E}[r]$ is the expected return computed from the Fama-French model. β denotes the stock return beta on the Fama-French factors and β_{MPK} the *mpk* beta on the same factors. *t*-statistics are in parentheses. Significance levels are denoted by: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Bootstrapped standard errors for Table 1. Table 1 reports standard errors that are two-way clustered by firm and year, but do not account for the estimation error in the measures of risk exposure used in the first stage, and possibly are limited by the parametric assumptions imposed by the regression specification. To address this concern, we have performed a custom bootstrapping procedure that runs a block-bootstrap three ways on our entire procedure, encompassing (i) our rolling-window estimation of firm risk exposures (betas), (ii) the estimation of firm risk premia from these betas, and (iii) the estimation of the regression of future MPK on firm risk premia. First, we randomly sample those firms in our dataset (with replacement) which report sufficient data to compute MPK and appear for long enough to compute measures of risk exposure. Second, we randomly sample the time-periods to use for our first and second

---

[65]The results are robust to using different asset pricing models to compute betas and expected returns, such as the CAPM and Hou et al. (2015) investment-CAPM models. The relationship is robust to a variety of different controls and industry definitions as well. Finally, the results are qualitatively similar when we use the inter-quartile range instead of the standard deviation as our measure of within-industry dispersion.

stage procedures. Third, within each backwards-looking rolling window used to compute the betas, we randomly sample the time periods of observations used (but use the same time periods for each firm).[66] We run this bootstrapping procedure 250 times for each specification. We compute the standard deviation of these estimates, adjusted for the sample size of the re-sampled regressions.[67] We find that the regression coefficients in Table 1 remain significant, except for those using the CAPM model. The implied t-statistics for specifications (1)-(6) are, respectively, 0.36, 2.29, 5.53, 0.36, 2.53, and 6.00.[68]

D   Baseline Model

This appendix provides detailed derivations for the baseline model and analysis.

D.1   Solution – No Adjustment Costs

The static labor choice solves

$$\max e^{\hat{z}_{it}+\hat{\beta}_i x_t} K_{it}^{\theta_1} N_{it}^{\theta_2} - W_t N_{it}$$

with the associated first order condition

$$N_{it} = \left(\frac{\theta_2 e^{\hat{z}_{it}+\hat{\beta}_i x_t} K_{it}^{\theta_1}}{W_t} \right)^{\frac{1}{1-\theta_2}}$$

Substituting for the wage with $W_t = X_t^\omega$ and rearranging gives operating profits

$$\Pi_{it} = G e^{\beta_i x_t + z_{it}} K_{it}^{\theta}$$

where $G \equiv (1-\theta_2) \theta_2^{\frac{\theta_2}{1-\theta_2}}$, $\beta_i = \frac{1}{1-\theta_2} \left(\hat{\beta}_i - \omega \theta_2 \right)$, $z_{it} = \frac{1}{1-\theta_2} \hat{z}_{it}$ and $\theta = \frac{\theta_1}{1-\theta_2}$, which is equation (11) in the text.

---

[66]This will account for potential correlation of estimation error across firms.

[67]Our random sampling of firms and years leads to, on average, fewer observations than in our baseline dataset. We adjust the estimated standard deviation for the lower average number of observations.

[68]In the case of the CAPM, the bootstrapping algorithm generates a few extreme outliers that lead to the high standard deviation and low t-statistics. If we were to use the percentiles of the distribution instead, the p-value would be lower than what the t-statistic implies. These extreme outliers do not occur with the Fama-French or q^5 models, which are known to be better at matching the cross-sectional distribution of risk premia.

The first order and envelope conditions associated with (1) give the Euler equation:

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

Substituting for m_{t+1} and rearranging,

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

and

$$\mathbb{E}_t \left[M_{t+1} \right] = \mathbb{E}_t \left[e^{\log \rho - \gamma_t \varepsilon_{t+1} - \frac{1}{2}\gamma_t^2 \sigma_\varepsilon^2} \right] = e^{\log \rho + \frac{1}{2}\gamma_t^2 \sigma_\varepsilon^2 - \frac{1}{2}\gamma_t^2 \sigma_\varepsilon^2} = \rho$$

so that

$$\theta G K_{it+1}^{\theta-1} = \frac{1 - (1-\delta)\rho}{e^{\log \rho + \rho_z z_{it} + \beta_i \rho_x x_t + \frac{1}{2}\sigma_{\tilde{\varepsilon}}^2 + \frac{1}{2}\beta_i^2 \sigma_\varepsilon^2 - \beta_i \gamma_t \sigma_\varepsilon^2}}$$

and rearranging and taking logs,

$$k_{it+1} = \frac{1}{1-\theta} \left(\tilde{\alpha} + \frac{1}{2}\sigma_{\tilde{\varepsilon}}^2 + \frac{1}{2}\beta_i^2 \sigma_\varepsilon^2 + \rho_z z_{it} + \beta_i \rho_x x_t - \beta_i \gamma_t \sigma_\varepsilon^2 \right)$$

where

$$\tilde{\alpha} = \log \theta + \log G - \alpha$$
$$\alpha = -\log \rho + \log (1 - (1-\delta)\rho) = \log (r_f + \delta)$$

Ignoring the variance terms gives equation (12).

The realized mpk is given by

$$mpk_{it+1} = \log \theta + \pi_{it+1} - k_{it+1}$$
$$= \log \theta + \log G + z_{it+1} + \beta_i x_{t+1} - (1-\theta) k_{it+1}$$
$$= \log \theta + \log G + z_{it+1} + \beta_i x_{t+1} - \tilde{\alpha} - \rho_z z_{it} - \beta_i \rho_x x_t + \beta_i \gamma_t \sigma_\varepsilon^2$$
$$= \alpha + \varepsilon_{it+1} + \beta_i \varepsilon_{t+1} + \beta_i \gamma_t \sigma_\varepsilon^2$$

The time t conditional expected mpk is

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

and the time t and mean cross-sectional variances are, respectively,

$$
\begin{aligned}
\sigma^2_{\mathbb{E}_t[mpk_{it+1}]} &= \sigma^2_\beta \left(\gamma_t \sigma^2_\varepsilon\right)^2 \\
\mathbb{E}\left[\sigma^2_{\mathbb{E}_t[mpk_{it+1}]}\right] &= \mathbb{E}\left[\sigma^2_\beta \left(\gamma_0 + \gamma_1 x_t\right)^2 \left(\sigma^2_\varepsilon\right)^2\right] = \sigma^2_\beta \left(\gamma_0^2 + \gamma_1^2 \sigma^2_x\right) \left(\sigma^2_\varepsilon\right)^2
\end{aligned}
$$

D.2 Solution – Adjustment Costs

With capital adjustment costs, the firm's investment problem takes the form

$$
\begin{aligned}
V\left(X_t, Z_{it}, K_{it}\right) = \max_{K_{it+1}} \quad & G X_t^{\beta_i} Z_t K_{it}^\theta - K_{it+1} + (1-\delta) K_{it} - \Phi\left(I_{it}, K_{it}\right) \tag{38} \\
& + \ \mathbb{E}_t\left[M_{t+1} V\left(X_{t+1}, Z_{it+1}, K_{it+1}\right)\right]
\end{aligned}
$$

Policy function. The first order and envelope conditions associated with (38) give the Euler equation:

$$
\begin{aligned}
1 + \xi\left(\frac{K_{it+1}}{K_{it}} - 1\right) &= \mathbb{E}_t\left[M_{t+1}\left(G\theta e^{z_{it+1}+\beta_i x_{t+1}} K_{it+1}^{\theta-1} + 1 - \delta - \frac{\xi}{2}\left(\frac{K_{it+2}}{K_{it+1}} - 1\right)^2 + \xi\left(\frac{K_{it+2}}{K_{it+1}} - 1\right)\frac{K_{it+2}}{K_{it+1}}\right)\right] \\
&= \mathbb{E}_t\left[M_{t+1}\left(G\theta e^{z_{it+1}+\beta_i x_{t+1}} K_{it+1}^{\theta-1} + 1 - \delta + \frac{\xi}{2}\left(\frac{K_{it+2}}{K_{it+1}}\right)^2 - \frac{\xi}{2}\right)\right]
\end{aligned}
$$

In the non-stochastic steady state,

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

Define the investment return:

$$
R^I_{it+1} = \frac{G\theta e^{z_{it+1}+\beta_i x_{t+1}} K_{it+1}^{\theta-1} + 1 - \delta + \frac{\xi}{2}\left(\frac{K_{it+2}}{K_{it+1}}\right)^2 - \frac{\xi}{2}}{1 + \xi\left(\frac{K_{it+1}}{K_{it}} - 1\right)}
$$

and log-linearizing,

$$
\begin{aligned}
r_{it+1}^I = {} & \rho G \theta K^{\theta-1} \left(z_{it+1} + \beta_i x_{t+1} \right) + \left(\rho G \theta \left(\theta - 1 \right) K^{\theta-1} - \xi \left(1 + \rho \right) \right) k_{it+1} + \rho \xi k_{it+2} + \xi k_{it} \\
& - \log \rho - \rho G \theta \left(\theta - 1 \right) K^{\theta-1} k
\end{aligned}
$$

where $k = \log K$. Rearranging and suppressing constants yields expression (25).

To derive the investment policy function, conjecture it takes the form

$$
k_{it+1} = \phi_{0i} + \phi_1 \beta_i x_t + \phi_2 z_{it} + \phi_3 k_{it}
$$

Then,

$$
k_{it+2} = \phi_{0i} \left(1 + \phi_3 \right) + \phi_1 \beta_i \left(\rho_x + \phi_3 \right) x_t + \phi_2 \left(\rho_z + \phi_3 \right) z_{it} + \phi_3^2 k_{it} + \phi_1 \beta_i \varepsilon_{t+1} + \phi_2 \varepsilon_{it+1}
$$

Substituting into the investment return,

$$
\begin{aligned}
r_{it+1}^I = {} & \left(\rho G \theta \left(\theta - 1 \right) K^{\theta-1} - \xi \left(1 - \rho \phi_3 \right) \right) \phi_{0i} - \log \rho - \rho G \theta \left(\theta - 1 \right) K^{\theta-1} k \\
& + \left(\rho G \theta K^{\theta-1} \rho_z + \left(\rho G \theta \left(\theta - 1 \right) K^{\theta-1} - \xi \left(1 + \rho \right) \right) \phi_2 + \rho \xi \left(\rho_z + \phi_3 \right) \phi_2 \right) z_{it} \\
& + \left(\rho G \theta K^{\theta-1} \rho_x + \left(\rho G \theta \left(\theta - 1 \right) K^{\theta-1} - \xi \left(1 + \rho \right) \right) \phi_1 + \rho \xi \left(\rho_x + \phi_3 \right) \phi_1 \right) \beta_i x_t \\
& + \left(\left(\rho G \theta \left(\theta - 1 \right) K^{\theta-1} - \xi \left(1 + \rho \right) \right) \phi_3 + \rho \xi \phi_3^2 + \xi \right) k_{it} \\
& + \left(\rho G \theta K^{\theta-1} + \rho \xi \phi_2 \right) \varepsilon_{it+1} + \left(\rho G \theta K^{\theta-1} + \rho \xi \phi_1 \right) \beta_i \varepsilon_{t+1}
\end{aligned}
$$

and

$$
\begin{aligned}
r_{it+1}^I + m_{it+1} = {} & \left(\rho G \theta \left(\theta - 1 \right) K^{\theta-1} - \xi \left(1 - \rho \phi_3 \right) \right) \phi_{0i} - \rho G \theta \left(\theta - 1 \right) K^{\theta-1} k - \frac{1}{2} \gamma_0^2 \sigma_\varepsilon^2 - \frac{1}{2} \gamma_1^2 \sigma_\varepsilon^2 x_t^2 \\
& + \left(\rho G \theta K^{\theta-1} \rho_z + \left(\rho G \theta \left(\theta - 1 \right) K^{\theta-1} - \xi \left(1 + \rho \right) \right) \phi_2 + \rho \xi \left(\rho_z + \phi_3 \right) \phi_2 \right) z_{it} \\
& + \left(\left(\rho G \theta K^{\theta-1} \rho_x + \left(\rho G \theta \left(\theta - 1 \right) K^{\theta-1} - \xi \left(1 + \rho \right) \right) \phi_1 + \rho \xi \left(\rho_x + \phi_3 \right) \phi_1 \right) \beta_i - \gamma_0 \gamma_1 \sigma_\varepsilon^2 \right) x_t \\
& + \left(\left(\rho G \theta \left(\theta - 1 \right) K^{\theta-1} - \xi \left(1 + \rho \right) \right) \phi_3 + \rho \xi \phi_3^2 + \xi \right) k_{it} \\
& + \left(\rho G \theta K^{\theta-1} + \rho \xi \phi_2 \right) \varepsilon_{it+1} + \left(\left(\rho G \theta K^{\theta-1} + \rho \xi \phi_1 \right) \beta_i - \gamma_0 - \gamma_1 x_t \right) \varepsilon_{t+1}
\end{aligned}
$$

The Euler equation governing the investment return implies

$$
\begin{aligned}
0 &= \mathbb{E}_t\left[r^I_{it+1} + m_{t+1}\right] + \frac{1}{2}\mathrm{var}\left(r^I_{it+1} + m_{it+1}\right) \\
&= \left(\rho G\theta\left(\theta-1\right)K^{\theta-1} - \xi\left(1-\rho\phi_3\right)\right)\phi_{0i} - \rho G\theta\left(\theta-1\right)K^{\theta-1}k \\
&\quad + \left(\rho G\theta K^{\theta-1}\rho_z + \left(\rho G\theta\left(\theta-1\right)K^{\theta-1} - \xi\left(1+\rho\right)\right)\phi_2 + \rho\xi\left(\rho_z+\phi_3\right)\phi_2\right)z_{it} \\
&\quad + \left(\rho G\theta K^{\theta-1}\rho_x + \left(\rho G\theta\left(\theta-1\right)K^{\theta-1} - \xi\left(1+\rho\right)\right)\phi_1 + \rho\xi\left(\rho_x+\phi_3\right)\phi_1 - \left(\rho G\theta K^{\theta-1} + \rho\xi\phi_1\right)\gamma_1\sigma_\varepsilon^2\right)\beta_i x_t \\
&\quad + \left(\left(\rho G\theta\left(\theta-1\right)K^{\theta-1} - \xi\left(1+\rho\right)\right)\phi_3 + \rho\xi\phi_3^2 + \xi\right)k_{it} \\
&\quad + \frac{1}{2}\left(\rho G\theta K^{\theta-1} + \rho\xi\phi_2\right)^2\sigma_{\tilde{\varepsilon}}^2 \\
&\quad + \frac{1}{2}\left(\rho G\theta K^{\theta-1} + \rho\xi\phi_1\right)^2\beta_i^2\sigma_\varepsilon^2 - \left(\rho G\theta K^{\theta-1} + \rho\xi\phi_1\right)\beta_i\gamma_0\sigma_\varepsilon^2
\end{aligned}
$$

and we can solve for the coefficients from:

$$
\begin{aligned}
0 &= \left(\rho G\theta\left(\theta-1\right)K^{\theta-1} - \xi\left(1-\rho\phi_3\right)\right)\phi_{0i} - \rho G\theta\left(\theta-1\right)K^{\theta-1}k \\
&\quad + \frac{1}{2}\left(\rho G\theta K^{\theta-1} + \rho\xi\phi_2\right)^2\sigma_{\tilde{\varepsilon}}^2 \\
&\quad + \frac{1}{2}\left(\rho G\theta K^{\theta-1} + \rho\xi\phi_1\right)^2\beta_i^2\sigma_\varepsilon^2 - \left(\rho G\theta K^{\theta-1} + \rho\xi\phi_1\right)\beta_i\gamma_0\sigma_\varepsilon^2 \\
&= \rho G\theta K^{\theta-1}\rho_z + \left(\rho G\theta\left(\theta-1\right)K^{\theta-1} - \xi\left(1+\rho\right)\right)\phi_2 + \rho\xi\left(\rho_z+\phi_3\right)\phi_2 \\
&= \rho G\theta K^{\theta-1}\rho_x + \left(\rho G\theta\left(\theta-1\right)K^{\theta-1} - \xi\left(1+\rho\right)\right)\phi_1 + \rho\xi\left(\rho_x+\phi_3\right)\phi_1 - \left(\rho G\theta K^{\theta-1} + \rho\xi\phi_1\right)\gamma_1\sigma_\varepsilon^2 \\
&= \left(\rho G\theta\left(\theta-1\right)K^{\theta-1} - \xi\left(1+\rho\right)\right)\phi_3 + \rho\xi\phi_3^2 + \xi
\end{aligned}
$$

Define $\hat{\xi} = \frac{\xi}{\rho G\theta K^{\theta-1}} = \frac{\xi}{1-\rho(1-\delta)}$. Then,

$$
\begin{aligned}
0 &= \left(\left(\theta-1\right) - \hat{\xi}\left(1+\rho\right)\right)\phi_3 + \rho\hat{\xi}\phi_3^2 + \hat{\xi} \\
\phi_1 &= \frac{\left(\rho_x - \gamma_1\sigma_\varepsilon^2\right)\phi_3}{\hat{\xi}\left(1 - \rho\rho_x\phi_3 + \rho\gamma_1\sigma_\varepsilon^2\phi_3\right)} \\
\phi_2 &= \frac{\rho_z\phi_3}{\hat{\xi}\left(1 - \rho\rho_z\phi_3\right)} \\
\phi_{0i} &= \phi_{00} - \phi_{01}\beta_i + \phi_{02}\beta_i^2
\end{aligned}
$$

where

$$\phi_{00} = \frac{\rho G\theta(1-\theta)K^{\theta-1}k + \frac{1}{2}\left(\rho G\theta K^{\theta-1} + \rho\xi\phi_2\right)^2\sigma_{\tilde{\varepsilon}}^2}{\rho G\theta(1-\theta)K^{\theta-1} + \xi(1-\rho\phi_3)}$$

$$\phi_{01} = \frac{\phi_3}{\hat{\xi}(1-\rho\phi_3)}\frac{\gamma_0\sigma_{\varepsilon}^2}{1-\rho\rho_x\phi_3 + \rho\gamma_1\sigma_{\varepsilon}^2\phi_3}$$

$$\phi_{02} = \frac{\rho G\theta K^{\theta-1}\rho\xi\phi_1 + \frac{1}{2}\left(\rho\xi\phi_1\right)^2 + \frac{1}{2}\left(\rho G\theta K^{\theta-1}\right)^2}{\rho G\theta(1-\theta)K^{\theta-1} + \xi(1-\rho\phi_3)}\sigma_{\varepsilon}^2$$

Note that $\frac{\phi_3}{\hat{\xi}}$ goes to $\frac{1}{1-\theta}$ as $\hat{\xi}$ goes to zero and zero as $\hat{\xi}$ goes to infinity. Again ignoring variance terms, and defining $\phi_4 = \frac{\phi_{01}}{\gamma_0\sigma_{\varepsilon}^2}$, the policy function is

$$k_{it+1} = \phi_1\beta_i x_t + \phi_2 z_{it} + \phi_3 k_{it} - \phi_4\beta_i\gamma_0\sigma_{\varepsilon}^2 + \text{constant}$$

which is equation (27) in the text.

Persistent MPK Dispersion. To derive expression (29), take the unconditional expectation of the policy function to obtain

$$\mathbb{E}[k_{it+1}] = -\frac{\phi_4\beta_i\gamma_0\sigma_{\varepsilon}^2}{1-\phi_3}$$

and thus the unconditional expected mpk as

$$\begin{aligned}\mathbb{E}[mpk_{it+1}] &= (\theta-1)\mathbb{E}[k_{it+1}] + \text{constant}\\ &= \frac{(1-\theta)\phi_3}{\hat{\xi}(1-\rho\phi_3)(1-\phi_3)}\frac{1}{1-\rho\phi_3(\rho_x - \gamma_1\sigma_{\varepsilon}^2)}\beta_i\gamma_0\sigma_{\varepsilon}^2 + \text{constant}\end{aligned}$$

where we have substituted using the definition of ϕ_4. Lastly, we can use the definition of ϕ_3 to show that the first fraction is equal to one and thus,

$$\mathbb{E}[mpk_{it+1}] = \frac{1}{1-\rho\phi_3(\rho_x - \gamma_1\sigma_{\varepsilon}^2)}\beta_i\gamma_0\sigma_{\varepsilon}^2 + \text{constant}$$

D.3 Aggregation

The first order condition on labor gives

$$N_{it} = \left(\frac{\theta_2 e^{\hat{\beta}_i x_t + \hat{z}_{it}}K_{it}^{\theta_1}}{W_t}\right)^{\frac{1}{1-\theta_2}}$$

and substituting for the wage,

$$N_{it} = \left(\theta_2 e^{\left(\hat{\beta}_i - \omega\right)x_t + \hat{z}_{it}} K_{it}^{\theta_1}\right)^{\frac{1}{1-\theta_2}}$$

Labor market clearing gives:

$$N_t = \int N_{it} di = \theta_2^{\frac{1}{1-\theta_2}} e^{-\frac{1}{1-\theta_2}\omega x_t} \int e^{\frac{1}{1-\theta_2}\hat{\beta}_i x_t + z_{it}} K_{it}^{\theta} di$$

so that

$$\theta_2^{\frac{\theta_2}{1-\theta_2}} e^{-\frac{\theta_2}{1-\theta_2}\omega x_t} = \left(\frac{N_t}{\int e^{\frac{1}{1-\theta_2}\hat{\beta}_i x_t + z_{it}} K_{it}^{\theta} di}\right)^{\theta_2}$$

Then,

$$
\begin{aligned}
Y_{it} &= e^{\hat{\beta}_i x_t + \hat{z}_{it}} K_{it}^{\theta_1} N_{it}^{\theta_2} = \theta_2^{\frac{\theta_2}{1-\theta_2}} e^{-\frac{\theta_2}{1-\theta_2}\omega x_t} e^{\frac{1}{1-\theta_2}\hat{\beta}_i x_t + z_{it}} K_{it}^{\theta} \\
&= \frac{e^{\frac{1}{1-\theta_2}\hat{\beta}_i x_t + z_{it}} K_{it}^{\theta}}{\left(\int e^{\frac{1}{1-\theta_2}\hat{\beta}_i x_t + z_{it}} K_{it}^{\theta} di\right)^{\theta_2}} N_t^{\theta_2}
\end{aligned}
$$

By definition,

$$MPK_{it} = \frac{\theta e^{\frac{1}{1-\theta_2}\hat{\beta}_i x_t + z_{it}} K_{it}^{\theta-1}}{\left(\int e^{\frac{1}{1-\theta_2}\hat{\beta}_i x_t + z_{it}} K_{it}^{\theta} di\right)^{\theta_2}} N_t^{\theta_2}$$

and rearranging,

$$K_{it} = \left(\frac{\theta e^{\frac{1}{1-\theta_2}\hat{\beta}_i x_t + z_{it}}}{MPK_{it}}\right)^{\frac{1}{1-\theta}} \left(\frac{N_t}{\int e^{\frac{1}{1-\theta_2}\hat{\beta}_i x_t + z_{it}} K_{it}^{\theta} di}\right)^{\frac{\theta_2}{1-\theta}}$$

Capital market clearing gives

$$K_t = \int K_{it} di = \theta^{\frac{1}{1-\theta}} \left(\frac{N_t}{\int e^{\frac{1}{1-\theta_2}\hat{\beta}_i x_t + z_{it}} K_{it}^{\theta} di}\right)^{\frac{\theta_2}{1-\theta}} \int e^{\frac{1}{1-\theta_2}\frac{1}{1-\theta}\hat{\beta}_i x_t + \frac{1}{1-\theta} z_{it}} MPK_{it}^{-\frac{1}{1-\theta}} di$$

so that

$$K_{it}^{\theta} = \left(\frac{e^{\frac{1}{1-\theta_2}\frac{1}{1-\theta}\hat{\beta}_i x_t + \frac{1}{1-\theta} z_{it}} MPK_{it}^{-\frac{1}{1-\theta}}}{\int e^{\frac{1}{1-\theta_2}\frac{1}{1-\theta}\hat{\beta}_i x_t + \frac{1}{1-\theta} z_{it}} MPK_{it}^{-\frac{1}{1-\theta}} di} K_t\right)^{\theta}$$

and substituting into the expression for Y_{it},

$$Y_{it} = \frac{e^{\frac{1}{1-\theta_2}\hat\beta_i x_t + z_{it}}\left(\frac{e^{\frac{1}{1-\theta_2}\frac{1}{1-\theta}\hat\beta_i x_t + \frac{1}{1-\theta}z_{it}}MPK_{it}^{-\frac{1}{1-\theta}}}{\int e^{\frac{1}{1-\theta_2}\frac{1}{1-\theta}\hat\beta_i x_t + \frac{1}{1-\theta}z_{it}}MPK_{it}^{-\frac{1}{1-\theta}}di}K_t\right)^{\theta}}{\left(\int e^{\frac{1}{1-\theta_2}\hat\beta_i x_t + z_{it}}\left(\frac{e^{\frac{1}{1-\theta_2}\frac{1}{1-\theta}\hat\beta_i x_t + \frac{1}{1-\theta}z_{it}}MPK_{it}^{-\frac{1}{1-\theta}}}{\int e^{\frac{1}{1-\theta_2}\frac{1}{1-\theta}\hat\beta_i x_t + \frac{1}{1-\theta}z_{it}}MPK_{it}^{-\frac{1}{1-\theta}}di}K_t\right)^{\theta}di\right)^{\theta_2}}N_t^{\theta_2}$$

$$= \frac{\frac{e^{\frac{1}{1-\theta_2}\frac{1}{1-\theta}\hat\beta_i x_t + \frac{1}{1-\theta}z_{it}}MPK_{it}^{-\frac{\theta}{1-\theta}}}{\left(\int e^{\frac{1}{1-\theta_2}\frac{1}{1-\theta}\hat\beta_i x_t + \frac{1}{1-\theta}z_{it}}MPK_{it}^{-\frac{1}{1-\theta}}di\right)^{\theta}}}{\left(\frac{\int e^{\frac{1}{1-\theta_2}\frac{1}{1-\theta}\hat\beta_i x_t + \frac{1}{1-\theta}z_{it}}MPK_{it}^{-\frac{\theta}{1-\theta}}di}{\left(\int e^{\frac{1}{1-\theta_2}\frac{1}{1-\theta}\hat\beta_i x_t + \frac{1}{1-\theta}z_{it}}MPK_{it}^{-\frac{1}{1-\theta}}di\right)^{\theta}}\right)^{\theta_2}}K_t^{\theta_1}N_t^{\theta_2}$$

Aggregate output is then

$$Y_t = \int Y_{it}di = A_t K_t^{\theta_1}N_t^{\theta_2}$$

where

$$A_t = \left(\frac{\int e^{\frac{1}{1-\theta_2}\frac{1}{1-\theta}\hat\beta_i x_t + \frac{1}{1-\theta}z_{it}}MPK_{it}^{-\frac{\theta}{1-\theta}}di}{\left(\int e^{\frac{1}{1-\theta_2}\frac{1}{1-\theta}\hat\beta_i x_t + \frac{1}{1-\theta}z_{it}}MPK_{it}^{-\frac{1}{1-\theta}}di\right)^{\theta}}\right)^{1-\theta_2}$$

Taking logs,

$$a_t = (1-\theta_2)\left(\log\int e^{\frac{1}{1-\theta_2}\frac{1}{1-\theta}\hat\beta_i x_t + \frac{1}{1-\theta}z_{it}}MPK_{it}^{-\frac{\theta}{1-\theta}}di - \theta\log\int e^{\frac{1}{1-\theta_2}\frac{1}{1-\theta}\hat\beta_i x_t + \frac{1}{1-\theta}z_{it}}MPK_{it}^{-\frac{1}{1-\theta}}di\right)$$

The first expression in braces is equal to

$$\frac{1}{1-\theta}\frac{1}{1-\theta_2}\bar{\hat\beta}x_t - \frac{\theta}{1-\theta}\bar{m}pk + \frac{1}{2}\left(\frac{1}{1-\theta}\right)^2\left(\left(\frac{1}{1-\theta_2}\right)^2 x_t^2\sigma_{\hat\beta}^2 + \sigma_z^2\right) + \frac{1}{2}\left(\frac{\theta}{1-\theta}\right)^2\sigma_{mpk}^2$$

$$- \frac{\theta}{(1-\theta)^2}\frac{1}{1-\theta_2}\sigma_{mpk,\hat\beta_i x_t + z_{it}}$$

and the second to

$$\frac{\theta}{1-\theta}\frac{1}{1-\theta_2}\bar{\hat{\beta}}x_t - \frac{\theta}{1-\theta}\bar{m}pk + \frac{1}{2}\theta\left(\frac{1}{1-\theta}\right)^2\left(\left(\frac{1}{1-\theta_2}\right)^2 x_t^2\sigma_{\hat{\beta}}^2 + \sigma_z^2\right) + \frac{1}{2}\theta\left(\frac{1}{1-\theta}\right)^2\sigma_{mpk}^2$$

$$-\frac{\theta}{(1-\theta)^2}\frac{1}{1-\theta_2}\sigma_{mpk,\hat{\beta}_ix_t+z_{it}}$$

and combining (and using $\sigma_\beta = \frac{1}{1-\theta_2}\sigma_{\hat{\beta}}$) gives

$$\begin{aligned}
a_t &= \bar{\hat{\beta}}x_t + (1-\theta_2)\left(\frac{1}{2}\frac{1}{1-\theta}\left(x_t^2\sigma_\beta^2 + \sigma_z^2\right) - \frac{1}{2}\frac{\theta}{1-\theta}\sigma_{mpk}^2\right) \\
&= a_t^* - \frac{1}{2}(1-\theta_2)\frac{\theta}{1-\theta}\sigma_{mpk}^2 \\
&= a_t^* - \frac{1}{2}\frac{\theta_1(1-\theta_2)}{1-\theta_1-\theta_2}\sigma_{mpk}^2
\end{aligned}$$

D.4 Stock Market Returns

We derive stock market returns in the environment with adjustment costs. This nests the simpler case without them when $\xi = 0$.

Dividends are equal to

$$D_{it+1} = e^{z_{it+1}+\beta_ix_{t+1}}K_{it+1}^\theta - K_{it+2} + (1-\delta)K_{it+1} - \frac{\xi}{2}\left(\frac{K_{it+2}}{K_{it+1}} - 1\right)^2 K_{it+1}$$

and log-linearizing,

$$d_{it+1} = \frac{\Pi}{D}(z_{it+1} + \beta_ix_{t+1}) + \left(\theta\frac{\Pi}{D} + (1-\delta)\frac{K}{D}\right)k_{it+1} - \frac{K}{D}k_{it+2} + \log D - \left(\theta\frac{\Pi}{D} - \delta\frac{K}{D}\right)k$$

where $k = \log K$. Substituting for k_{it+1} and k_{it+2} from Appendix D.2 and rearranging,

$$d_{it+1} = A_{0i} + \tilde{A}_1z_{it} + A_1\beta_ix_t + \tilde{A}_2\varepsilon_{it+1} + A_2\beta_i\varepsilon_{t+1} + A_3k_{it}$$

where

$$A_{0i} = \log D - \left(\theta\frac{\Pi}{D} - \delta\frac{K}{D}\right)(k - \phi_{0i}) - \frac{K}{D}\phi_{0i}\phi_3$$

$$A_1 = \frac{\Pi}{D}\rho_x + \left(\theta\frac{\Pi}{D} + \frac{K}{D}(1 - \delta - \rho_x - \phi_3)\right)\phi_1$$

$$\tilde{A}_1 = \frac{\Pi}{D}\rho_z + \left(\theta\frac{\Pi}{D} + \frac{K}{D}(1 - \delta - \rho_z - \phi_3)\right)\phi_2$$

$$A_2 = \frac{\Pi}{D} - \frac{K}{D}\phi_1$$

$$\tilde{A}_2 = \frac{\Pi}{D} - \frac{K}{D}\phi_2$$

$$A_3 = \left(\theta\frac{\Pi}{D} + \frac{K}{D}(1 - \delta - \phi_3)\right)\phi_3$$

By definition, returns are equal to

$$R_{it+1} = \frac{D_{it+1} + P_{it+1}}{P_{it}}$$

and log-linearizing,

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

Conjecture the stock price takes the form

$$p_{it} = c_{0i} + c_1\beta_i x_t + c_2 z_{it} + c_3 k_{it}$$

Then,

$$\begin{aligned}
r_{it+1} = {}& -\log\rho + (1 - \rho)\left(\log\frac{P}{D} + A_{0i} - c_{0i}\right) + \rho c_3\phi_{0i} \\
& + \left((\rho\rho_z - 1)c_2 + \rho c_3\phi_2 + (1 - \rho)\tilde{A}_1\right)z_{it} \\
& + ((\rho\rho_x - 1)c_1 + \rho c_3\phi_1 + (1 - \rho)A_1)\beta_i x_t \\
& + ((\rho\phi_3 - 1)c_3 + (1 - \rho)A_3)k_{it} \\
& + \left(\rho c_2 + (1 - \rho)\tilde{A}_2\right)\varepsilon_{it+1} + (\rho c_1 + (1 - \rho)A_2)\beta_i\varepsilon_{t+1}
\end{aligned}$$

and the (log) excess return is the (negative of the) conditional covariance with the SDF:

$$\log\mathbb{E}_t\left[R^e_{it+1}\right] = (\rho c_1 + (1 - \rho)A_2)\beta_i\gamma_t\sigma_\varepsilon^2$$

To solve for the coefficients, use the Euler equation. First,

$$
\begin{aligned}
r_{it+1} + m_{it+1} \;=\;& (1-\rho)\left(\log\frac{P}{D} + A_{0i} - c_{0i}\right) + \rho c_3\phi_{0i} - \frac{1}{2}\gamma_0^2\sigma_\varepsilon^2 \\
& + \left((\rho\rho_z - 1)c_2 + \rho c_3\phi_2 + (1-\rho)\tilde{A}_1\right)z_{it} \\
& + \left(((\rho\rho_x - 1)c_1 + \rho c_3\phi_1 + (1-\rho)A_1)\beta_i - \gamma_0\gamma_1\sigma_\varepsilon^2\right)x_t \\
& + ((\rho\phi_3 - 1)c_3 + (1-\rho)A_3)k_{it} \\
& - \frac{1}{2}\gamma_1^2\sigma_\varepsilon^2 x_t^2 \\
& + \left(\rho c_2 + (1-\rho)\tilde{A}_2\right)\varepsilon_{it+1} \\
& + ((\rho c_1 + (1-\rho)A_2)\beta_i - \gamma_0 - \gamma_1 x_t)\varepsilon_{t+1}
\end{aligned}
$$

The Euler equation implies

$$
\begin{aligned}
0 \;=\;& \mathbb{E}_t\left[r_{it+1} + m_{it+1}\right] + \frac{1}{2}\mathrm{var}\left(r_{it+1} + m_{it+1}\right) \\
\;=\;& (1-\rho)\left(\log\frac{P}{D} + A_{0i} - c_{0i}\right) + \rho c_3\phi_{0i} + \frac{1}{2}(\rho c_1 + (1-\rho)A_2)^2\beta_i^2\sigma_\varepsilon^2 - (\rho c_1 + (1-\rho)A_2)\beta_i\gamma_0\sigma_\varepsilon^2 \\
& + \frac{1}{2}\left(\rho c_2 + (1-\rho)\tilde{A}_2\right)^2\sigma_{\tilde\varepsilon}^2 \\
& + \left((\rho\rho_z - 1)c_2 + \rho c_3\phi_2 + (1-\rho)\tilde{A}_1\right)z_{it} \\
& + \left((\rho\rho_x - 1)c_1 + \rho c_3\phi_1 + (1-\rho)A_1 - (\rho c_1 + (1-\rho)A_2)\gamma_1\sigma_\varepsilon^2\right)\beta_i x_t \\
& + ((\rho\phi_3 - 1)c_3 + (1-\rho)A_3)k_{it}
\end{aligned}
$$

and so by undetermined coefficients,

$$
\begin{aligned}
0 \;=\;& (1-\rho)\left(\log\frac{P}{D} + A_{0i} - c_{0i}\right) + \rho c_3\phi_{0i} + \frac{1}{2}(\rho c_1 + (1-\rho)A_2)^2\beta_i^2\sigma_\varepsilon^2 - (\rho c_1 + (1-\rho)A_2)\beta_i\gamma_0\sigma_\varepsilon^2 \\
& + \frac{1}{2}\left(\rho c_2 + (1-\rho)\tilde{A}_2\right)^2\sigma_{\tilde\varepsilon}^2 \\
\;=\;& (\rho\rho_z - 1)c_2 + \rho c_3\phi_2 + (1-\rho)\tilde{A}_1 \\
\;=\;& (\rho\rho_x - 1)c_1 + \rho c_3\phi_1 + (1-\rho)A_1 - (\rho c_1 + (1-\rho)A_2)\gamma_1\sigma_\varepsilon^2 \\
\;=\;& (\rho\phi_3 - 1)c_3 + (1-\rho)A_3
\end{aligned}
$$

or

$$c_3 = \frac{(1 - \rho) A_3}{1 - \rho \phi_3}$$

$$c_2 = \frac{\rho c_3 \phi_2 + (1 - \rho) \tilde{A}_1}{1 - \rho \rho_z}$$

$$c_1 = \frac{\rho c_3 \phi_1 + (1 - \rho) (A_1 - A_2 \gamma_1 \sigma_\varepsilon^2)}{1 - \rho \rho_x + \rho \gamma_1 \sigma_\varepsilon^2}$$

Substituting for c_1 we can solve for

$$\log \mathbb{E}_t \left[R_{it+1}^e \right] = \frac{\rho^2 c_3 \phi_1 + (1 - \rho) (\rho A_1 + (1 - \rho \rho_x) A_2)}{1 - \rho \rho_x + \rho \sigma_\varepsilon^2 \gamma_1} \beta_i \gamma_t \sigma_\varepsilon^2$$

Solving for

$$\rho A_1 + (1 - \rho \rho_x) A_2 = \frac{\frac{1}{\rho} + \delta - 1 - \rho \theta \phi_1 \phi_3}{\frac{1}{\rho} + \delta (1 - \theta) - 1}$$

$$\rho^2 c_3 \phi_1 = \theta \frac{\rho^2 (1 - \rho) \phi_1 \phi_3}{1 - \rho \phi_3} \frac{\frac{1}{\rho} - \phi_3}{\frac{1}{\rho} + \delta (1 - \theta) - 1}$$

substituting into the return equation and simplifying, we obtain

$$\log \mathbb{E}_t \left[R_{it+1}^e \right] = \psi \beta_i \gamma_t \sigma_\varepsilon^2$$

where

$$\psi = \frac{\frac{1}{\rho} + \delta - 1}{\frac{1}{\rho} + \delta (1 - \theta) - 1} \frac{1 - \rho}{1 - \rho \rho_x + \rho \gamma_1 \sigma_\varepsilon^2}$$

which is equation (21) in the text.

The Sharpe ratio is the ratio of expected excess returns to the conditional standard deviation of the return:

$$SR_{it} = \frac{\psi \beta_i \gamma_t \sigma_\varepsilon^2}{\sqrt{\left(\rho c_2 + (1 - \rho) \tilde{A}_2 \right)^2 \sigma_{\tilde{z}}^2 + \psi^2 \beta_i^2 \sigma_\varepsilon^2}}$$

We can solve for

$$\rho c_2 + (1 - \rho) \tilde{A}_2 = \frac{\frac{1}{\rho} + \delta - 1}{\frac{1}{\rho} + \delta (1 - \theta) - 1} \frac{1 - \rho}{1 - \rho \rho_z}$$

and substituting and rearranging gives the expression in footnote 29.

For a perfectly diversified portfolio (i.e., the integral over individual returns) idiosyncratic shocks cancel, i.e., $\sigma_{\tilde{\varepsilon}}^2 = 0$ and $SR_{mt} = \gamma_t \sigma_\varepsilon$.

D.5 Autocorrelation of Investment

To derive the autocorrelation of investment, define net investment as $\Delta k_{it+1} = k_{it+1} - k_{it}$. We use the following:

$$
\begin{aligned}
\operatorname{cov}\left(\Delta z_{it}, z_{it}\right) &= \operatorname{cov}\left(\left(\rho_z - 1\right) z_{it-1} + \varepsilon_{it} . \rho_z z_{it-1} + \varepsilon_{it}\right) \\
&= \rho_z \left(\rho_z - 1\right) \sigma_z^2 + \sigma_{\bar{\varepsilon}}^2 \\
&= \frac{1}{1 + \rho_z} \sigma_{\bar{\varepsilon}}^2 \\
\operatorname{cov}\left(\Delta k_{it}, z_{it}\right) &= \operatorname{cov}\left(\Delta k_{it}, \rho_z z_{it-1} + \varepsilon_{it}\right) \\
&= \rho_z \operatorname{cov}\left(\Delta k_{it}, z_{it-1}\right) \\
&= \rho_z \operatorname{cov}\left(\phi_1 \beta_i \Delta x_{t-1} + \phi_2 \Delta z_{it-1} + \phi_3 \Delta k_{it-1}, z_{it-1}\right) \\
&= \rho_z \left(\operatorname{cov}\left(\phi_2 \Delta z_{it-1}, z_{it-1}\right) + \phi_3 \operatorname{cov}\left(\Delta k_{it-1}, z_{it-1}\right)\right) \\
&= \rho_z \phi_2 \frac{1}{1 + \rho_z} \sigma_{\bar{\varepsilon}}^2 + \rho_z \phi_3 \operatorname{cov}\left(\Delta k_{it-1}, z_{it-1}\right)
\end{aligned}
$$

so that

$$
\mathbb{E}\left[\operatorname{cov}\left(\Delta k_{it}, z_{it}\right)\right] = \frac{\rho_z}{1 + \rho_z} \frac{\phi_2 \sigma_{\bar{\varepsilon}}^2}{1 - \phi_3 \rho_z}
$$

Next,

$$
\begin{aligned}
\operatorname{cov}\left(\Delta k_{it+1}, \Delta z_{it+1}\right) &= \operatorname{cov}\left(\phi_1 \beta_i \Delta x_t + \phi_2 \Delta z_{it} + \phi_3 \Delta k_{it} . \left(\rho_z - 1\right) z_{it} + \varepsilon_{it+1}\right) \\
&= \phi_2 \left(\rho_z - 1\right) \operatorname{cov}\left(\Delta z_{it}, z_{it}\right) + \phi_3 \left(\rho_z - 1\right) \operatorname{cov}\left(\Delta k_{it}, z_{it}\right) \\
&= \phi_2 \left(\rho_z - 1\right) \frac{1}{1 + \rho_z} \sigma_{\bar{\varepsilon}}^2 + \frac{\rho_z}{1 + \rho_z} \frac{\phi_3 \left(\rho_z - 1\right) \phi_2 \sigma_{\bar{\varepsilon}}^2}{1 - \phi_3 \rho_z} \\
&= \frac{\rho_z - 1}{1 + \rho_z} \frac{\phi_2 \sigma_{\bar{\varepsilon}}^2}{1 - \phi_3 \rho_z}
\end{aligned}
$$

Similar steps give

$$
\operatorname{cov}\left(\Delta k_{it+1}, \Delta x_{t+1}\right) = \frac{\rho_x - 1}{1 + \rho_x} \frac{\phi_1 \beta_i \sigma_{\varepsilon}^2}{1 - \phi_3 \rho_x}
$$

Combining these gives the variance of investment:

$$
\begin{aligned}
\sigma_{\Delta k}^2 &= \phi_1^2 \beta_i^2 \mathrm{var}\left(\Delta x_t\right) + \phi_2^2 \mathrm{var}\left(\Delta z_{it}\right) + \phi_3^2 \sigma_{\Delta k}^2 \\
&+ 2\phi_1 \phi_3 \beta_i \mathrm{cov}\left(\Delta x_t, \Delta k_{it}\right) + 2\phi_2 \phi_3 \mathrm{cov}\left(\Delta z_{it}, \Delta k_{it}\right) \\
&= \phi_1^2 \beta_i^2 \frac{2}{1+\rho_x}\sigma_\varepsilon^2 + \phi_2^2 \frac{2}{1+\rho_z}\sigma_{\tilde\varepsilon}^2 + \phi_3^2 \sigma_{\Delta k}^2 \\
&+ \frac{2\phi_1^2 \phi_3 \beta_i^2 \sigma_\varepsilon^2}{1-\phi_3 \rho_x}\frac{\rho_x-1}{1+\rho_x} + \frac{2\phi_2^2 \phi_3 \sigma_{\tilde\varepsilon}^2}{1-\phi_3 \rho_z}\frac{\rho_z-1}{1+\rho_z} \\
&= \phi_3^2 \sigma_{\Delta k}^2 + 2\phi_1^2 \beta_i^2 \frac{\sigma_\varepsilon^2}{1+\rho_x}\left(1 + \frac{\phi_3\left(\rho_x-1\right)}{1-\phi_3\rho_x}\right) + 2\phi_2^2 \frac{\sigma_{\tilde\varepsilon}^2}{1+\rho_z}\left(1 + \frac{\phi_3\left(\rho_z-1\right)}{1-\phi_3\rho_z}\right) \\
&= \frac{2}{1+\phi_3}\left(\phi_1^2 \beta_i^2 \sigma_\varepsilon^2 \frac{1}{1+\rho_x}\frac{1}{1-\phi_3\rho_x} + \phi_2^2 \sigma_{\tilde\varepsilon}^2 \frac{1}{1+\rho_z}\frac{1}{1-\phi_3\rho_z}\right)
\end{aligned}
$$

Next,

$$
\begin{aligned}
\mathrm{cov}\left(\Delta k_{it+1}, \Delta k_{it}\right) &= \mathrm{cov}\left(\phi_1 \beta_i \Delta x_t + \phi_2 \Delta z_{it} + \phi_3 \Delta k_{it}, \Delta k_{it}\right) \\
&= \phi_1 \beta_i \mathrm{cov}\left(\Delta x_t, \Delta k_{it}\right) + \phi_2 \mathrm{cov}\left(\Delta z_{it}, \Delta k_{it}\right) + \phi_3 \sigma_{\Delta k}^2 \\
&= \phi_1^2 \beta_i^2 \sigma_\varepsilon^2 \frac{\rho_x-1}{1+\rho_x}\frac{1}{1-\phi_3\rho_x} + \phi_2^2 \sigma_{\tilde\varepsilon}^2 \frac{\rho_z-1}{1+\rho_z}\frac{1}{1-\phi_3\rho_z} + \phi_3 \sigma_{\Delta k}^2
\end{aligned}
$$

and the autocorrelation is:

$$
\mathrm{corr}\left(\Delta k_{it+1}, \Delta k_{it}\right) = \phi_3 + \frac{1+\phi_3}{2}\frac{\phi_1^2 \beta_i^2 \sigma_\varepsilon^2 \frac{\rho_x-1}{1+\rho_x}\frac{1}{1-\phi_3\rho_x} + \phi_2^2 \sigma_{\tilde\varepsilon}^2 \frac{\rho_z-1}{1+\rho_z}\frac{1}{1-\phi_3\rho_z}}{\phi_1^2 \beta_i^2 \sigma_\varepsilon^2 \frac{1}{1+\rho_x}\frac{1}{1-\phi_3\rho_x} + \phi_2^2 \sigma_{\tilde\varepsilon}^2 \frac{1}{1+\rho_z}\frac{1}{1-\phi_3\rho_z}} \tag{39}
$$

Notice that this approaches

$$
\mathrm{corr}\left(\Delta k_{it+1}, \Delta k_{it}\right) = \phi_3 + \left(1-\phi_3\right)\frac{\rho_x-1}{2}
$$

as ρ_z and ρ_x become close. Further, in the case both shocks follow a random walk, the autocorrelation is simply equal to ϕ_3.

E   Numerical Procedure

Our numerical approach to parameterize the model is as follows. To accurately capture the properties of the time-varying risk premium, we solve for returns numerically using a fourth-order approximation in Dynare++. For a given set of the parameters γ_0, γ_1, ξ and σ_β^2, we solve the model for a wide grid of beta-types centered around the mean beta. We use an 11 point grid ranging from -3 to 7 (the results are not overly sensitive to the width of the grid). We

simulate a time series of excess returns for a large number of firms of each type, which results in a large panel of excess returns. Averaging returns across these firms in each time period yields a series for the market excess return. We can then compute the mean and standard deviation (i.e., Sharpe ratio) of the market return.

Next, we compute the expected return for each beta-type in each time period directly as the conditional expectation $\mathbb{E}_t \left[R_{it+1} \right] = \mathbb{E}_t \left[\frac{D_{it+1} + P_{it+1}}{P_{it}} \right]$ and then average over the time periods to obtain the average expected return for firms of each type. We then use these values to calculate the dispersion in expected returns, σ_{Er}^2, interpolating for values of β that are not on the grid. We use a simulated investment series to calculate the autocorrelation of investment.

Finally, we find the set of the four parameters, γ_0, γ_1, σ_β^2 and ξ that make the simulated moments consistent with the empirical ones, i.e., (i) market excess return, (ii) market Sharpe ratio, (iii) cross-sectional dispersion in expected returns and (iv) the autocorrelation of investment. As shown in column (1) of the bottom panel of Table 4, the simulated moments are quite close to their empirical counterparts.

F   Extensions

Our baseline framework in Section 3 features (i) a single source of aggregate risk and (ii) a tight connection between financial market conditions and the "real" side of the economy – indeed, the state of technology determined both the common component of firm-level productivities and the price of risk simultaneously. In this appendix, we generalize that setup to allow for (i) multiple risk factors and (ii) more flexible formulations of the determinants of financial conditions. Although empirically disciplining the additional factors added here may be challenging, we demonstrate that the same insights from the baseline analysis go through. We also study versions of the model where heterogeneity in risk premia stem from "alphas" or "mis-pricing" in addition to betas and from differences across firms in exposure to capital price shocks, rather than productivity shocks. and show that our main results continue to hold.

F.1   Multifactor Model

There are J aggregate risk factors in the economy. Firms have heterogeneous loadings on these factors, so that the profit function (in logs) takes the form

$$\pi_{it} = \beta_i x_t + z_{it} + \theta k_{it} \tag{40}$$

70

where β_i is a vector of factor loadings of firm i, e.g., the j-th element of β_i is the loading of firm i profits on factor j, and x_t is the vector of factor realizations at time t, i.e.,

$$\beta_i = \begin{bmatrix} \beta_{1i} \\ \beta_{2i} \\ \vdots \\ \beta_{Ji} \end{bmatrix}' \qquad x_t = \begin{bmatrix} x_{1t} \\ x_{2t} \\ \vdots \\ x_{Jt} \end{bmatrix}'$$

Each factor, indexed by j, follows an AR(1) process

$$x_{jt+1} = \rho_j x_{jt} + \varepsilon_{jt+1}, \quad \varepsilon_{jt+1} \sim \mathcal{N}\left(0, \sigma_{\varepsilon_j}^2\right) \tag{41}$$

where the innovations are potentially correlated across factors. Denote by Σ_f the covariance matrix of factor innovations, i.e.,

$$\Sigma_f = \begin{bmatrix} \sigma_{\varepsilon_1}^2 & \sigma_{\varepsilon_1,\varepsilon_2} & \cdots & \sigma_{\varepsilon_1,\varepsilon_J} \\ \sigma_{\varepsilon_2,\varepsilon_1} & \sigma_{\varepsilon_2}^2 & \cdots & \sigma_{\varepsilon_2,\varepsilon_J} \\ \vdots & \vdots & & \vdots \\ \sigma_{\varepsilon_J,\varepsilon_1} & \sigma_{\varepsilon_J,\varepsilon_2} & \cdots & \sigma_{\varepsilon_J}^2 \end{bmatrix}$$

The idiosyncratic component of firm productivity follows

$$z_{it+1} = \rho_z z_{it} + \varepsilon_{it+1}, \quad \varepsilon_{it+1} \sim \mathcal{N}\left(0, \sigma_{\tilde{\varepsilon}}^2\right) \tag{42}$$

The stochastic discount factor takes the form

$$m_{t+1} = \log \rho - \gamma \varepsilon_{t+1} - \frac{1}{2}\gamma \Sigma_f \gamma' \tag{43}$$

where γ is a vector of factor exposures, e.g., element γ_j captures the exposure of the SDF to the j-th factor, and ε_{t+1} is the vector of innovations in each factor, i.e.,

$$\gamma = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_J \end{bmatrix}' \qquad \varepsilon_{t+1} = \begin{bmatrix} \varepsilon_{1t+1} \\ \varepsilon_{2t+1} \\ \vdots \\ \varepsilon_{Jt+1} \end{bmatrix}$$

For purposes of illustration, we assume γ is constant through time and there are no adjustment costs, although these assumptions are easily relaxed. Expressions (40), (41), (42) and (43) are

simple extensions of (11), (9) and (10).

Following a similar derivation as Appendix D.1, we can derive the realized mpk:

$$mpk_{it+1} = \alpha + \varepsilon_{it+1} + \beta_i \varepsilon_{t+1} + \beta_i \Sigma_f \gamma'$$

where β_i and ε_{t+1} denote vectors of factor loadings and shocks. The expected mpk and its cross-sectional dispersion are given by

$$\mathbb{E}_t[mpk_{it+1}] = \alpha + \beta_i \Sigma_f \gamma', \quad \sigma^2_{\mathbb{E}_t[mpk]} = \gamma \Sigma'_f \Sigma_\beta \Sigma_f \gamma'$$

where Σ_β is the covariance matrix of factor loadings across firms, i.e.,

$$\Sigma_\beta = \begin{bmatrix} \sigma^2_{\beta_1} & \sigma_{\beta_1,\beta_2} & \cdots & \sigma_{\beta_1,\beta_J} \\ \sigma_{\beta_2,\beta_1} & \sigma^2_{\beta_2} & \cdots & \sigma_{\beta_2,\beta_J} \\ \vdots & \vdots & & \vdots \\ \sigma_{\beta_J,\beta_1} & \sigma_{\beta_J,\beta_2} & \cdots & \sigma^2_{\beta_J} \end{bmatrix}$$

This is the natural analog of expression (16): (i) expected mpk is determined by the firm's exposure to (all) the aggregate risk factors in the economy and the risk prices of those factors, and (ii) mpk dispersion is a function of the dispersion in those exposures across firms as captured by Σ_β.

Similar steps as Appendix D.4 gives the following (approximate) expression for expected excess stock returns and the cross-sectional dispersion in expected returns:

$$Er^e_{it+1} = \beta_i \psi \Sigma_f \gamma', \quad \sigma^2_{Er^e_t} = \gamma \Sigma'_f \psi' \Sigma_\beta \psi \Sigma_f \gamma'$$

where ψ is a diagonal matrix with

$$\psi_{jj} = \frac{\frac{1}{\rho} + \delta - 1}{\frac{1}{\rho} + (1-\theta)\delta - 1} \frac{1-\rho}{1-\rho\rho_j}$$

where ρ_j denotes the persistence of factor j. These are the analogs of expressions (21) and (22) – expected returns depend on factor exposures and the risk prices of those factors. Expected return dispersion depends on the dispersion in those exposures, here captured by Σ_β.

Thus, the same insights from the single factor model go through – dispersion in $Empk$ and expected returns are both determined by variation in exposures to the set of aggregate factors and hence, there is a tight relationship between the two. To quantify the impact of these factors on mpk dispersion, however, we would need to know all the primitives governing the dynamics

of the factors, e.g., the vector of persistences ρ and the covariance matrix Σ_f, and exposures, i.e., the exposures of the SDF, γ, and the vectors of firm loadings, Σ_β. This would likely entail taking a stand on the nature of each factor, computing their properties from the data and calibrating/estimating the γ vector and the covariance matrix of firm exposures, Σ_β.

F.2 Financial Shocks

Our baseline model tightly linked financial conditions, for example, the price of risk, to macroeconomic conditions, i.e., the state of aggregate technology. However, financial conditions may not co-move one-for-one with the "real" business cycle. Here, we extend the setup to include pure financial shocks. The stochastic discount factor takes the form

$$
\begin{aligned}
m_{t+1} &= \log\rho - \gamma_t\varepsilon_{t+1} - \frac{1}{2}\gamma_t^2\sigma_\varepsilon^2 \\
\gamma_t &= \gamma_0 + \gamma_f f_t ,
\end{aligned}
\tag{44}
$$

where

$$
f_{t+1} = \rho_f f_t + \varepsilon_f, \quad \varepsilon_f \sim \mathcal{N}\left(0, \sigma_{\varepsilon_f}^2\right) .
$$

In this formulation, f_t denotes the time-varying state of financial conditions, which is now disconnected from the state of aggregate technology. These financial factors may be correlated with real conditions, x_t, but need not be perfectly so. Thus, there is scope for changes in financial conditions, independent of those in real conditions, to affect the price of risk and through this channel, the allocation of capital.[69] Note the difference between this setup and the one in Section F.1 – here, the financial factor, f_t, does not directly enter the profit function of the firm, it only affects the price of risk. Thus, it is a shock purely to financial market conditions. In contrast, the factors considered in Section F.1 directly affected firm profitability.

Keeping the remainder of the environment the same as Section 3, we can derive exactly the same expressions for expected mpk and its cross-sectional variance, i.e.,

$$
\mathbb{E}_t\left[mpk_{it+1}\right] = \alpha + \beta_i\gamma_t\sigma_\varepsilon^2, \quad \sigma_{\mathbb{E}_t[mpk_{it+1}]}^2 = \sigma_\beta^2\left(\gamma_t\sigma_\varepsilon^2\right)^2 ,
$$

where now γ_t is a function of financial market conditions. When credit market conditions tighten (i.e., when f_t is small/negative since $\gamma_f < 0$), γ_t is high and mpk dispersion will rise.

Just as in Section 3, the conditional expectation of one-period ahead TFP is given by

$$
\mathbb{E}_t\left[a_{t+1}\right] = \mathbb{E}_t\left[a_{t+1}^*\right] - \frac{1}{2}\frac{\theta_1\left(1-\theta_2\right)}{1-\theta_1+\theta_2}\sigma_\beta^2\left(\gamma_t\sigma_\varepsilon^2\right)^2
\tag{45}
$$

---

[69]Our baseline model is the nested case where $\gamma_f = \gamma_1$ and f_t and x_t are perfectly correlated.

which illustrates the effects of a deterioration in financial conditions on macroeconomic performance – when credit market conditions tighten and risk premia rise (i.e., f_t falls), the resulting increase in mpk dispersion leads to a fall in aggregate TFP.

Finally, the average long-run level of $Empk$ dispersion and aggregate TFP are given by

$$\mathbb{E}\left[\sigma^2_{Empk_t}\right] = \sigma^2_\beta \left(\gamma_0^2 + \gamma_f^2 \sigma_f^2\right)\left(\sigma^2_{\varepsilon_f}\right)^2, \quad \bar{a} = a^* - \frac{1}{2}\frac{\theta_1\left(1-\theta_2\right)}{1-\theta_1-\theta_2}\sigma^2_\beta \left(\gamma_0^2 + \gamma_f^2 \sigma_f^2\right)\left(\sigma^2_{\varepsilon_f}\right)^2,$$

where $\sigma_f^2 = \frac{\sigma^2_{\varepsilon_f}}{1-\rho_f^2}$. The expressions reveal a tight connection between financial conditions and long-run performance of the economy – higher financial volatility ($\sigma^2_{\varepsilon_f}$), even independent of the state of the macroeconomy, induces greater persistent MPK dispersion and depresses the average level of achieved productivity.

G   Other Distortions

With other distortions, the derivations are similar to those in Appendix D.1. The Euler equation is given by

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

Idiosyncratic distortions.   Substituting for m_{t+1} and τ_{it+1} and rearranging,

$$\begin{aligned}
\mathbb{E}_t\left[e^{m_{t+1}+\tau_{it+1}+z_{it+1}+\beta_i x_{t+1}}\right] &= \mathbb{E}_t\left[e^{\log\rho-\gamma_t\varepsilon_{t+1}-\frac{1}{2}\gamma_t^2\sigma_\varepsilon^2-\nu_1 z_{it+1}-\eta_{it+1}+z_{it+1}+\beta_i x_{t+1}}\right]\\
&= \mathbb{E}_t\left[e^{\log\rho+(1-\nu)\rho_z z_{it}+(1-\nu_1)\varepsilon_{it+1}+\beta_i\rho_x x_t+(\beta_i-\gamma_t)\varepsilon_{t+1}-\frac{1}{2}\gamma_t^2\sigma_\varepsilon^2-\eta_{it+1}}\right]\\
&= e^{\log\rho+(1-\nu_1)\rho_z z_{it}+\beta_i\rho_x x_t+\frac{1}{2}(1-\nu_1)^2\sigma_{\tilde{\varepsilon}}^2+\frac{1}{2}\beta_i^2\sigma_\varepsilon^2-\beta_i\gamma_t\sigma_\varepsilon^2-\eta_{it+1}}
\end{aligned}$$

so that

$$\theta GK_{it+1}^{\theta-1} = \frac{1-(1-\delta)\rho}{e^{\log\rho+(1-\nu_1)\rho_z z_{it}+\beta_i\rho_x x_t+\frac{1}{2}(1-\nu_1)^2\sigma_{\tilde{\varepsilon}}^2+\frac{1}{2}\beta_i^2\sigma_\varepsilon^2-\beta_i\gamma_t\sigma_\varepsilon^2-\eta_{it+1}}}$$

and rearranging and taking logs,

$$k_{it+1} = \frac{1}{1-\theta}\left(\tilde{\alpha}+\frac{1}{2}(1-\nu_1)^2\sigma_{\tilde{\varepsilon}}^2+\frac{1}{2}\beta_i^2\sigma_\varepsilon^2+(1-\nu_1)\rho_z z_{it}+\beta_i\rho_x x_t-\beta_i\gamma_t\sigma_\varepsilon^2-\eta_{it+1}\right)$$

where $\tilde{\alpha}$ and α are as defined in Appendix D.1.

The realized mpk is given by (ignoring the variance terms)

$$
\begin{aligned}
mpk_{it+1} &= \log\theta + \pi_{it+1} - k_{it+1} \\
&= \log\theta + \log G + z_{it+1} + \beta_i x_{t+1} - (1-\theta)\, k_{it+1} \\
&= \log\theta + \log G + z_{it+1} + \beta_i x_{t+1} - \tilde{\alpha} - (1-\nu_1)\,\rho_z z_{it} - \beta_i \rho_x x_t + \beta_i \gamma_t \sigma_\varepsilon^2 + \eta_{it+1} \\
&= \alpha + \varepsilon_{it+1} + \beta_i \varepsilon_{t+1} + \nu_1 \rho_z z_{it} + \beta_i \gamma_t \sigma_\varepsilon^2 + \eta_{it+1}
\end{aligned}
$$

The conditional expected mpk is

$$
\mathbb{E}_t\left[mpk_{it+1}\right] = \alpha + \nu_1 \rho_z z_{it} + \beta_i \gamma_t \sigma_\varepsilon^2 + \eta_{it+1}
$$

and the cross-sectional variance is

$$
\sigma^2_{\mathbb{E}_t[mpk_{it+1}]} = \left(\nu_1 \rho_z\right)^2 \sigma_z^2 + \sigma_\eta^2 + \left(\gamma_t \sigma_\varepsilon^2\right)^2 \sigma_\beta^2 \tag{46}
$$

Deriving stock returns follows closely the steps in Appendix D.4. Dividends are equal to

$$
D_{it+1} = e^{\tau_{it+1} + z_{it+1} + \beta_i x_{t+1}} K_{it+1}^\theta - K_{it+2} + (1-\delta)\, K_{it+1} - \frac{\xi}{2}\left(\frac{K_{it+2}}{K_{it+1}} - 1\right)^2 K_{it+1}
$$

and log-linearizing,

$$
d_{it+1} = \frac{\Pi}{D}\left(\tau_{it+1} + z_{it+1} + \beta_i x_{t+1}\right) + \left(\theta\frac{\Pi}{D} + (1-\delta)\frac{K}{D}\right) k_{it+1} - \frac{K}{D} k_{it+2} + \log D - \left(\theta\frac{\Pi}{D} - \delta\frac{K}{D}\right) k
$$

where $k = \log K$.

Substituting for k_{it+1} and k_{it+2} from above,

$$
d_{it+1} = A_0 + \tilde{A}_1 z_{it} + A_1 \beta_i x_t + \tilde{A}_2 \varepsilon_{it+1} + A_2 \beta_i \varepsilon_{t+1} + A_3 \eta_{it+1} + A_4 \eta_{it+2}
$$

where

$$A_0 = \log D - \left(\theta\frac{\Pi}{D} - \delta\frac{K}{D}\right)\left(k - \frac{\tilde{\alpha}}{1-\theta}\right)$$

$$A_1 = \frac{1}{1-\theta}\left(\frac{\Pi}{D} + (1-\delta-\rho_x)\frac{K}{D}\right)\rho_x - \frac{1}{1-\theta}\left(\theta\frac{\Pi}{D} + (1-\delta-\rho_x)\frac{K}{D}\right)\gamma_1\sigma_\varepsilon^2$$

$$\tilde{A}_1 = \frac{1-\nu_1}{1-\theta}\left(\frac{\Pi}{D} + (1-\delta-\rho_z)\frac{K}{D}\right)\rho_z$$

$$A_2 = \frac{\Pi}{D} - \frac{1}{1-\theta}\frac{K}{D}\rho_x + \frac{1}{1-\theta}\frac{K}{D}\gamma_1\sigma_\varepsilon^2$$

$$\tilde{A}_2 = \left(\frac{\Pi}{D} - \frac{1}{1-\theta}\frac{K}{D}\right)(1-\nu_1)\rho_z$$

$$A_3 = -\frac{1}{1-\theta}\left(\theta\frac{\Pi}{D} + (1-\delta)\frac{K}{D}\right)$$

$$A_4 = \frac{1}{1-\theta}\frac{K}{D}$$

Using the log-linearized return equation,

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

and conjecturing the stock price takes the form

$$p_{it} = c_{0i} + c_1\beta_i x_t + c_2 z_{it} + c_3\eta_{it+1}$$

gives

$$\begin{aligned}
r_{it+1} = {}& -\log\rho + (1-\rho)\left(\log\frac{P}{D} + A_0 - c_0\right) \\
& + \left((\rho\rho_z - 1)c_2 + (1-\rho)\tilde{A}_1\right)z_{it} \\
& + ((\rho\rho_x - 1)c_1 + (1-\rho)A_1)\beta_i x_t \\
& + \left(\rho c_2 + (1-\rho)\tilde{A}_2\right)\varepsilon_{it+1} + (\rho c_1 + (1-\rho)A_2)\beta_i\varepsilon_{t+1} \\
& + (\rho c_3 + (1-\rho)A_4)\eta_{it+2} + ((1-\rho)A_3 - c_3)\eta_{it+1}
\end{aligned}$$

The (log) excess return is the (negative of the) conditional covariance with the SDF:

$$\log\mathbb{E}_t\left[R_{it+1}^e\right] = (\rho c_1 + (1-\rho)A_2)\beta_i\gamma_t\sigma_\varepsilon^2$$

A_2 is independent of ν_1 and η. Following the same steps as in Appendix D.4, it is easily verified that c_1 is independent of these terms as well. Thus, expected returns are independent

of distortions.

Aggregate distortions. Consider the first formulation, i.e.,

$$\tau_{it+1} = -\nu_1 z_{it+1} - \nu_2 x_{t+1} - \eta_{it+1}$$

Similar steps as above give expression (46). Dispersion in expected stock market returns are similarly unaffected.

Next, consider the second formulation:

$$\tau_{it+1} = -\nu_1 z_{it+1} - \nu_3 \beta_i x_{t+1} - \eta_{it+1}$$

In this case, similar steps as above give the conditional expected mpk as

$$\mathbb{E}_t \left[mpk_{it+1} \right] = \alpha + \nu_1 \rho_z z_{it} + \nu_3 \beta_i \rho_x x_t + (1 - \nu_3) \beta_i \gamma_t \sigma_\varepsilon^2 + \eta_{it+1}$$

and expected excess stock market returns as

$$\log \mathbb{E}_t \left[R_{it+1}^e \right] = (1 - \nu_3) \psi \beta_i \gamma_t \sigma_\varepsilon^2$$

where ψ is as defined in expression (21). In other words, the risk-premium effect on expected mpk, as well as expected returns, are both scaled by a factor $1 - \nu_3$.

The mean level of expected mpk and return dispersion are, respectively,

$$
\begin{aligned}
\mathbb{E} \left[\sigma_{\mathbb{E}_t[mpk_{it+1}]}^2 \right] &= \sigma_\eta^2 + (\nu_1 \rho_z)^2 \sigma_z^2 + (\nu_3 \rho_x)^2 \sigma_x^2 \sigma_\beta^2 \\
&\quad + \left((1 - \nu_3) \sigma_\varepsilon^2 \right)^2 \left(\gamma_0^2 + \gamma_1^2 \sigma_x^2 \right) \sigma_\beta^2 + 2\nu_3 (1 - \nu_3) \rho_x \sigma_x^2 \gamma_1 \sigma_\varepsilon^2 \sigma_\beta^2 \\
\mathbb{E} \left[\sigma_{\log \mathbb{E}_t[R_{it+1}^e]}^2 \right] &= \left((1 - \nu_3) \psi \sigma_\varepsilon^2 \right)^2 \left(\gamma_0^2 + \gamma_1^2 \sigma_x^2 \right) \sigma_\beta^2
\end{aligned}
$$

The last two terms of the first equation capture the mpk effects of risk premia. The last term there is new and does not have a counterpart in the second equation – in other words, using dispersion in expected returns would give the second to last term, as usual, but not the last. If $\nu_3 < 0$, it is straightforward to verify that that term is positive (recall that γ_1 is negative). Then, we may be understating risk premium effects. If $\nu_3 > 0$, the last terms is negative and we may be overstating them.

In this latter case, we can obtain an upper bound on the extent of the potential bias as follows: holding the other parameters fixed, the term is most negative for $\nu_3 = 0.5$. Using this value, along with the estimated values of the other parameters, yields a value of the bias that

is at most about 0.03.

H   Robustness – Productivity Betas

In this appendix, we investigate the potential effects of (i) mis-measurement of firm-level capital and (ii) unobserved heterogeneity in θ on our estimates of productivity betas in Section 4.4. First, to see the effects of mis-measured capital or measurement error, assume that the measured capital stock is $\hat{k}_{it} = k_{it} + e_{it}$, where k_{it} is true capital and e_{it} the mis-measurement. Measured firm-level productivity growth is then equal to $\Delta z_{it} + \beta_i \Delta x_t - \theta \Delta e_{it}$. Regressing this on measures of aggregate productivity, i.e., Δx_t, it is straightforward to see that the estimated β's would be unaffected so long as changes in mis-measurement at the firm-level (Δe_{it}) are uncorrelated with the business cycle, which may be a reasonable conjecture. Put another way, mis-measured capital in this analysis leads to measurement errors in the dependent variable, which, under relatively mild conditions, are innocuous.[70]

How about unobserved heterogeneity in θ? It turns out this will have small effects as well. To see this, let θ_i denote the true firm-specific parameter and θ our assumed common value. Measured firm-level productivity growth is then equal to $\Delta z_{it} + \beta_i \Delta x - (\theta - \theta_i) \Delta k_{it}$ and regressing this on aggregate productivity growth gives $\beta_i - \frac{(\theta - \theta_i)\text{cov}(\Delta k_{it}, \Delta x_t)}{\text{var}(\Delta x_t)}$, where $\beta_i = \frac{\hat{\beta}_i - \theta_{2i}\omega}{1 - \theta_{2i}}$ is the effective true β (see Section 5 and Appendix I for a further discussion of this expression). The second term represents the potential bias, which depends on the covariance between investment and changes in aggregate productivity. How large is this covariance? As one example, consider the case with no adjustment costs and a constant price of risk. We can use the firm's optimality condition to analytically characterize the covariance, which gives the bias term to be $-\frac{(\theta - \theta_i)\text{cov}(\Delta k_{it}, \Delta x_t)}{\text{var}(\Delta x_t)} = \frac{1}{2}\frac{\theta - \theta_i}{1 - \theta_i}\beta_i \rho_x (1 - \rho_x)$. This term turns out to be negligible. Intuitively, because of time-to-build, investment in period $t + 1$ capital is chosen in period t, before the innovation in productivity is realized. Because of this, the covariance between changes in capital and contemporaneous productivity is quite small and is only non-zero due to mean reversion in the AR(1) process (indeed, if productivity follows a random walk or is iid, i.e., $\rho_x = 1$ or $\rho_x = 0$, the bias term is zero).[71] To verify this result quantitatively, we have simulated data under the extreme case where heterogeneity in θ_i is the only source of beta dispersion (we use the distribution of θ described in Appendix I). As described there, the true

---

[70]Moreover, a non-zero correlation between Δe_{it} and Δx_t is not itself sufficient to bias the estimates of beta dispersion. In the case of a non-zero correlation, the regression yields $\beta_i - \theta \frac{\text{cov}(\Delta e_{it}, \Delta x_t)}{\text{var}(\Delta x_t)}$. Thus, if the stochastic process on e_{it} is common across firms, this will add a constant bias to the beta estimates, but will not affect our estimates of dispersion.

[71]Industry-level heterogeneity is a special case where θ varies across industries but not across firms within an industry.

standard deviation of beta is 1.35; the biased estimate is 1.38. Although analytic expressions are not available in the full model with adjustment costs and time-varying risk, we have simulated this case as well – the biased estimate remains extremely close to the truth, 1.39. In sum, because the productivity betas are estimated off of covariances, they are extremely robust to concerns of both measurement of capital and unobserved parameter heterogeneity.

I The Sources of Betas

Heterogeneous technologies. With heterogeneity in input elasticities, the production function for firm i is

$$Y_{it} = X_t^{\hat{\beta}_i} \hat{Z}_{it} K_{it}^{\theta_{1i}} N_{it}^{\theta_{2i}} \tag{47}$$

In this case, we must make a distinction between mpk and the average product of capital, $apk = y_{it} - k_{it}$, which is the object we measure in the data. With common parameters, these are proportional. With parameter heterogeneity, they are not. Following similar steps as in the baseline analysis, we can derive

$$apk_{it+1} = -\log \theta_{1i} + \varepsilon_{it+1} + \beta_i \varepsilon_{t+1} + \beta_i \gamma_t \sigma_\varepsilon^2 + \text{constant} \tag{48}$$

where

$$\beta_i \equiv \frac{\hat{\beta}_i - \theta_{2i}\omega}{1 - \theta_{2i}}$$

In other words, an expression analogous to (14) holds, with two differences: first, variation in capital elasticities, θ_{1i}, will directly lead to apk dispersion through the first term in (48). Second, the effective beta is now a combination of the direct sensitivity to the aggregate shock, $\hat{\beta}_i$, and the firm-specific labor elasticity, θ_{2i}. Variation in θ_{2i} leads firms to have different exposures to changes in labor market conditions, captured through the cyclicality of wages, ω. To gain intuition, consider the extreme case where all heterogeneity in business cycle exposure comes through θ_{2i}, i.e., $\hat{\beta}_i = 1 \, \forall \, i$. Then, $\beta_i = \frac{1 - \theta_{2i}\omega}{1 - \theta_{2i}}$. It is straightforward to show that β_i is increasing in θ_{2i} as long as $\omega < 1$, i.e., holding all else equal, labor intensive firms are more exposed to cyclical movements in wages, which in and of itself leads to a higher risk premium.[72] Given this simple reinterpretation of beta, a version of the analysis in Section 3 continues to hold. In particular, we can derive an expression for expected stock markets returns that is analogous to (21), but which now also reflects the variation in θ_{2i} – in other words, this type of heterogeneity should be picked up by our empirical measure of variation in risk premia.

---

[72]Donangelo et al. (2018) explore a related mechanism and provide empirical support for the connection between "labor leverage" and risk premia. They also find that a necessary condition for this relationship to hold is that wages are less than perfectly procyclical, i.e., ω must be less than one.

How much of the observed beta dispersion can be attributed to variation in production function parameters? Although precisely pinning down its contribution is challenging, we can reach one (likely over-) estimate as follows. First, under the (admittedly strong) assumption that all cross-firm variation in labor's share of income comes from heterogeneity in θ_{2i}, we have $\frac{W_t N_{it}}{Y_{it}} = \theta_{2i}$. This is likely to be an upper bound, since there are many other reasons that labor's share may differ across firms (e.g., labor market frictions or distortions). Donangelo et al. (2018) (Table XII, Panel C) report a cross-sectional standard deviation of labor's share among Compustat firms of 0.186. Using this as an estimate of the dispersion in θ_{2i}, we can calculate the implied beta dispersion. Specifically, we assume that θ_{2i} is normally distributed and discretize the distribution on a seven point grid following the method suggested in Kennan (2006). This yields a range of values for θ_{2i} from 0.31 to 0.84 with standard deviation 0.183. Next, we compute the implied betas as $\beta_i = \frac{1-\theta_{2i}\omega}{1-\theta_{2i}}$. The standard deviation of the betas is 1.35, which represents about 12% of the overall standard deviation of betas in Section 4.[73]

Heterogeneous markups. As is well known in the literature, the production function in expression (8) with decreasing returns to scale is isomorphic to a revenue function that arises with monopolistically competitive firms that produce differentiated products and face constant elasticity demand functions. Specifically, assume that demand and production for firm i take the forms

$$Q_{it} = P_{it}^{-\mu_i}, \quad Y_{it} = X_t^{\tilde{\beta}_i} \tilde{Z}_{it} K_{it}^{\tilde{\theta}_1} N_{it}^{\tilde{\theta}_2}$$

where μ_i denotes the (potentially firm-specific) elasticity of demand and $\tilde{\theta}_j$, $j = 1, 2$ the techno-logical parameters in the production function, which for this section are assumed to be common across firms. It is straightforward to derive the following expression for firm revenues:

$$P_{it} Y_{it} = X_t^{\hat{\beta}_i} \hat{Z}_{it} K_{it}^{\theta_{1i}} N_{it}^{\theta_{2i}}$$

where $\hat{\beta}_i = \left(1 - \frac{1}{\mu_i}\right) \tilde{\beta}_i$, $\hat{Z}_{it} = \tilde{Z}_{it}^{1-\frac{1}{\mu_i}}$ and $\theta_{ji} = \left(1 - \frac{1}{\mu_i}\right) \tilde{\theta}_j$, $j = 1, 2$. With these re-interpretations of parameters, this is equivalent to (47) (there, the common price of output is equal to one). Note that for the case of a common demand elasticity, i.e., $\mu_i = \mu$, the analysis from Section 3 goes through exactly. With heterogeneity in demand elasticities, the analysis takes the same form as with technology heterogeneity – variation in technology and markups show up in the same way. Thus, markup dispersion across firms is an additional candidate for heterogeneous exposures and, indeed, should be picked up in our measures of firm-level risk

---

[73]David and Venkateswaran (2019) investigate technology heterogeneity in detail in a related framework and provide a sharper upper bound on the extent of this heterogeneity. We have also used their estimate for Compustat firms and found similar, though somewhat smaller, results.

premia. All else equal, firms facing a high demand elasticity (so setting a low markup, which is equal to $\frac{\mu_i}{\mu_i - 1}$) respond more strongly to shocks and so show greater sensitivity to them.

Even with no additional heterogeneity in $\tilde{\beta}_i$, the firm's beta in the revenue function is given by $\hat{\beta}_i = 1 - \frac{1}{\mu_i} = \frac{1}{markup_i}$, i.e., is the inverse of the markup. How much of the measured beta dispersion can variation in markups explain? Recent estimates of the within-industry standard deviation of (log) markups among Compustat firms yield values of about 0.20 (e.g., David and Venkateswaran (2019)).[74] Following a similar approach as in our analysis of technology heterogeneity, we can compute the resulting dispersion in betas. Specifically, we discretize the distribution of log markups on a five point grid. The lowest value on the grid implies a markup less than one, which we set to 1.01. We choose the standard deviation of the distribution so that the standard deviation of the truncated distribution is 0.20. This yields a range of markups from 1.01 to 1.63. After optimizing over labor, the implied beta for firm i is given by $\beta_i = \frac{\hat{\beta}_i - \theta_{2i}\omega}{1 - \theta_{2i}} = \frac{1 - \tilde{\theta}_2\omega}{markup_i - \tilde{\theta}_2}$. We set $\tilde{\theta}_2$ to a standard value of 0.67 and compute the standard deviation of these betas, which is 0.71. This accounts for about 6% of the overall standard deviation calculated in Section 4.

Other parameter heterogeneity. We have also examined the potential effects of two other forms of parameter heterogeneity – in the depreciation rate, δ, and the properties of idiosyncratic shocks, i.e., their persistence and volatility, ρ_z and $\sigma_{\tilde{\varepsilon}}^2$. To a first-order, the latter two parameters do not enter our estimates of beta dispersion/risk premia anywhere – idiosyncratic shocks, while extremely important in determining firm dynamics, do not affect covariances and so do not lead to risk premia. Expression (21) shows that δ does play a role in determining expected stock returns (through the denominator of ψ, which, with heterogeneity in δ will be firm-specific), but a numerical simulation suggests these effects are small. For example, allowing δ to range from 0.04 to 0.16 (so half and double the baseline value) generates a spread in expected returns of 1.6%, which is modest relative to the extent of expected return dispersion in the data. For example, Table 5 in Appendix B.2 shows that interquartile range of expected returns is almost 12%. Halving/doubling ρ_z and $\sigma_{\tilde{\varepsilon}}^2$ also leads to only limited spreads in expected returns (1.3% and 2.3%, respectively). These results suggest that unobserved heterogeneity in these parameters seems unlikely to account for the substantial dispersion in risk premia observed in the data. Moreover, note that our calculation of productivity betas in Section 4.4 is independent of these parameters, further emphasizing that the majority of the empirical beta dispersion is

---

[74]The statistics reported in Edmond et al. (2018) imply a roughly similar figure. Haltiwanger et al. (2018) find the same value using a different empirical method (namely, estimating a variable elasticity of substitution demand system using detailed data on prices and quantities) on a sample drawn from the Census of Manufactures.

unlikely to stem from these parameters.[75]

Heterogeneous demand sensitivities. SIC 5812 is defined as "Establishments primarily engaged in the retail sale of prepared food and drinks for on-premise or immediate consumption" and includes food service establishments ranging from fast food (e.g., McDonalds) to high-end restaurants (e.g., Ruth's Chris Steak House). We gathered data (where available) on average check per person (usually proxied by total check divided by the number of entrees ordered) from publicly available sources, including company SEC filings and surveys performed by Citi Research and Morgan Stanley. The data are generally from 2014 to 2015. Matching these prices to the Compustat data yielded a sample of 20 publicly traded firms in SIC 5812 with data on prices, betas, expected returns and MPK.

We first extracted the set of all firms in SIC 5812 for which we have sufficient quarterly observations to compute our measures of risk exposure (20 consecutive quarters are required). Next we obtained data on average check per person. These data are primarily from surveys performed by Citi Research and Morgan Stanley, downloaded from `https://finance.yahoo.com/news/much-costs-eat-every-major-201809513.html`, dated September 2015. Of the firms in the Compustat sample, this gave us pricing data for 8 firms: McDonalds (MCD), Wendy's (WEN), Sonic (SON), Chipotle (CHP), Cheesecake Factory (CHE), Texas Roadhouse (TEX), BJ's Restaurants (BJR) and Red Robin (ROB). We supplemented these data with figures reported in company 10-K filings with the SEC for the year 2014 for Jack in the Box (JCK), Panera Bread (PAN), Carrol's Restaurant Group (the largest Burger King franchisee; BKG), Chili's (CHL), Cracker Barrel (CRA), Bob Evans (BOB), Ruth's Chris Steakhouse (RUT), Denny's (DEN), Famous Dave's (FAM), Kona Grill (KON), Granite City (GRA) and Darden (DAR). Data on Granite City are from its 2013 10-K filing, where we calculated the average of the reported range across markets. Darden owns Eddie V's, Capital Grille, Seasons's 52, Bahama Breeze, Olive Garden, Longhorn Steakhouse, Fleming's, Bonefish Grill, Carraba's and Outback Steakhouse. It reports an average check for each of these chains separately, which we combined into a single value using a sales-weighted average. The largest among this group is Olive Garden. We excluded chains that were confined to a very limited geographic area and those for which we could not obtain average check data. In total, our sample consists of 20

---

[75]We have also explored the effects of adjustment costs alone by simulating a panel of firms with a common beta and computing the mean of period-by-period expected return dispersion. We find that adjustment costs on their own lead to very little dispersion in expected returns (average standard deviation of about 0.015 compared to 0.127 in the data), suggesting that it is unlikely that our estimates of beta are reflecting the effects of these costs. We have also verified that this result goes through for larger levels of these costs (e.g., $\xi = 3$, compared to 0.04 in the baseline). Note that this is in line with our approximate expression for expected returns in equation (21): that expression show that to a first-order, expected returns are completely independent of adjustment costs.

firms. We computed average betas, expected returns and MPK for these firms over the period 2010-2015.

Figure 2 illustrates the main results from this exercise. The top two panels of the figure plot average check against CAPM and demand betas, along with the lines of best fit. Both plots show a strong positive relationship – higher quality restaurants, as proxied by price, have higher exposure to aggregate shocks, measured using either stock market or operating data. Firms on the low end include McDonalds (MCD), Burger King (BUR), Wendy's (WEN), Sonic (SON), etc., and towards the higher end Kona Grill (KON), Famous Dave's (FAM) and Cheesecake Factory (CHE). The highest-price restaurant in the sample is Ruth's Chris Steak House (RUT).[76] The bottom two panels of the figure go one step further and additionally link quality to measures of expected returns and MPK. Again, there is a strong positive relationship: higher quality restaurants – which the top panel shows tend to be those with higher exposure to aggregate shocks – have higher expected returns and MPK.

Table 10 presents the full set of correlations across average check, betas, expected returns and MPK for this set of firms (we also add a measure of beta constructed from the Fama-French factors, which gives similar results). The table shows strong positive correlations between average check and the various beta measures, as well as between average check and returns and MPK. Further, the positive correlations between beta, expected returns and MPK show that high beta and high expected return firms tend to have MPK. However, Figure 2 neatly summarizes the key message – differences in the responsiveness of firm-level demand to aggregate conditions due to quality variation and "trading down" seems a promising explanation of beta dispersion.

Table 10: Correlations – SIC 5812, Eating Places

	Ln(avg. check)	CAPM Beta	Demand Beta	FF Beta	Expected Return	Ln(MPK)
Ln(avg. check)	1.00					
CAPM Beta	0.47	1.00				
Demand Beta	0.64	0.41	1.00			
FF Beta	0.61	0.85	0.47	1.00		
Expected Return	0.63	0.42	0.37	0.57	1.00	
Ln(MPK)	0.65	0.65	0.63	0.77	0.64	1.00

---

[76]Ruth's Chris is somewhat of an outlier with a price of $76.00 per person, about three times larger than the next highest. We have verified that omitting Ruth's Chris does not significantly change the results.
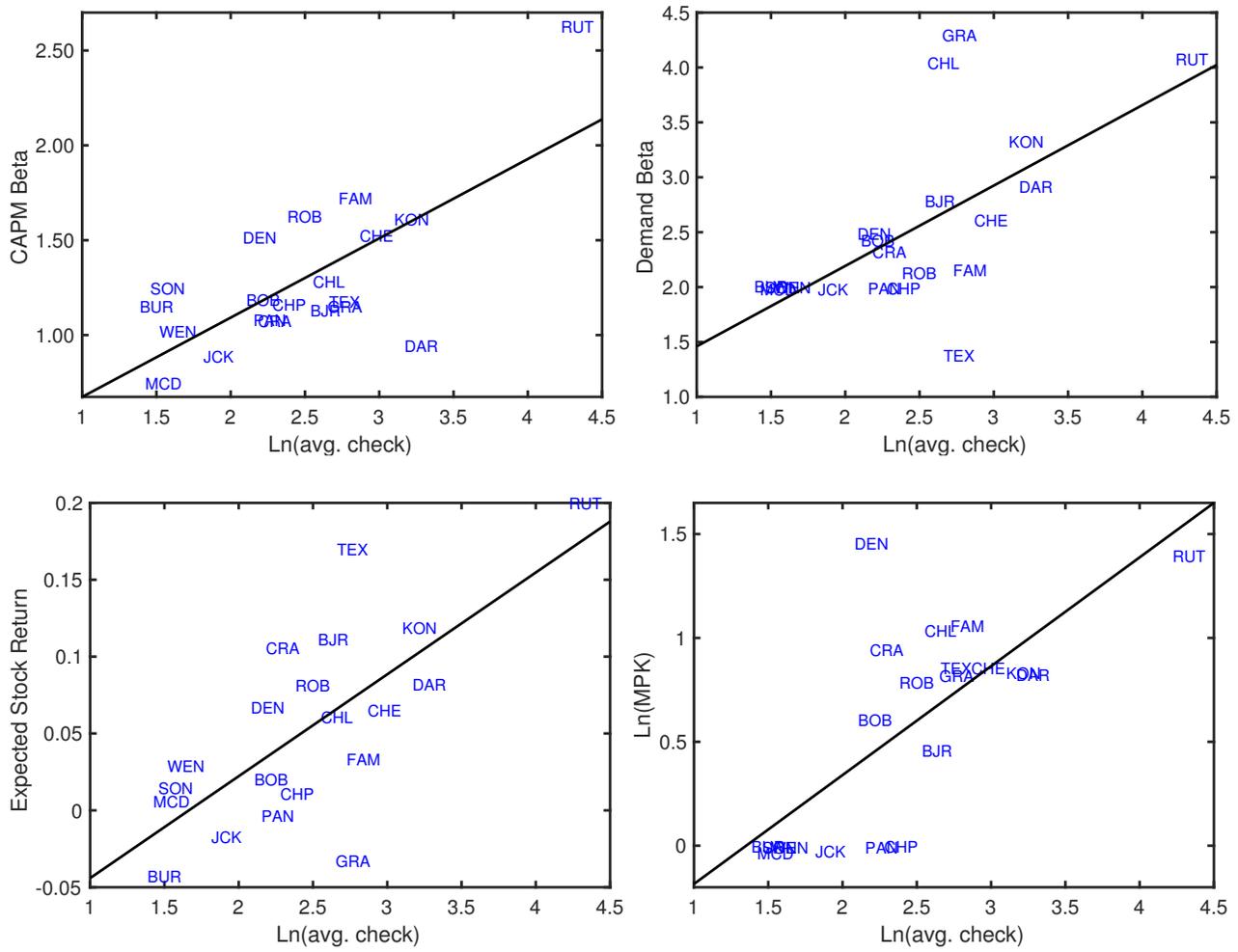
Figure 2: Average Check, Beta, Expected Returns and MPK in SIC 5812, Eating Places