



2013

Factor-Loading Uncertainty and Expected Returns

Christopher S. Armstrong
University of Pennsylvania

Snehal Banerjee

Carlos Corona

Follow this and additional works at: https://repository.upenn.edu/accounting_papers



Part of the [Accounting Commons](#), [Corporate Finance Commons](#), and the [Finance and Financial Management Commons](#)

Recommended Citation

Armstrong, C. S., Banerjee, S., & Corona, C. (2013). Factor-Loading Uncertainty and Expected Returns. *The Review of Financial Studies*, 26 (1), 158-207. <http://dx.doi.org/10.1093/rfs/hhs102>

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/accounting_papers/91
For more information, please contact repository@pobox.upenn.edu.

Factor-Loading Uncertainty and Expected Returns

Abstract

Firm-specific information can affect expected returns if it affects investor uncertainty about risk-factor loadings. We show that a stock's expected return is decreasing in factor-loading uncertainty, controlling for the average level of its factor loading. When loadings are persistent, learning by investors can induce time-series variation in price-dividend ratios, expected returns, and idiosyncratic volatility, even when the aggregate risk-premium is constant and fundamental shocks are homoscedastic. Consistent with our predictions, we estimate that average annual returns of a firm with the median level of factor-loading uncertainty are 400 to 525 basis points lower than a comparable firm without factor-loading uncertainty.

Disciplines

Accounting | Corporate Finance | Finance and Financial Management

Factor-Loading Uncertainty and Expected Returns

Chris Armstrong

Wharton School, University of Pennsylvania

Snehal Banerjee

Kellogg School of Management, Northwestern University

Carlos Corona

Tepper School of Business, Carnegie Mellon University

Abstract

Firm-specific information can affect expected returns if it affects investor uncertainty about risk-factor loadings. We show that a stock's expected return is decreasing in factor-loading uncertainty, controlling for the average level of its factor loading. When loadings are persistent, learning by investors can induce time-series variation in price-dividend ratios, expected returns, and idiosyncratic volatility, even when the aggregate risk-premium is constant and fundamental shocks are homoscedastic. Consistent with our predictions, we estimate that average annual returns of a firm with the median level of factor-loading uncertainty are 400 to 525 basis points lower than a comparable firm without factor-loading uncertainty. (JEL: G12, G14)

We have benefited from comments from Pietro Veronesi (the editor) and the anonymous referee. We would also like to thank Andres Almazan, Ryan Ball (discussant), Jules van Binsbergen, Anna Cieslak, John Core, Alexander David, Mike Faulkender, Mike Fishman, Lorenzo Garlappi, Jeremy Graveline, Yaniv Konchitchki, Bob Korajczyk, Arvind Krishnamurthy, Praveen Kumar (discussant), Christian Leuz, Dmitry Livdan (discussant), Dimitris Papanikolaou, Jonathan Parker, Lubos Pástor, Mitchell Petersen, Tano Santos, Ernst Schaumburg, Paul Tetlock, Viktor Todorov, Laura Veldkamp, X. Frank Zhang, and seminar participants at the University of Toronto, Loyola University, the University of Illinois at Chicago, Columbia University, the 2009 Financial Economics and Accounting Conference, the 2010 Annual AFA Meetings, the 2010 Tel Aviv University Finance Conference, and the 2011 SFS Cavalcade Conference. Previous versions of this paper were circulated under the titles "Information Quality, Systematic Risk and the Cost of Capital," and "Information Quality and the Cross-Section of Expected Returns." Send correspondence to Snehal Banerjee, Kellogg School of Management, Northwestern University, 2001 Sheridan Rd., Evanston, IL 60201; telephone: 847-491-8335; fax: 847-491-5719. E-mail: snehal-banerjee@kellogg.northwestern.edu.

Does publicly available information about a firm affect its expected return? The standard theoretical approach in the literature considers the relation between information and expected returns in the context of a single firm.¹ While these models generally predict a negative relation between the quality of public information and expected returns, the empirical evidence on this predicted negative relation has been mixed.² Moreover, it is not clear whether the intuition from these single-firm models extends to a large economy with multiple firms, since firm-specific information may be diversifiable (e.g., Hughes, Liu, and Liu 2007).

Our insight is to recognize that firm-specific information can affect expected returns if it is about the firm's systematic risk-factor loadings. In a benchmark model, we show that a firm's expected returns are decreasing in the uncertainty about the firm's risk-factor loadings, even after controlling for the average level of the factor loading. The basic intuition for our main result is easily described using a simple example.³ Consider a stock that pays annual dividends, with current dividend $D = \$1$, and an annual dividend growth rate $g = 5\%$. Further, suppose the CAPM holds, the risk-free rate is 3%, and the market risk premium is 5%. If investors know, with certainty, that the beta of the stock is 1, then the CAPM implies an expected return of 8% (i.e., $3\% + 1(5\%) = 8\%$), and the Gordon growth formula implies a price of $P = \frac{D(1+g)}{r-g} = \frac{1.05}{8\%-5\%} = \35 . If, instead, investors are uncertain about the stock's beta, then this uncertainty affects both its price and expected return. For instance, suppose investors believe that the beta of the stock is either 0.5 or 1.5, each with probability 50%, and they learn its value at the end of the year. Conditional on each possible value of beta, the price is given by the standard Gordon growth formula as follows:

$$P_{\beta=0.5} = \frac{1.05}{[3\%+0.5(5\%)]-5\%} = \$210, \text{ and } P_{\beta=1.5} = \frac{1.05}{[3\%+1.5(5\%)]-5\%} = \$19.09,$$

and the price of the stock is given by the expectation over the possible values of beta:

$$P = 0.5 \times P_{\beta=0.5} + 0.5 \times P_{\beta=1.5} = \$114.545.$$

Importantly, since the stock price is a convex function of its beta, uncertainty about beta leads to a higher price than would be implied by the average beta (i.e., $P = \$35$). Moreover, the expected return on the stock over the next year is given by

$$\mathbb{E}[r] = \frac{0.5 [P_{\beta=0.5} + D] (1 + g) + 0.5 [P_{\beta=1.5} + D] (1 + g)}{P} - 1 = 5.92\%.$$

As such, when investors are uncertain about the firm's risk-factor loading (or beta), its expected return is lower than what is predicted by its average factor loading.

¹Examples include Diamond and Verrecchia (1991), Baiman and Verrecchia (1996), and Easley and O'Hara (2004). Lambert, Leuz, and Verrecchia (2007) extend this approach to a multiple firm setting.

²On the one hand, a number of papers, including Easley and O'Hara (2004), Francis, LaFond, Olsson, and Schipper (2004), Francis, LaFond, Olsson, and Schipper (2005), Barth, Konchitchki, and Landsman (2011), and Francis, Nanda, and Olsson (2008), document that proxies of higher information quality or increased transparency are associated with lower expected returns. On the other hand, papers including Botosan (1997), Botosan and Plumlee (2002), Core, Guay, and Verdi (2008), Duarte and Young (2009), and Ben-Rephael, Kadan, and Wohl (2010) find either limited or no evidence of a relation between information quality or disclosure and cost of capital.

³We thank the editor for suggesting the example.

We develop a dynamic, partial equilibrium model with time-varying factor loadings and learning by investors that not only captures the intuition from this simple example but also allows us to analyze the effects of investor uncertainty and learning about factor loadings on prices and returns more generally. As has been suggested in the literature, estimating factor loadings and discount rates is difficult for investors, especially given that they vary over time. For instance, Fama and French (1997) point out that “there is strong variation through time in the CAPM and three-factor risk loadings of industries. . . . And industries give an understated picture of the problems that will arise in estimating risk loadings for individual firms and investment projects.” Our model suggests that factor-loading uncertainty, and learning about these factors, can have economically important implications for the pricing of returns, both in the cross-section and over time.

We begin with a benchmark model with i.i.d. shocks to highlight the economic mechanism through which factor-loading uncertainty affects expected returns in a standard asset pricing framework. A firm’s risk-factor loading determines the covariance between its cash flows and the pricing kernel, and the firm’s expected return decreases in this covariance.⁴ We depart from the standard setup by assuming that investors face uncertainty about firms’ future risk-factor loadings, and we show that a firm’s expected return decreases in this factor-loading uncertainty. This result follows from the fact that the covariance between a firm’s cash flows and the pricing kernel, and hence the price of these cash flows, is a convex function of the firm’s future risk-factor loading. Intuitively, in bad states of the world (i.e., when the pricing kernel is high), a decrease in the firm’s factor loading, or systematic risk, leads to a bigger change in the present value of its cash flows than an equivalent increase in its factor loading. However, in good states of the world, the opposite is true—the effect of an increase in the factor loading on the present value is larger than the effect of an equivalent decrease. As a result, the overall effect of increasing uncertainty about the risk-factor loading of next period’s cash flow is to increase the present value of this cash flow, or equivalently, reduce its discount rate (or expected return).

We then generalize the benchmark model to allow for persistence in factor loadings, and model various ways in which investors can learn about these loadings over time. We show that the effect of factor-loading uncertainty on expected returns can be significantly amplified when factor loadings are persistent. Since price-dividend ratios depend on investors’ conditional beliefs about factor loadings, learning about these loadings over time generates time-series variation in price-dividend ratios (and returns), even though we assume that the aggregate risk premium is constant. We also show that when the quality of information available to investors changes over time, this naturally generates time-series variation in idiosyncratic return volatility. Finally, we consider an extension of the model in which investors condition on realized dividends and returns to update their beliefs about firms’ factor loadings. In this case, even though all fundamental shocks are homoscedastic, the information structure and learning by investors generates *stochastic* idiosyncratic return volatility, since investors’ conditional beliefs about a firm’s factor loadings depend on realizations of the aggregate risk factor, and these beliefs endogenously determine the firm’s return dynamics.

⁴Cash flows in bad states of the world (i.e., when the pricing kernel, or marginal utility of the representative investor, is high) are discounted less—hence, expected returns are decreasing in the covariance between cash flows and the pricing kernel.

The primary contribution of our paper is to provide a benchmark model in which to analyze the effect of investor uncertainty and learning about firm-specific risk-factor loadings on the pricing of returns. Specifically, we show that uncertainty about factor loadings can generate additional cross-sectional variation in expected returns that is not captured by the level of the factor loadings. Since our model predicts a negative relation between firm-specific uncertainty and expected returns, in contrast to the single-firm models discussed earlier, our results may help reconcile the apparently conflicting empirical evidence documented in the existing literature. Our model also provides a rational mechanism that can generate the seemingly “puzzling” negative relation between idiosyncratic volatility and expected returns documented by Ang, Hodrick, Xing, and Zhang (2006) and others, since, in our model, factor-loading uncertainty is negatively related to expected returns but positively related to idiosyncratic volatility.⁵

More generally, we show that learning about risk-factor loadings can generate time-series variation in firms’ price-dividend ratios, expected returns, and idiosyncratic return volatility, even when the aggregate risk premium is constant. While time-series variation in aggregate risk premia obviously has important economic implications, by shutting down this channel, our analysis highlights the mechanism through which learning firm-specific information can affect the dynamics of prices and returns. In fact, we show that when investors condition on realized dividends and returns, learning about risk-factor loadings naturally leads to stochastic variation in idiosyncratic return volatility, even when the fundamental shocks are all assumed to be homoscedastic. Moreover, each firm’s stochastic volatility is affected by shocks to the aggregate risk factor (or returns to the factor-mimicking portfolio), and therefore the cross-section of idiosyncratic volatility exhibits a factor structure that is similar to what has been empirically documented in the literature (e.g., Campbell, Lettau, Malkiel, and Xu 2001). As such, our model may serve as a useful benchmark in which to study how learning can endogenously give rise to stochastic volatility at the firm level.

We document evidence that is consistent with our main prediction, using Fama-MacBeth regressions on the cross-section of CRSP stock returns. In the first stage, we estimate sixty month rolling window firm-specific regressions to estimate firm-specific factor loadings on the (log) excess market return, and use the standard error of the factor-loading estimate as a firm-specific proxy for factor-loading uncertainty. In the second-stage cross-sectional regressions, we show that uncertainty about market-factor loadings is negatively related to expected returns. We also find that this relation is statistically significant and is robust to including factor loadings on the SMB, HML, and UMD portfolios and aggregate volatility risk, and a number of firm characteristics (including size, market-to-book ratio, turnover, debt-to-equity ratio, volatility of earnings, return on assets, accruals, and bid-ask spreads). Moreover, the effect appears to be economically significant. In our sample of firms from January 1964 to December 2011, our estimates imply that a firm with the median level of uncertainty about the market-factor loading should have average annualized returns that are 400 to 525 basis points lower than an otherwise comparable firm with the same average factor loading but no uncertainty.

The next section discusses the marginal contribution of our paper in relation to the existing

⁵As Ang, Hodrick, Xing, and Zhang (2006) discuss, most models that incorporate market frictions or behavioral biases imply that idiosyncratic volatility should be positively related to expected returns. As such, they state that their “results on idiosyncratic volatility represent a substantive puzzle.”

literature. Section 2 presents the benchmark model, which we use to illustrate the intuition behind our main result. Section 3 presents the general model, which incorporates persistence in factor loadings and learning by investors, and characterizes the economic magnitudes of the effects we describe. Section 4 reports the results from our empirical analysis of the model’s main prediction and summarizes additional implications, and Section 5 concludes. All proofs are in Appendix A.

1. Related Literature

Our paper relates most closely to Pástor and Veronesi (2003) and Xia (2001), but our results differ significantly from theirs. Pástor and Veronesi (2003) show that since a firm’s stock price is convex in its cash-flow growth, higher uncertainty about profitability leads to higher market-to-book ratios. Like Pástor and Veronesi (2003), we find that firm-specific uncertainty leads to higher valuations, and when investors learn about this source of firm-specific uncertainty, price-dividend ratios and idiosyncratic volatility exhibit time-series variation. However, Pástor and Veronesi (2003) and Johnson (2004) argue that for unlevered firms, learning about firm-specific profitability has no effect on expected returns.⁶ In contrast to these earlier papers, we show that firm-specific uncertainty can have an effect on expected returns even in the absence of leverage. As we discuss further in Section 2, this is because the firm-specific uncertainty in our model is about the covariance between cash-flow growth and the pricing kernel, and, importantly, because this firm-specific uncertainty is resolved *simultaneously* with aggregate uncertainty.

As in Xia (2001), investors in our model learn about the predictability of expected returns, and this leads to stochastic variance of the factor loading. However, in contrast to Xia (2001), investors’ learning affects the pricing of returns in our model, and hence the stochastic conditional variance that results from learning leads to stochastic volatility in returns. While the effect of learning in generating stochastic return volatility has been explored at the aggregate level (e.g., David 1997; Veronesi 1999), to our knowledge, our model is the first to highlight a natural link between learning about factor loadings from past dividends / returns and stochastic volatility.

Our paper is also related to the large literature on parameter uncertainty and estimation risk. Kumar, Sorescu, Boehme, and Danielsen (2008) extend this literature by considering an economy in which investors are uncertain about the quality of information they can learn from. They endogenously derive a conditional CAPM in which the market risk premium, market volatility, and systematic risk of stocks depend on the information available to investors. Adrian and Franzoni (2009) extend the conditional CAPM by introducing unobservable long-run changes in factor loadings. They show that when investors learn about these long-run changes over time, the model generates low-frequency variation in betas that can help explain

⁶Pástor and Veronesi (2003) show that there is no effect of uncertainty about profitability on expected returns in the case of no dividends, and a small, local effect in the presence of dividends. Johnson (2004) shows that in levered firms, an increase in idiosyncratic volatility (which increases total volatility but keeps the risk premium constant) decreases the expected return on their levered equity.

the value and size premia in the cross-section of stock returns. Our analysis, which we view as complementary to these studies, takes the pricing kernel as given and focuses on the effect of firm-specific uncertainty about factor loadings. As a result, although firm-specific information has an effect on expected returns, it is not through an effect on the market risk premium (in contrast to Kumar et. al. 2008). Moreover, by explicitly modeling dividends in a standard asset pricing setup, we are able to derive implications of uncertainty and learning for not only expected returns but also other firm-specific attributes such as price-dividend ratios and return volatility.

2. Benchmark Model

2.1 Model setup

In this section, we consider a simple benchmark model in which dividend growth is assumed to be independent and identically distributed over time. Although this assumption is made primarily for tractability, it also makes the effect of uncertainty about risk-factor loadings on expected returns transparent. In the next section, we generalize the model to allow for persistence in factor loadings and learning by investors.

The evolution of the pricing kernel is given by:

$$M_{t+1} = M_t \exp \left\{ -r_f - \frac{1}{2}V_m - m_{t+1} \right\} \text{ where } m_{t+1} \sim \mathcal{N}(0, V_m), \quad (1)$$

and the m_{t+1} are i.i.d., normally distributed variables with mean zero and variance V_m . The aggregate source of risk, or risk factor, in the economy is therefore driven by the random variable m_{t+1} . This implies that the unconditional (log) risk-free rate is given by r_f , since

$$-\log \left(E \left[\frac{M_{t+1}}{M_t} \right] \right) = r_f + \frac{1}{2}V_m - \frac{1}{2}V_m = r_f. \quad (2)$$

The existence of a pricing kernel relies only on the assumption that there is no arbitrage in the economy, which makes the current setup quite general. In particular, this representation can capture a variety of pricing models, including consumption-based models and factor-based models, such as the CAPM.⁷

⁷For instance, in an endowment economy in which aggregate consumption growth follows a log-normal process, i.e., $\Delta c_{t+1} \equiv \log(C_{t+1}/C_t) \sim N(\mu, \sigma_c^2)$, and the representative investor has power utility with a coefficient of relative risk-aversion γ and discount factor e^δ , the pricing kernel (or stochastic discount factor) is given by the marginal utility of the representative investor, $\frac{M_{t+1}}{M_t} = e^\delta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} = \exp\{\delta - \gamma \Delta c_{t+1}\}$, and the risk-free rate is given by: $R_f = \exp\{r_f\} \equiv 1/E \left[\frac{M_{t+1}}{M_t} \right] = \exp\{-\delta + \gamma\mu - \frac{1}{2}\gamma^2\sigma_c^2\}$, which, in turn, implies that

$$M_{t+1} = M_t \exp \left\{ -r_f - \frac{1}{2}\gamma^2\sigma_c^2 - \gamma(\Delta c_{t+1} - \mu) \right\} \equiv M_t \exp \left\{ -r_f - \frac{1}{2}V_m - m_{t+1} \right\},$$

where $m_{t+1} = \gamma(\Delta c_{t+1} - \mu) \sim N(0, \gamma^2\sigma_c^2)$ and $V_m = \gamma^2\sigma_c^2$.

At the beginning of each period, investors' beliefs about firm i 's dividends (or cash flows), at date $t + 1$ are given by:

$$D_{i,t+1} = D_{i,t} \exp \left\{ \bar{d}_i + \beta_{i,t+1} m_{t+1} + d_{i,t+1} - \frac{1}{2} (\beta_{i,t+1}^2 V_m + V_{d,i}) \right\}, \text{ where } d_{i,t+1} \sim \mathcal{N}(0, V_{d,i}) \quad (3)$$

and $\beta_{i,t+1}$ is assumed to have an i.i.d. distribution given by

$$\beta_{i,t+1} \sim \mathcal{N}(b_i, V_{\beta,i}), \quad (4)$$

and m_{t+1} , $d_{i,t+1}$, and $\beta_{i,t+1}$ are independent. The specification in equation (3) implies that dividend growth consists of a constant component (i.e., \bar{d}_i), a systematic component (i.e., $\beta_{i,t+1} m_{t+1}$), and an idiosyncratic component (i.e., $d_{i,t+1}$). The systematic component of dividend growth depends on the firm's *stochastic* factor loading, $\beta_{i,t+1}$.⁸ Unlike standard models, we relax the assumption that investors know (with certainty) the future factor loading that drives dividend growth at time $t + 1$. We first assume that factor loadings are i.i.d. over time to highlight the intuition for our main result, but we relax this assumption in the next section.

Our specification also ensures that (log) expected dividend growth is given by \bar{d}_i . In the standard case, when there is no uncertainty about $\beta_{i,t+1} = b_i$, one can adjust the dividend growth process by a convexity adjustment term $\frac{1}{2} b_i^2 V_m + \frac{1}{2} V_{d,i}$ to ensure that (log) expected dividend growth is given by \bar{d}_i , i.e.,

$$D_{i,t+1} = D_{i,t} \exp \left\{ \bar{d}_i + b_i m_{t+1} + d_{i,t+1} - \frac{1}{2} (b_i^2 V_m + V_{d,i}) \right\} \Rightarrow \mathbb{E} \left[\frac{D_{i,t+1}}{D_{i,t}} \right] = \exp \{ \bar{d}_i \} \quad (5)$$

In our setup, the convexity adjustment naturally generalizes to $\frac{1}{2} \beta_{i,t+1}^2 V_m + \frac{1}{2} V_{d,i}$, which implies that the log expected dividend growth is given by \bar{d}_i , since

$$\mathbb{E} \left[\frac{D_{i,t+1}}{D_{i,t}} \right] = \mathbb{E} \left[\mathbb{E} \left[\exp \left\{ \bar{d}_i + \beta_{i,t+1} m_{t+1} + d_{i,t+1} - \frac{1}{2} (\beta_{i,t+1}^2 V_m + V_{d,i}) \right\} \middle| \beta_{i,t+1} \right] \right] = \exp \{ \bar{d}_i \}. \quad (6)$$

This convexity adjustment ensures that the effect of uncertainty about $\beta_{i,t+1}$ that we derive is not simply a consequence of the difference between log expected returns and expected log returns.⁹

2.2 The pricing equation and expected returns

By the definition of the pricing kernel, the price of firm i 's stream of dividends is given by:

$$P_{i,t} = \mathbb{E}_t \left[\sum_{s=1}^{\infty} \frac{M_{t+s}}{M_t} D_{i,t+s} \right], \quad (7)$$

⁸This is an example of what is known as a random coefficients model (e.g., Cooley and Prescott 1976). Ang and Chen (2007) use a similar setup to model the uncertainty faced by an econometrician when estimating time-varying betas, and use this to argue that there is little evidence of a value premium in the long run.

⁹One can alternatively define the convexity adjustment as $\log \mathbb{E} \left[\frac{D_{i,t+1}}{D_{i,t}} \right] - \mathbb{E} \left[\log \left(\frac{D_{i,t+1}}{D_{i,t}} \right) \right]$. As we show in Appendix B, while this alternative specification implies a similar (but larger) effect of uncertainty about $\beta_{i,t+1}$ on the expected return, it leads to less-transparent and less-intuitive expressions for expected returns.

and the *unconditional* expected return on the firm's dividend stream is then given by:

$$\mathbb{E} [R_{i,t+1}] = \mathbb{E} \left[\frac{P_{i,t+1} + D_{i,t+1}}{P_{i,t}} \right]. \quad (8)$$

Under an appropriate transversality condition, we have our benchmark result.

Proposition 1. *Suppose the sequence of prices P_t satisfies the following transversality condition:*

$$\lim_{T \rightarrow \infty} \mathbb{E}_t \left[\frac{M_{t+T}}{M_t} P_{t+T} \right] = 0. \quad (9)$$

Then, given that factor loadings are *i.i.d.*, we have the following results:

(a) firm i 's price-dividend ratio is constant, and given by $PD_{i,t} \equiv P_{i,t}/D_{i,t} = \frac{p_i}{1-p_i}$, where

$$p_i = \mathbb{E} \left[\frac{M_{t+1}}{M_t} \frac{D_{i,t+1}}{D_{i,t}} \right] = \exp \left\{ \bar{d}_i - r_f - b_i V_m + \frac{1}{2} V_{\beta,i} V_m^2 \right\}, \quad (10)$$

(b) firm i 's return process is of the form:

$$R_{i,t+1} = \exp \left\{ \mu_i + \beta_{i,t+1} m_{t+1} + d_{i,t+1} - \frac{1}{2} (\beta_{i,t+1}^2 V_m + V_{d,i}) \right\}, \quad (11)$$

(c) firm i 's expected return is given by:

$$\mathbb{E} [R_{i,t+1}] = \exp \{ \mu_i \} = \exp \left\{ r_f + b_i V_m - \frac{1}{2} V_{\beta,i} V_m^2 \right\}. \quad (12)$$

The result implies that conditional on the (mean) level of the factor loading (i.e., b_i), higher uncertainty about the factor loading (i.e., $V_{\beta,i}$) decreases firm i 's expected return. This result follows from the effect of uncertainty about $\beta_{i,t+1}$ on the covariance between dividend growth and the pricing kernel. To understand this effect, we derive the expression for the covariance explicitly:

$$\begin{aligned} \text{cov} \left(\frac{M_{t+1}}{M_t}, \frac{D_{i,t+1}}{D_{i,t}} \right) &= \mathbb{E} \left[\text{cov} \left(\frac{M_{t+1}}{M_t}, \frac{D_{i,t+1}}{D_{i,t}} \mid \beta_{i,t+1} \right) \right] + \text{cov} \left(\mathbb{E} \left[\frac{M_{t+1}}{M_t} \mid \beta_{i,t+1} \right], \mathbb{E} \left[\frac{D_{i,t+1}}{D_{i,t}} \mid \beta_{i,t+1} \right] \right) \\ &= \mathbb{E} \left[\text{cov} \left(\frac{M_{t+1}}{M_t}, \frac{D_{i,t+1}}{D_{i,t}} \mid \beta_{i,t+1} \right) \right] \end{aligned} \quad (13)$$

$$= \mathbb{E} \left[\exp \{ \bar{d}_i - r_f \} (\exp \{ -\beta_{i,t+1} V_m \} - 1) \right] \quad (14)$$

$$= \exp \{ \bar{d}_i - r_f \} (\exp \{ -b_i V_m + \frac{1}{2} V_m^2 V_{\beta,i} \} - 1). \quad (15)$$

Since the expected value of the pricing kernel and dividend growth conditional on a realization of $\beta_{i,t+1}$ are uncorrelated, the unconditional covariance between the pricing kernel and dividend growth is just the expectation of the covariance conditional on a realization of $\beta_{i,t+1}$ (i.e., it is given by (13)). Moreover, since conditional on a specific realization of $\beta_{i,t+1}$, $\frac{M_{t+1}}{M_t}$ and $\frac{D_{i,t+1}}{D_{i,t}}$ are lognormal random variables, the covariance between them is a convex function of $\beta_{i,t+1}$ (see the expression in equation (14)). As a result, the unconditional covariance between the pricing kernel and dividend growth is increasing in uncertainty about $\beta_{i,t+1}$ through a Jensen's inequality effect.

The convexity of the conditional covariance (given by equation (14)) in $\beta_{i,t+1}$ follows from our assumption that both the pricing kernel and dividend growth are log-linear (convex) functions of the common shock m_{t+1} . This is a common assumption in much of the asset pricing literature, which relies on the exponential-affine structure of the setup. Intuitively, the log-linearity implies that the pricing kernel $\frac{M_{t+1}}{M_t}$ is steeper in bad states of the world (i.e., low m_{t+1} , or high $\frac{M_{t+1}}{M_t}$) than in good states of the world (i.e., high m_{t+1} , or low $\frac{M_{t+1}}{M_t}$). In a representative agent model, this can be interpreted as the convexity of the representative agent's marginal utility—the agent's marginal utility is steeper in bad states of the world (i.e., low aggregate consumption) than in good states of the world.

The convexity of the pricing kernel implies that the present value of a dividend $D_{i,t+1}$ in the next period, given by $\frac{M_{t+1}}{M_t} D_{i,t+1}$, is convex in the factor loading $\beta_{i,t+1}$. As a result, a mean-preserving spread in $\beta_{i,t+1}$ increases the present value of the dividend. Intuitively, in bad states of the world, the effect of a unit decrease in $\beta_{i,t+1}$ (i.e., a decrease in systematic risk) leads to a bigger increase in the present value of the dividend than the decrease in its present value from a unit increase in $\beta_{i,t+1}$. In good states of the world, the opposite is true i.e., the effect of an increase in $\beta_{i,t+1}$ is bigger. As a result, a mean-preserving spread in $\beta_{i,t+1}$ increases the present value of the dividend in both good and bad states of the world. Figure 1 illustrates these effects for a specific set of parameters. The bottom panel plots the present value of the dividend across states of the world given a mean-preserving spread in $\beta_{i,t+1}$ (of ± 0.1 around one). For bad states of the world (i.e., $m_{t+1} \leq 0$), the effect of a decrease in $\beta_{i,t+1}$ on the present value is larger, while for good states of the world, the effect of an increase in $\beta_{i,t+1}$ is larger. Consequently, the effect of a mean-preserving spread is to increase the present value of the dividend (i.e., the average of the dashed and dotted lines lies above the solid line), or, equivalently, decrease the expected return.

As we show in the proof of Proposition 1, since the pricing kernel and cash-flow growth are conditionally independent over time, the price-dividend ratio (denoted by $PD_{i,t}$) is constant. Moreover, the effect of uncertainty on the present value of the dividend is also reflected in the expression for the price-dividend ratio. In particular, a firm's price-dividend ratio $PD_{i,t}$ is increasing in the price-dividend ratio of a dividend strip, p_i , which itself is increasing in uncertainty about $\beta_{i,t+1}$ (see equation (10)). This effect of uncertainty on valuations is analogous to the effect of uncertainty about profitability on market-to-book ratios in Pástor and Veronesi (2003).

Since the price-dividend ratio is constant, the Euler equation $\mathbb{E} \left[\frac{M_{t+1}}{M_t} R_{i,t+1} \right] = 1$ implies that the expected excess return on security i decreases with the covariance between its dividend growth and the pricing kernel, which, in turn, implies that it is decreasing in the uncertainty about its factor loading, since

$$\mathbb{E} [R_{i,t+1}] - R_f = -R_f \text{cov} \left(\frac{M_{t+1}}{M_t}, R_{i,t+1} \right) = -R_f \frac{1+PD_{i,t}}{PD_{i,t}} \text{cov} \left(\frac{M_{t+1}}{M_t}, \frac{D_{i,t+1}}{D_{i,t}} \right) \quad (16)$$

$$= -R_f \frac{1}{p_i} \text{cov} \left(\frac{M_{t+1}}{M_t}, \frac{D_{i,t+1}}{D_{i,t}} \right) = R_f \left(\exp \left\{ b_i V_m - \frac{1}{2} V_{\beta,i} V_m^2 \right\} - 1 \right). \quad (17)$$

As we show in the next section, if the factor loading is persistent, the price-dividend ratio varies over time, and this amplifies the effect of factor-loading uncertainty on expected returns.

However, by initially shutting down this additional effect, we are able to characterize the effect of factor-loading uncertainty on expected returns more transparently.

Expression (11) for returns in Proposition 1 explicitly accounts for the appropriate convexity adjustments (through the $\frac{1}{2}(\beta_{i,t+1}^2 V_m + V_{d,i})$ term), and as such, our main result is not simply an artifact of the difference between log expected returns and expected log returns.¹⁰ Finally, since there is only one source of aggregate uncertainty, the risk-factor loading that drives firm i 's cash-flow growth also drives the covariance of the firm's return with the pricing kernel. Accordingly, if there is a tradable factor-mimicking portfolio, then our model implies a log-linear, uncertainty-adjusted CAPM that is described in the following corollary.

Corollary 1. *Suppose $R_{M,t+1}$ is the return on the factor-mimicking portfolio—i.e.,*

$$R_{M,t+1} = \exp \left\{ \mu_M + m_{t+1} - \frac{1}{2} V_m \right\}, \quad (18)$$

where $\mu_M = \log(\mathbb{E}(R_{M,t+1}))$ denotes the log expected return. Then the log expected return $\mu_i = \log(\mathbb{E}(R_{i,t+1}))$ on firm i 's stock with cash-flow loading $\beta_{i,t+1} \sim N(b_i, V_{\beta,i})$ is characterized by the following expression:

$$\mu_i - r_f = b_i(\mu_M - r_f) - \frac{1}{2} V_{\beta,i} V_m^2. \quad (19)$$

Moreover, conditional on $\beta_{i,t+1}$, the covariance between $r_{i,t+1} = \log(R_{i,t+1})$ and $r_{M,t+1} = \log(R_{M,t+1})$ is given by $\beta_{i,t+1} V_m$.

The above result immediately follows from the definition of the factor-mimicking portfolio $R_{M,t+1}$ and Proposition 1. The corollary highlights that, in our model, risk-factor loadings in cash-flow growth are identical to risk-factor loadings in log returns. The corollary also highlights that while firm-specific uncertainty about factor loadings (i.e., $V_{\beta,i}$) affects expected returns, uncertainty about the idiosyncratic component of dividend growth (i.e., $V_{d,i}$) does not. Finally, in the special case of our model where investors face no uncertainty about risk-factor loadings (i.e., $V_{\beta,i} = 0$), we recover the standard log-linear CAPM relation, which is given by

$$\mu_i - r_f = b_i(\mu_M - r_f). \quad (20)$$

2.3 Discussion

Although our model is very similar to those in Pástor and Veronesi (2003) and Johnson (2004), in contrast to those studies, we find that firm-specific parameter uncertainty has an effect on expected returns even in the absence of leverage.¹¹ As noted in these earlier papers, whether parameter uncertainty affects expected returns depends not only on what parameters investors are uncertain about, but also *when* the parameter uncertainty is resolved. In our model,

¹⁰Specifically, the object of interest is the log expected return, μ_i , which is characterized in equation (12).

¹¹As mentioned earlier, Pástor and Veronesi (2003) show that there is no effect of uncertainty about profitability on expected returns in the case of no dividends, and a small, local effect in the presence of dividends. However, the mechanism through which firm-specific uncertainty affects expected returns in our model is different and, as we discuss in the next section, is potentially substantial in magnitude.

uncertainty about $\beta_{i,t+1}$ affects expected returns because (i) $\beta_{i,t+1}$ affects prices and returns *nonlinearly*, and (ii) the risk generated by the parameter uncertainty is *resolved simultaneously* with aggregate uncertainty (i.e., uncertainty about $\beta_{i,t+1}$ and m_{t+1} are resolved simultaneously). While the log-normal framework we use is tractable and helps us relate our results to the literature, these observations suggest that uncertainty about factor loadings should have an effect on expected returns in other settings if conditions (i) and (ii) are satisfied.¹² In a related point, note that, in our model, investors face uncertainty about the covariance between cash-flow growth and the pricing kernel, and the nonlinearity this introduces is important for our results. As such, the effect of uncertainty on expected returns would be different in a model where investors face uncertainty about the *level* of cash flows or prices, instead of the risk-factor loadings.

It is also important to note that investors do not “eventually learn” the factor loading of a firm’s dividends in our model. Given the distribution of $\beta_{i,t+1}$ in (4), investors always face residual uncertainty about the future factor loadings of a firm’s dividends. Moreover, since factor loadings are stochastic, regressing *past* dividend growth on *past* realizations of the aggregate risk factor only provides investors with an estimate of *past* factor loadings, but does not perfectly reveal *future* factor loadings, which are relevant for setting current prices. As we show in the next section, this is true even when factor loadings are persistent.

3. Learning About Persistent Factor Loadings

In this section, we generalize the model by allowing for persistence in firm’s risk-factor loadings. We show that the effect of factor-loading uncertainty on expected returns is generally amplified when $\beta_{i,t}$ are persistent. We also show that since price-dividend ratios depend on investors’ conditional beliefs, learning about factor loadings generates time-series variation in price-dividend ratios and, as a result, is an additional source of variation in returns.

We maintain most of the assumptions of the benchmark model. In particular, investors’ beliefs about the pricing kernel and firm i ’s dividend process are given by (1) and (3), respectively. However, we generalize the process for the risk-factor loading to allow for persistence by assuming that $\beta_{i,t+1}$ evolves according to the following process:

$$\beta_{i,t+1} = (1 - a_i) b_i + a_i \beta_{i,t} + \phi_i y_t + \nu_{i,t+1}, \text{ where } \nu_{i,t+1} \sim N(0, V_{\beta,i}), y_t \sim N(0, V_y) \quad (21)$$

and $\nu_{i,t+1}$ and y_t are independent of each other and independent of m_{t+1} and $d_{i,t+1}$. We assume that the persistence, a_i , the unconditional mean, b_i , the fundamental variance of factor loadings, $V_{\beta,i}$, the conditioning variable y_t , and the corresponding weight, ϕ_i , are known to all investors at date t , but that the realization of $\beta_{i,t}$ is not. Without loss of generality, we assume that the conditioning variable y_t is mean zero, so that the unconditional mean of $\beta_{i,t+1}$ is b_i . We restrict attention to the simplest specification that allows for persistence in factor loadings to maintain tractability and to highlight the effect of uncertainty and learning in a transparent manner.

¹²For instance, our example in the introduction illustrates the effect of factor-loading uncertainty on expected returns in a setting that does not assume log-normality.

The specification in equation (21) is a special case of Adrian and Franzoni (2009) applied to our log-linear setting, and, as discussed in their paper, nests a number of conditional CAPM models as special cases (including Jagannathan and Wang 1996; Lettau and Ludvigson 2001; Jostova and Philipov 2005; Ang and Chen 2007).

It is important to note that we maintain the assumption that V_m is constant over time, which implies that the aggregate risk premium (or, market price of risk) for the economy remains constant. While we believe that time-series variation in the aggregate risk premium is empirically and theoretically important, we make this assumption to highlight the effects of learning about factor loadings in our setting. In particular, any time-series variation in expected returns and price-dividend ratios will not be driven by time-series variation in the market price of risk, but by the learning process of investors.

In the following subsections, we consider different models of learning. Subsection 3.1 describes the general version of the model in which investors learn about the current realization of $\beta_{i,t}$ using exogenous signals. In Subsection 3.2, we consider two special cases to refine the intuition from the general model, and to gauge the economic magnitude of the effects we characterize. In Subsection 3.3, we present an extension of the general model in which investors condition on realized dividends and returns to update their beliefs about $\beta_{i,t}$. In this case, we show that learning can induce stochastic volatility in the return process, even if all fundamental shocks are assumed to be homoscedastic.

3.1 Learning from signals

Although we explicitly model how investors can learn about factor loadings from realized dividends and returns in a later extension, we begin with a stylized information structure that provides a useful benchmark. We assume that investors do not observe the factor loadings themselves, but that they instead observe a signal $Y_{i,t+1}$ at date $t + 1$, given by

$$Y_{i,t+1} = \beta_{i,t+1} + \varepsilon_{i,t+1}, \text{ where } \varepsilon_{i,t+1} \sim N(0, V_{\varepsilon,i,t}), \quad (22)$$

and $\varepsilon_{i,t+1}$ is uncorrelated with m_{t+1} , y_{t+1} , $\beta_{i,t+1}$, and $d_{i,t+1}$. Although we allow $\varepsilon_{i,t+1}$ to have time-varying volatility, we assume that conditional on date t information, the volatility of $\varepsilon_{i,t+1}$ is known to investors. This assumption allows us to parsimoniously model the possibility of time-series variation in information quality, without introducing too much additional notation or complexity. One can interpret the information structure in equation (22) as a reduced-form representation of investors learning about factor loadings from a variety of information sources that we do not model explicitly, including fundamental analysis, earnings announcements and guidance, and firm-level disclosures (e.g., see Beaver, Kettler, and Scholes 1970; Bowman 1979). The stylized nature of (22) allows us to characterize the first-order effects of learning in a tractable manner, and we leave more nuanced modeling of richer information structures for future work.

Denote investors' distribution of $\beta_{i,t+1}$ conditional on date t information by

$$\beta_{i,t+1} | \mathcal{I}_t \sim N(b_{i,t}, \Sigma_{i,t}) \quad (23)$$

Given the evolution of factor loadings and the information structure above, an immediate application of the Kalman filter gives us the following result.

Lemma 1. *The expectation of the risk-factor loading at date $t + 2$, conditional on the information available to investors at date $t + 1$, is given by:*

$$b_{i,t+1} \equiv \mathbb{E}_{t+1} [\beta_{i,t+2}] = (1 - a_i) b_i + a_i (b_{i,t} + \lambda_{i,t} (Y_{i,t+1} - b_{i,t})) + \phi_i y_{t+1} \quad (24)$$

$$= (1 - a_i) b_i + a_i (b_{i,t} + \lambda_{i,t} (\beta_{i,t+1} + \varepsilon_{i,t+1} - b_{i,t})) + \phi_i y_{t+1} \quad (25)$$

where $\lambda_{i,t} = \frac{\Sigma_{i,t}}{\Sigma_{i,t} + V_{\varepsilon,i,t}}$. The variance of the risk-factor loading at date $t + 2$, conditional on date $t + 1$ information, is given by:

$$\Sigma_{i,t+1} \equiv \text{var}_{t+1} (\beta_{i,t+2}) = V_{\beta,i} + a_i^2 \Sigma_{i,t} (1 - \lambda_{i,t}). \quad (26)$$

Moreover, if the quality of information available to investors is constant (i.e., $V_{\varepsilon,i,t} = V_{\varepsilon,i}$), then the steady-state conditional variance is given by:

$$\Sigma_i = \frac{1}{1 - a_i^2} V_{\beta,i} = \frac{1}{2} \left(V_{\beta,i} - V_{\varepsilon,i} (1 - a_i^2) + \sqrt{4V_{\beta,i}V_{\varepsilon,i} + (V_{\beta,i} - V_{\varepsilon,i} (1 - a_i^2))^2} \right), \quad (27)$$

where $\lambda_i = \frac{\Sigma_i}{\Sigma_i + V_{\varepsilon,i}}$.

The Kalman gain parameter, $\lambda_{i,t}$, is a convenient measure of the quality of information available to investors at date t , since it is decreasing in the noise of the signal, $V_{\varepsilon,i,t}$, and is bounded between zero and one. The gain $\lambda_{i,t}$ is zero when the noise in the signal $Y_{i,t}$ is infinite (i.e., $V_{\varepsilon,i,t} \rightarrow \infty$) and investors learn nothing about the factor loading from their signal. Conversely, the gain is one if the noise in the signal is zero, and investors perfectly learn the current realization of $\beta_{i,t}$ from their signal. As expected, the conditional variance of factor loadings at date t is decreasing in the quality of information available to investors at date t (i.e., $\lambda_{i,t}$).

If the quality of information available to investors is constant (i.e., $V_{\varepsilon,i,t} = V_{\varepsilon,i}$), then we can characterize the steady-state value of the conditional variance in closed form (see equation (27)). Straightforward calculations also imply the following intuitive results:

$$\lim_{\lambda_i \rightarrow 1} \Sigma_i = V_{\beta,i}, \quad \lim_{\lambda_i \rightarrow 0} \Sigma_i = \frac{1}{1 - a_i^2} V_{\beta,i}, \quad \frac{\partial}{\partial V_{\beta,i}} \Sigma_i > 0, \quad \text{and} \quad \frac{\partial}{\partial V_{\varepsilon,i}} \Sigma_i > 0. \quad (28)$$

If the quality of information available to investors is perfect (i.e., $\lambda_i = 1$, or $V_{\varepsilon,i} = 0$), then investors observe the realization of the current factor loading perfectly, and the conditional variance of the next period's factor loading is given by $V_{\beta,i}$. As such, even if investors perfectly observe current factor loadings, they still face uncertainty about future loadings that are relevant for pricing returns. At the other extreme, if the signal $Y_{i,t}$ is infinitely noisy (i.e., $V_{\varepsilon,i} \rightarrow \infty$, or $\lambda_i \rightarrow 0$), then investors have no information about the current factor loading, and the conditional variance of the next period's factor loading is given by the unconditional variance $\frac{1}{1 - a_i^2} V_{\beta,i}$. This quantity characterizes the upper bound on the uncertainty that investors face about factor loadings.

Given the evolution of beliefs described in Lemma 1, we can calculate firm i 's price-dividend ratio and returns by log-linearizing the price-dividend ratio around its long-run mean (e.g., Campbell and Shiller 1988). Under the appropriate transversality condition, we have our main result.

Proposition 2. *Suppose the sequence of prices P_t satisfies the following transversality condition:*

$$\lim_{T \rightarrow \infty} \mathbb{E}_t \left[\frac{M_{t+T}}{M_t} P_{t+T} \right] = 0. \quad (29)$$

Then, given that factor loadings evolve according to (21), and investors' beliefs evolve according to the results of Lemma 1, a log-linearization of the price-dividend ratio around its long-run mean implies the following:

(a) *firm i 's log price-dividend ratio is given by:*

$$pd_{i,t} = \log(PD_{i,t}) = A_{i,t} - \frac{1}{1-\rho_i} b_{i,t} V_m, \quad (30)$$

(b) *firm i 's return process is of the form:*

$$R_{i,t+1} = \exp \left\{ \alpha_{i,t} + \beta_{i,t+1} m_{t+1} + d_{i,t+1} - \frac{1}{2} (\beta_{i,t+1}^2 V_m + V_{d,i}) + \gamma_i b_{i,t+1} \right\}, \quad (31)$$

(c) *firm i 's expected return, conditional on date t information, is given by:*

$$\mathbb{E}_t [R_{i,t+1}] = \exp \left\{ r_f + b_{i,t} V_m - \frac{1}{2} \left(1 + 2\lambda_{i,t} \frac{a_i \rho_i}{1-a_i \rho_i} \right) \Sigma_{i,t} V_m^2 \right\}, \quad (32)$$

where $A_{i,t}$, $\alpha_{i,t}$, and γ_i are known to investors at date t , and defined in Appendix A, and $\rho_i \equiv \frac{\exp\{\bar{pd}\}}{1+\exp\{\bar{pd}\}} \in (0, 1)$, where \bar{pd} is the long-run mean of $pd_{i,t}$.

The above proposition presents our main results, and generalizes the benchmark results of Proposition 1. As before, the price-dividend ratio is decreasing in the conditional mean $b_{i,t}$, and increasing in the conditional variance of the factor loading (i.e., $\Sigma_{i,t}$), as can be seen from expression (A10) for $A_{i,t}$ in Appendix A. As in the i.i.d. case, since the systematic shocks to returns are driven by $\beta_{i,t+1} m_{t+1}$, the covariance between firm i 's return and the factor-mimicking portfolio is still driven by the factor loading—i.e., the covariance between $r_{i,t+1} = \log(R_{i,t+1})$ and $r_{M,t+1} = \log(R_{M,t+1})$ (as defined by (18)) is given by $\beta_{i,t+1} V_m$. Finally, as in Proposition 1, firm i 's expected return is decreasing in investors' (conditional) uncertainty about its factor loadings (i.e., $\Sigma_{i,t}$), even after controlling for the conditional mean level of its factor loading (i.e., $b_{i,t}$).

However, allowing for persistence in firms' factor loadings generates a number of additional effects that are not present in the benchmark case of Proposition 1. First, a firm's price-dividend ratio is no longer constant, but instead depends on investors' conditional beliefs about its factor loadings. This implies that a firm's price-dividend ratio varies over time, even when all fundamental shocks have constant variance by assumption, and in particular, the aggregate risk premium (i.e., V_m) is constant. Moreover, this time-series variation in price-dividend ratios

leads to an additional source of variation in returns. In particular, a firm's returns depend not only on shocks to its dividend growth (as captured by $\beta_{i,t+1}m_{t+1} + d_{i,t+1}$), but also on variation in investors' beliefs about its factor loading (as captured by $b_{i,t+1}$).

The effect of investors' uncertainty about factor loadings (i.e., $V_{\beta,i}$) on expected returns is increasing in the persistence of factor loadings (i.e., a_i). This amplification effect manifests through two complementary channels. To illustrate the intuition for this effect, consider the following expression for firm i 's expected excess return:¹³

$$\mathbb{E}_t [R_{i,t+1}] - R_f = -R_f \text{cov}_t \left(\frac{M_{t+1}}{M_t}, R_{i,t+1} \right) = -R_f \text{cov}_t \left(\frac{M_{t+1}}{M_t}, \frac{1+PD_{i,t+1}}{PD_{i,t}} \frac{D_{i,t+1}}{D_{i,t}} \right) \quad (33)$$

$$= -R_f \mathbb{E}_t \left[\text{cov}_t \left(\frac{M_{t+1}}{M_t}, \frac{D_{i,t+1}}{D_{i,t}} \middle| \beta_{i,t+1} \right) \frac{1+PD_{i,t+1}}{PD_{i,t}} \right] \quad (34)$$

$$= -R_f \mathbb{E}_t \left[\exp \{ \bar{d}_i - r_f \} (\exp \{ -\beta_{i,t+1} V_m \} - 1) \frac{1+PD_{i,t+1}}{PD_{i,t}} \right] \quad (35)$$

The first channel through which persistence affects expected returns is through its effect on the conditional covariance between the pricing kernel and dividend growth. As in the i.i.d. case, the covariance between the pricing kernel and dividend growth, conditional on a realization of $\beta_{i,t+1}$, is a convex function of $\beta_{i,t+1}$ (see equation (35)). Moreover, as Lemma 1 implies, the conditional variance of factor loadings (i.e., $\Sigma_{i,t}$) ranges from a lower bound of $V_{\beta,i}$ to an upper bound of $\frac{1}{1-a_i^2} V_{\beta,i}$, and is increasing in $V_{\beta,i}$. Hence, even if price-dividend ratios were constant (as in the i.i.d. case), the effect of the same level of fundamental uncertainty (i.e., $V_{\beta,i}$) on the conditional covariance is at least as large as in the i.i.d. case, but is likely to be larger, when factor loadings are persistent (i.e., $a_i > 0$).

Second, a larger variance of shocks to factor loadings (i.e., a higher $V_{\beta,i}$) leads to a higher volatility of conditional expectations (i.e., higher $\text{var}_t(b_{i,t+1})$), which translates into a larger multiplicative effect on the expected excess return through the $\frac{1+PD_{i,t+1}}{PD_{i,t}}$ term, since $PD_{i,t+1}$ is convex in the conditional expectation $b_{i,t+1}$. Intuitively, since factor loadings are persistent, price-dividend ratios are correlated over time. This implies that, when investors learn about the current factor loading, $\beta_{i,t}$, they update their beliefs about the covariance between dividends and the pricing kernel not only for the next period, but for all future periods. The effect of this amplification on the current price-dividend ratio is captured by the factor $\frac{1}{1-\rho_i a_i}$, which scales $b_{i,t} V_m$ in equation (30). Similarly, the amplification effect on expected returns is captured by the $2\lambda_{i,t} \frac{a_i \rho_i}{1-a_i \rho_i}$ term in equation (32).

Proposition 2 also highlights that in a dynamic setting, an increase in fundamental uncertainty about factor loadings (i.e., higher $V_{\beta,i}$) and a decrease in information quality about factor loadings (i.e., lower λ_i , or higher $V_{\varepsilon,i}$) can have different effects. In static settings, as intuition suggests, an increase in $V_{\beta,i}$ has a similar effect as a decrease in λ_i , since both lead to an increase in conditional uncertainty (i.e., Σ_i). However, in a dynamic setting, information quality affects

¹³Mathematically, note that since both the price-dividend ratio (i.e., $PD_{i,t+1}$) and the covariance between the pricing kernel and dividend growth, conditional on $\beta_{i,t+1}$ (i.e., $\text{cov}_t \left(\frac{M_{t+1}}{M_t}, \frac{D_{i,t+1}}{D_{i,t}} \middle| \beta_{i,t+1} \right)$) are log-convex in $\beta_{i,t+1}$, so is their product (see equation (34)). This, in turn, implies that the product $\text{cov}_t \left(\frac{M_{t+1}}{M_t}, \frac{D_{i,t+1}}{D_{i,t}} \middle| \beta_{i,t+1} \right) \frac{1+PD_{i,t+1}}{PD_{i,t}}$ is convex in $\beta_{i,t+1}$, and so the expected excess return is decreasing in uncertainty about $\beta_{i,t+1}$.

future uncertainty in two offsetting ways. To see these channels explicitly in our model, note that Proposition 2 implies that the effect of uncertainty on expected returns in the steady state is given by the following term:

$$-\frac{1}{2} \left(1 + 2\lambda_i \frac{a_i \rho_i}{1 - a_i \rho_i} \right) \Sigma_i V_m^2, \text{ where } \Sigma_i = \frac{1}{1 - a_i^2 (1 - \lambda_i)} V_{\beta, i}. \quad (36)$$

This expression shows that information quality has two offsetting effects on expected returns when factor loadings are persistent. First, as in a static model, higher information quality (i.e., higher λ_i) decreases the conditional uncertainty about future factor loadings (i.e., lowers Σ_i). However, higher information quality also increases the sensitivity of expected returns to uncertainty about factor loadings through the $2\lambda_i \frac{a_i \rho_i}{1 - a_i \rho_i}$ term. This is because, in a dynamic setting, higher information quality also implies that investors put more weight on their signals when updating their beliefs, which, in turn, implies that conditional expectations are more volatile in the future (i.e., higher variance of $b_{i,t}$). The overall effect of information quality (i.e., λ_i) on expected returns depends on both of these effects. In particular, by differentiating expression (36) with respect to λ_i , one can show that the relation between expected returns and information quality after controlling for the level of factor-loading uncertainty (i.e., $V_{\beta, i}$) is positive when factor-loading persistence is high enough—i.e., when

$$a_i - 2\rho_i + a_i^2 \rho_i > 0, \text{ or equivalently, } a_i > \frac{1}{2\rho_i} \left(\sqrt{1 + 8\rho_i^2} - 1 \right), \quad (37)$$

but negative when factor-loading persistence is low (i.e., when $a_i < \frac{1}{2\rho_i} \left(\sqrt{1 + 8\rho_i^2} - 1 \right)$).¹⁴

The subtle difference between the notions of uncertainty and information quality in a dynamic setting may have important empirical implications. For instance, our model predicts that the cross-sectional effect of an increase in fundamental uncertainty about factor loadings (i.e., $V_{\beta, i}$) is to decrease expected returns. However, after controlling for the level of fundamental uncertainty (i.e., $V_{\beta, i}$), our model predicts that expected returns decrease with information quality when factor-loading persistence (i.e., a_i) is low, but increase with information quality when factor-loading persistence is high.

Similarly, one must be careful when distinguishing between the effects of uncertainty and information quality on return volatility, as the next result suggests.

Corollary 2. *Conditional on date t information, and adjusting for convexity effects, the idiosyncratic volatility of log returns $r_{i,t+1} = \log(R_{i,t+1})$, is given by $\sigma_{i,t}$, where*

$$\begin{aligned} \sigma_{i,t}^2 &\equiv \text{var}_t \left(r_{i,t+1} - \left(\beta_{i,t+1} m_{i,t+1} - \frac{1}{2} \beta_{i,t+1}^2 V_m \right) \right) \\ &= V_{di} + \left(\frac{\rho_i}{1 - \rho_i a_i} V_m \right)^2 \left(\frac{a_i^2 \Sigma_{i,t}^2}{\Sigma_{i,t} + V_{\varepsilon, i, t}} + \phi_i^2 V_y \right). \end{aligned} \quad (38)$$

¹⁴This mechanism through which higher information quality can lead to higher uncertainty can be found in other dynamic settings. For instance, Banerjee (2011) highlights a similar effect in a dynamic version of Hellwig (1980)'s rational expectations model. In the model, investors receive private signals about next period's dividends, but also condition on prices to update their beliefs. When the quality of investors' information improves, their posterior uncertainty about next period's dividends decreases, but the total risk they face may increase. This is because, as a result of more precise signals, prices are more sensitive to investors' signals and are therefore more volatile. Since investors care about future dividends and future prices, the total risk they face may actually increase as information quality improves.

Note that since factor loadings are stochastic in our model, the systematic component of return volatility is stochastic by construction. Since we want to identify the effect of factor-loading uncertainty and learning on volatility that is not mechanically driven by our assumptions, we focus on the volatility of the idiosyncratic component of returns.

The result shows that idiosyncratic volatility is driven by the volatility of idiosyncratic shocks to dividend growth (i.e., $V_{d,i}$) and the volatility of investors' future conditional expectations of factor loadings (i.e., $\text{var}_t(b_{i,t+1})$). As a result, idiosyncratic volatility is increasing in investors' conditional variance of factor loadings (i.e., $\Sigma_{i,t}$), the variance of shocks to dividend growth (i.e., $V_{d,i}$), and the variance of conditioning variables that investors use to forecast factor loadings (i.e., V_y). Idiosyncratic volatility is also increasing in the persistence of factor loadings (i.e., a_i) and the volatility of the aggregate risk factor (i.e., V_m), since an increase in either variable implies that $b_{i,t+1}$ has a bigger impact on current returns. Finally, the result implies that if the quality of information available to investors is time-varying (i.e., if $V_{\varepsilon,i,t}$ varies over time), then idiosyncratic volatility also exhibits time variation. Moreover, all else equal, the conditional idiosyncratic volatility at date t increases with information quality $\lambda_{i,t}$ (i.e., decreases with $V_{\varepsilon,i,t}$), after controlling for uncertainty (i.e., $\Sigma_{i,t}$). As discussed above, this is because lower information quality (i.e., higher $V_{\varepsilon,i,t}$) implies a noisier signal $Y_{i,t+1}$, which, in turn, implies a lower sensitivity of $b_{i,t+1}$ to $Y_{i,t+1}$, and therefore lower variability of $b_{i,t+1}$ (i.e., lower $\text{var}_t(b_{i,t+1})$).

Note that Proposition 2 and Corollary 2 together imply that controlling for the variation of factor-loading uncertainty, variation in information quality (i.e., $\lambda_{i,t}$) leads to a positive relation between expected returns and idiosyncratic volatility when factor-loading persistence is high enough (i.e., $a_i > \frac{1}{2\rho_i} \left(\sqrt{1 + 8\rho_i^2} - 1 \right)$), but a negative relation when factor-loading persistence is low. As such, we expect to find the relation between expected returns and idiosyncratic volatility, after controlling for factor-loading uncertainty, to be more positive for firms with more persistent factor loadings. In Section 4, we test this prediction on the cross-section of stock returns and find evidence consistent with the prediction (see Table 4).

3.2 Special cases

In this subsection, we consider two extreme special cases of the general model to develop a better understanding of the effects of uncertainty and learning on return dynamics, and to gauge the economic magnitude of these effects. In the first case, we assume that investors receive no signals about the current realization of the factor loading (i.e., $\beta_{i,t}$), and so they price securities based on the unconditional distribution of $\beta_{i,t}$. In the second case, we assume that investors observe $\beta_{i,t}$ perfectly at date t . Together, these two cases characterize the range of possible effects that uncertainty and learning can have on prices and returns. Without much loss of generality, we simplify the exposition of this analysis by setting ϕ_i to zero.

When investors receive no signals about factor loadings at date t (or equivalently, if the noise in signals is infinite—i.e., $V_{\varepsilon,i,t} \rightarrow \infty$), then investors calculate prices and expected returns based

on the unconditional distribution of $\beta_{i,t+1}$, which is given by

$$\beta_{i,t+1} \sim N\left(b_i, \frac{1}{1-a_i^2} V_{\beta,i}\right). \quad (39)$$

Since investors have no information about either the current or future values of $\beta_{i,t}$, investor uncertainty about factor loadings is highest in this case. Moreover, price-dividend ratios and expected returns are constant over time, and are characterized by the following corollary to Proposition 1.

Corollary 3. *Suppose the sequence of prices P_t satisfies the following transversality condition:*

$$\lim_{T \rightarrow \infty} \mathbb{E}_t \left[\frac{M_{t+T}}{M_t} P_{t+T} \right] = 0. \quad (40)$$

Then, given that factor loadings evolve according to (21), and investors' beliefs are given by (39), we have the following:

(a) firm i 's price-dividend ratio is constant, and given by $PD_{i,t} \equiv P_{i,t}/D_{i,t} = \frac{p_i}{1-p_i}$, where

$$p_i = \exp \left\{ \bar{d}_i - r_f - b_i V_m + \frac{1}{2(1-a_i^2)} V_{\beta,i} V_m^2 \right\}, \quad (41)$$

(b) firm i 's return process is of the form

$$R_{i,t+1} = \exp \left\{ \mu_i + \beta_{i,t+1} m_{t+1} + d_{i,t+1} - \frac{1}{2} (\beta_{i,t+1}^2 V_m + V_{d,i}) \right\}, \quad (42)$$

(c) firm i 's expected return is given by

$$\mathbb{E}_t [R_{i,t+1}] = \exp \{ \mu_i \} = \exp \left\{ r_f + b_i V_m - \frac{1}{2(1-a_i^2)} V_{\beta,i} V_m^2 \right\}, \quad \text{and} \quad (43)$$

(d) the idiosyncratic volatility of log returns, $r_{i,t+1} = \log(R_{i,t+1})$ is given by $\sigma_{i,t}$, where

$$\sigma_{i,t}^2 \equiv \text{var}_t \left(r_{i,t+1} - (\beta_{i,t+1} m_{i,t+1} - \frac{1}{2} \beta_{i,t+1}^2 V_m) \right) = V_{d,i}. \quad (44)$$

When investors do not receive any information about firms' factor loadings, their conditional and unconditional beliefs coincide, and as a result, the expressions for prices and expected returns look similar to those in the benchmark model in Section 2. However, the effect of uncertainty is magnified by a factor of $1/(1-a_i^2)$ relative to the i.i.d. case, since the unconditional variance of $\beta_{i,t+1}$ is given by $\frac{1}{(1-a_i^2)} V_{\beta,i}$. Finally, since there is no time-series variation in the conditional expectation of future factor loadings, the idiosyncratic volatility of log returns is simply given by the volatility of idiosyncratic shocks to dividend growth (i.e., $V_{d,i}$).

At the other extreme, the investor uncertainty about factor loadings is lowest when investors can perfectly observe the realization of $\beta_{i,t}$ at date t (or equivalently, if the noise in signals is zero—i.e., $V_{\varepsilon,i,t} = 0$). In this case, the conditional distribution of $\beta_{i,t+1}$, given investors' information at date t , is given by

$$\beta_{i,t+1} | \mathcal{I}_t \sim N((1-a_i) b_i + a_i \beta_{i,t}, V_{\beta,i}), \quad (45)$$

and we have the following result.

Corollary 4. *Suppose the sequence of prices P_t satisfies the following transversality condition:*

$$\lim_{T \rightarrow \infty} \mathbb{E}_t \left[\frac{M_{t+T}}{M_t} P_{t+T} \right] = 0. \quad (46)$$

Then, given that factor loadings evolve according to (21), and investors' beliefs are given by (45), a log-linearization of the price-dividend ratio around its long-run mean implies the following:

(a) *firm i 's price-dividend ratio is given by*

$$pd_{i,t} = \log(PD_{i,t}) = A_{i,t} - \frac{1}{1-\rho_i a_i} ((1-a_i)b_i + a_i \beta_{i,t}) V_m \quad (47)$$

(b) *firm i 's return process is of the form*

$$R_{i,t+1} = \exp \left\{ \alpha_{i,t} + \beta_{i,t+1} m_{t+1} + d_{i,t+1} - \frac{1}{2} (\beta_{i,t+1}^2 V_m + V_{d,i}) + \gamma_i b_{i,t+1} \right\}, \quad (48)$$

(c) *firm i 's expected return, conditional on date t information, is given by*

$$\mathbb{E}_t [R_{i,t+1}] = \exp \left\{ r_f + ((1-a_i)b_i + a_i \beta_{i,t}) V_m - \frac{1}{2} \left(1 + 2 \frac{a_i \rho_i}{1-a_i \rho_i} \right) V_{\beta,i} V_m^2 \right\}, \quad \text{and} \quad (49)$$

(d) *the idiosyncratic volatility of log returns, $r_{i,t+1} = \log(R_{i,t+1})$ is given by $\sigma_{i,t}$, where*

$$\sigma_{i,t}^2 \equiv \text{var}_t \left(r_{i,t+1} - (\beta_{i,t+1} m_{i,t+1} - \frac{1}{2} \beta_{i,t+1}^2 V_m) \right) = V_{d,i} + \left(\frac{\rho_i a_i}{1-\rho_i a_i} V_m \right)^2 V_{\beta,i}, \quad (50)$$

where $A_{i,t}$, $\alpha_{i,t}$, and γ_i are known to investors at date t , and defined as in Proposition 2 (with $b_{i,t} = (1-a_i)b_i + a_i \beta_{i,t}$ and $\Sigma_{i,t} = V_{\beta,i}$), and $\rho_i \equiv \frac{\exp\{\bar{pd}\}}{1+\exp\{\bar{pd}\}} \in (0,1)$, where \bar{pd} is the long-run mean of $pd_{i,t}$.

Although investors observe the current realization of factor loadings perfectly, they still face uncertainty about future realizations of $\beta_{i,t}$. As in the general case in Proposition 2, the price-dividend ratio is log-linear in the conditional expectation of next period's factor loading, which is given by $((1-a_i)b_i + a_i \beta_{i,t})$ in this case. As before, the effect of uncertainty on expected returns is increasing in $V_{\beta,i}$ and a_i , since persistence in factor loadings implies that uncertainty about next period's factor loading leads to uncertainty about factor loadings in all future periods. Finally, as in the general case with imperfect learning, idiosyncratic volatility is driven by the volatility of idiosyncratic shocks to dividend growth (i.e., $V_{d,i}$) and the volatility in the conditional expectations of future factor loadings, which in this case is given by $a_i^2 V_{\beta,i}$.

The two extreme cases characterized by Corollaries 3 and 4 imply that introducing persistence in factor loadings increases the effect of uncertainty on expected returns relative to the i.i.d. case from Section 2. However, as discussed in the previous subsection, the effect on expected returns is not monotonic in the information quality of investors' signal $Y_{i,t}$. Specifically, whether the effect of uncertainty is larger in the full-information case of Corollary 4 or the no-information case of Corollary 3 depends on the persistence in factor loadings, a_i , and the parameter ρ_i . All else equal (and assuming that $\beta_{i,t} = b_i$), expected returns are lower in the

no-information case relative to the full-information case if and only if $\rho_i < \frac{a_i}{2-a_i^2}$. The expressions in Corollaries 3 and 4 also highlight the result that, all else equal, idiosyncratic volatility decreases with the variance of the noise in the signals (i.e., $V_{\varepsilon,i,t}$), since noisier signals about factor loadings imply lower variability in conditional expectations about future factor loadings (i.e., $\text{var}_t(b_{i,t+1})$).

Corollaries 3 and 4 also allow us to characterize the economic magnitude of these effects in a tractable manner. We present a numerical example in Figure 2. We plot expected returns as a function of $V_{\beta,i}$ for the i.i.d. case presented in Section 2, the no-information benchmark presented in Corollary 3, and the full-information benchmark presented in Corollary 4. We set the risk-free rate to 2%, the variance of the pricing kernel V_m to 0.2, the unconditional mean factor loading (i.e., b_i) to 1, and the long-run mean price-dividend ratio $PD_{i,t}$ to 4, which implies $\rho_i = 0.80$. The assumed value of V_m is consistent with a Sharpe ratio on the aggregate stock market of at least 0.5.¹⁵ Although there is significant cross-sectional variation in firms' price-dividend ratios, our estimate of the long-run mean $PD_{i,t}$ is chosen to be conservative and to highlight the non-monotonic dependence of expected returns on information quality as described above.

The plots in Figure 2 suggest that the effect of uncertainty about factor loadings on expected returns can be economically significant, especially when factor loadings are persistent. In the i.i.d. benchmark model of Section 2, the effect is relatively small—a unit increase in the variance of $\beta_{i,t}$ (i.e., $V_{\beta,i} = 1$) leads to a decrease in annualized expected returns of about 200 basis points. However, once we allow for persistence, this effect is amplified considerably. For instance, with a persistence of $a_i = 0.85$, a unit increase in $V_{\beta,i}$ corresponds to a decrease in annualized expected returns of between 520 and 850 basis points (for the no-information and full-information cases, respectively). And increasing the persistence to $a_i = 0.95$ implies that a unit increase in $V_{\beta,i}$ leads to a decrease of between 12 and 19 percent in annualized expected returns. The choice of mean $PD_{i,t}$ is conservative in that we expect the price-dividend ratios of individual firms to be larger, which implies a larger value of ρ_i and, consequently, a larger effect of uncertainty on expected returns (see equation (32)).

For a sense of how representative these parameter ranges are, we can compare them with estimates from portfolio-level analysis in the empirical literature. For example, Jostova and Philipov (2005) find that persistence in betas range from around 0.04 for their “Other” sector portfolio to 0.93 for their “Finance” sector portfolio, while the volatility of shocks to beta range from 0.04 for their “Other” sector portfolio to 0.23 for their “Wholesale and Retail” portfolio.¹⁶ Andersen, Bollerslev, Diebold, and Wu (2005) document that the large-growth and large-value portfolios in their sample exhibit autocorrelation coefficients of 0.91 and 0.97, respectively. In their sample, Ang and Chen (2007) estimate the persistence in the value and

¹⁵This is because the Hansen-Jagannathan bound in this case is given by

$$\sigma\left(\frac{M_{t+1}}{M_t}\right) / \mathbb{E}\left(\frac{M_{t+1}}{M_t}\right) = \frac{\sqrt{(\exp\{V_m\}-1)\exp\{2(-r_f-\frac{1}{2}V_m)+V_m\}}}{\exp\{-r_f\}} = \sqrt{(\exp\{V_m\}-1)} > 0.5,$$

which implies $V_m > \log(1 + 0.5^2) \approx 0.22$.

¹⁶Their “Other” sector consists of all firms excluding those in the “Manufacturing”, “Utilities”, “Wholesale and Retail”, and “Finance” sectors.

growth portfolios as 0.988 and 0.981, respectively, and the volatility of shocks to beta for these portfolios as 0.168 and 0.132, respectively. Adrian and Franzoni (2009) find that persistence in betas range from -0.26 for their large-growth portfolio to 0.35 for their large-value portfolio and that the variance in shocks to betas range from 0.01 for their large-growth portfolio to 0.66 for their small-value portfolio. Moreover, while mapping these empirical estimates to our model parameters, one must keep in mind that these estimates are based on portfolio returns, which are likely to substantially underestimate the volatility of firm-level factor loadings. Therefore, these estimates suggest that there is substantial variation in the persistence and volatility of factor loadings across firms and that the effect of factor-loading uncertainty on the cross-section of expected returns implied by our model is potentially large and economically important.

The plots in Figure 2 also highlight the effect of the interaction between persistence and information quality on expected returns as discussed above. In particular, note that in the first and second panels of Figure 2, a_i is too small relative to condition (37), and as a result, the effect of uncertainty is bigger in the full-information case (i.e., the full-information expected returns are lower, all else equal). In contrast, in the third panel of Figure 2, a_i satisfies condition (37) and, as a result, the effect of uncertainty is bigger in the no-information case. As these panels illustrate, all else equal, and controlling for the level of factor-loading uncertainty (i.e., $\Sigma_{i,t}$), expected returns are increasing in information quality when factor-loading persistence is high, and decreasing in information quality when factor-loading persistence is low.

3.3 Learning from realized dividends and returns

As before, we assume that the pricing kernel evolves according to (1), the dividend process for firm i evolves according to (3), and the factor loading of firm i evolves according to (21). To focus on the effect of learning from realized dividends, we assume that investors do not observe signals of the form (22), but instead can only condition their beliefs on realized dividends and returns. In particular, we assume that returns on a factor-mimicking portfolio, $R_{M,t+1}$, are observable, where

$$R_{M,t+1} = \exp \left\{ \mu_M + m_{t+1} - \frac{1}{2} V_m \right\}. \quad (51)$$

This implies that investors can observe the realization of m_{t+1} at date $t+1$.¹⁷ Moreover, if they can observe the realization of $D_{i,t+1}$, they can recover a signal about $\beta_{i,t+1}$ of the form

$$Y_{i,t+1} = \beta_{i,t+1} m_{t+1} + d_{i,t+1}. \quad (52)$$

Given this information structure, we can again apply the Kalman filter to characterize investors' conditional distribution of factor loadings as follows.

Lemma 2. *The expectation of the risk-factor loading at date $t+2$, conditional on the infor-*

¹⁷This assumption holds in most models of complete information. For instance, in the CAPM, this amounts to observing the return of the market portfolio, $R_{M,t+1}$, at date $t+1$. In a consumption-CAPM, this assumption is equivalent to the representative investor observing consumption growth c_{t+1} at date $t+1$.

mation available to investors at date $t + 1$, is given by:

$$b_{i,t+1} \equiv \mathbb{E}_{t+1} [\beta_{i,t+2}] = (1 - a_i) b_i + a_i \left(b_{i,t} + \lambda_{i,t+1} \left(\frac{1}{m_{t+1}} Y_{i,t+1} - b_{i,t} \right) \right) + \phi_i y_{t+1} \quad (53)$$

$$= (1 - a_i) b_i + a_i \left(b_{i,t} + \lambda_{i,t+1} \left(\beta_{i,t+1} - b_{i,t} + \frac{d_{i,t+1}}{m_{t+1}} \right) \right) + \phi_i y_{t+1} \quad (54)$$

where $\lambda_{i,t+1} = \frac{m_{t+1}^2 \Sigma_{i,t}}{m_{t+1}^2 \Sigma_{i,t} + V_{d,i}}$. The variance of the risk-factor loading at date $t + 2$, conditional on date $t + 1$ information, is given by:

$$\Sigma_{i,t+1} \equiv \text{var}_{t+1} (\beta_{i,t+2}) = V_{\beta,i} + a_i^2 \Sigma_{i,t} (1 - \lambda_{i,t+1}). \quad (55)$$

In contrast to the results in Section 3.1 where the noise in the signal was assumed to be uncorrelated with other fundamentals, noise in the signal $Y_{i,t+1}$ is now the idiosyncratic component of log dividend growth (i.e., $d_{i,t+1}$). More importantly, the above result also implies that the Kalman gain, $\lambda_{i,t+1}$, is not only time-varying, but is also stochastic and depends on realizations of the aggregate risk-factor m_{t+1} . This, in turn, implies that the conditional variance of factor loadings, $\Sigma_{i,t+1}$, is also stochastic, and depends on realizations of m_{t+1} . The stochastic conditional variance of factor loadings is similar to the filtering result in Xia (2001). Intuitively, investors learn more about the factor loading of a given firm when there are big moves in the pricing kernel, or equivalently, large realizations (both positive and negative) of the factor-mimicking portfolio. Moreover, because the conditional variance of the factor loadings depends on the realization of the aggregate risk factor m_{t+1} , there is a common component across stocks—all else equal, the conditional factor-loading uncertainty of all stocks decreases by a larger amount after a large realization of m_{t+1}^2 .

So far we have assumed that investors use only realized dividends, and not additional signals, to keep the analysis tractable. However, as should be apparent from our earlier analysis, we can easily generalize the setup to allow investors to learn from both realized dividends and signals of the form in equation (22). In this more general model, the Kalman gain and the conditional variance would still be stochastic, as in the current setup. Given the evolution of beliefs described above, we can approximate the price-dividend ratio and the expected return as specified in the following result.

Proposition 3. *Suppose the sequence of prices P_t satisfies the following transversality condition:*

$$\lim_{T \rightarrow \infty} \mathbb{E}_t \left[\frac{M_{t+T}}{M_t} P_{t+T} \right] = 0. \quad (56)$$

Then, given that factor loadings evolve according to (21), and investors' beliefs evolve according to the results of Lemma 2, a linearization and a first-order Taylor expansion in m_{t+1} and $d_{i,t+1}$ of the (log) price-dividend ratio gives the following approximations:

(a) *firm i 's price-dividend ratio is approximately given by*

$$pd_{i,t} = \log(PD_{i,t}) = A_{i,t} - \frac{1}{1 - \rho_i a_i} b_{i,t} V_m + F_i \Sigma_{i,t} \quad (57)$$

(b) *firm i 's return process is approximately given by*

$$R_{i,t+1} = \exp \left\{ \alpha_{i,t+1} + \beta_{i,t+1} m_{t+1} + d_{i,t+1} - \frac{1}{2} \left(\beta_{i,t+1}^2 V_m + V_{d,i} \right) + \gamma_i b_{i,t+1} + \rho_i F_i \Sigma_{i,t+1} \right\}, \text{ and} \quad (58)$$

(c) firm i 's expected return, conditional on date t information, is approximately given by

$$\mathbb{E}_t [R_{i,t+1}] = \exp \left\{ r_f + b_{i,t} V_m - \frac{1}{2} V_m^2 \Sigma_{i,t} \right\}, \quad (59)$$

where $F_i = \frac{1}{2(1-a_i^2 \rho_i)} V_m^2$, $\gamma_i = -\frac{\rho_i}{1-\rho_i a_i} V_m$, and $A_{i,t}$, $\alpha_{i,t+1}$, and $\rho_i \in (0, 1)$ are known to investors at date t , and are defined in Appendix A.

As in Proposition 2, the price-dividend ratio varies over time, returns depend on both shocks to dividend growth and conditional expectations about factor loadings (i.e., $b_{i,t+1}$), and expected returns decrease with investors' conditional uncertainty about factor loadings (i.e., $\Sigma_{i,t}$), even after controlling for the level of factor loadings. Moreover, as before, the effect of uncertainty on expected returns increases with the aggregate risk premium, V_m , and the persistence in factor loadings, a_i (since $\Sigma_{i,t}$ increases in a_i), and decreases with the volatility of idiosyncratic shocks to dividend growth, $V_{d,i}$ (since $\lambda_{i,t+1}$ decreases with $V_{d,i}$), holding all else fixed. However, unlike our earlier results, the conditional variance of factor loadings, $\Sigma_{i,t}$, is stochastic in this case because it depends on realizations of the aggregate risk factor m_t . This implies that $\Sigma_{i,t}$ is an additional state variable, and affects the (log) price-dividend ratio and returns nonlinearly. As a result, unlike Proposition 2, a closed-form expression for the exact log-linearized price-dividend ratio is not analytically tractable. However, using a first-order Taylor expansion of the log price-dividend ratio (in m_{t+1} and $d_{i,t+1}$), we can derive approximations that are tractable and intuitive, and parsimoniously capture the first-order effects of learning about factor loadings from realized dividends and returns.¹⁸

Note that the difference in expected returns across stocks, after controlling for the (mean) level of their factor loadings, is likely to be smaller following large realizations of m_t^2 . This is because, all else equal, one expects the conditional variance $\Sigma_{i,t}$ to be smaller following large realizations of m_t^2 , since investors learn more about firms' factor loadings if they can condition on large realizations of the aggregate risk factor in either direction (i.e., large positive or negative realizations of m_t). As such, the model generates time-series variation in the cross-section of expected returns after controlling for the level of factor loadings, or alphas, that is driven by the aggregate risk factor, even though we assume that the aggregate risk premium (i.e., V_m) is constant.

The stochastic nature of $\Sigma_{i,t}$ (and $\lambda_{i,t+1}$) when investors condition on realized dividends and returns implies that idiosyncratic return volatility is stochastic, even though all shocks to fundamentals are homoscedastic (i.e., have constant variance).

Corollary 5. *Conditional on date t information, adjusting for convexity effects, and given the approximations in Proposition 3, the idiosyncratic volatility of log returns $r_{i,t+1} = \log(R_{i,t+1})$, denoted by $\sigma_{i,t}$, is approximately given by*

$$\sigma_{i,t}^2 \approx \gamma_i^2 \phi_i^2 V_y + \gamma_i^2 a_i^2 \Sigma_{i,t} \mathbb{E}_t [\lambda_{i,t+1}^2] + V_{d,i} \mathbb{E}_t \left[\left(1 + \frac{\lambda_{i,t+1}}{m_{t+1}} \gamma_i a_i \right)^2 \right] + \rho_i^2 F_i^2 \text{var}_t (\Sigma_{i,t+1}) \quad (60)$$

where $\gamma_i = -\frac{\rho_i}{1-\rho_i a_i} V_m$ and $F_i = \frac{1}{2(1-a_i^2 \rho_i)} V_m^2$.

¹⁸As we discuss in the proof, we expect higher-order terms of m_{t+1} and $d_{i,t+1}$ to be small relative to the other terms in the log price-dividend ratio, and dropping them allows us to cleanly characterize the approximate price-dividend ratio and expected returns.

This result follows from an application of the law of total variance.¹⁹ As before, the above result implies that idiosyncratic volatility increases in the volatility of idiosyncratic shocks to dividend growth (i.e., $V_{d,i}$), the variance of conditioning variables that investors use to forecast factor loadings (i.e., V_y), the persistence of factor loadings (i.e., a_i), and the volatility of m_{t+1} (i.e., V_m). The model also implies that idiosyncratic volatility is stochastic, since it depends on past realizations of m_t and, moreover, is decreasing in past realizations of m_t^2 . This implies that there is a common factor to idiosyncratic volatility across firms that is driven by realizations of the aggregate risk factor. This result is consistent with empirical evidence of a factor structure in the cross-section of idiosyncratic volatility that is correlated with the business cycle (e.g., Campbell, Lettau, Malkiel, and Xu 2001; Duarte, Kamara, Siegel, and Sun 2011).

While the assumptions of a constant aggregate risk premium (i.e., constant V_m) and homoscedastic shocks to fundamentals are restrictive, they serve to isolate the mechanism through which learning about factor loadings can endogenously generate stochastic volatility. As such, our model provides a useful benchmark for understanding how stochastic volatility can arise at the firm level, and we view our analysis as complementary to models of endogenous stochastic volatility at the aggregate level (e.g., David 1997; Veronesi 1999).

4. Empirical Implications

As the analysis in the previous sections suggests, allowing for uncertainty and learning about factor loadings in a standard asset pricing framework yields a rich set of empirical predictions. Our main result is that controlling for the average level of factor loadings, higher fundamental uncertainty about a firm’s factor loadings (i.e., higher $V_{\beta,i}$) leads to lower expected returns. In Subsection 4.1, we present empirical evidence that is both qualitatively and quantitatively consistent with this primary prediction. In Subsection 4.2, we summarize additional implications of our model that we do not test, but instead leave for future work.

4.1 The effect of factor-loading uncertainty on expected returns

We follow the standard empirical approach of estimating Fama-MacBeth regressions to test whether uncertainty about factor loadings has explanatory power in the cross-section of stock returns, after controlling for the level of firms’ factor loadings. To test this prediction, we ideally need to identify the following: (1) a correctly specified pricing kernel that is used by investors, (2) estimates of each firm’s risk-factor loadings, and (3) estimates of investors’ uncertainty about each firm’s risk-factor loadings. Unfortunately, since none of these variables are directly observable to an econometrician, testing the model’s predictions is challenging. Accordingly, our empirical analysis is accompanied by a number of important caveats.

First, to keep the empirical specification as parsimonious as possible, we adopt the (log) CAPM as the benchmark pricing model. This specification implies that the relevant aggregate risk-factor loading from investors’ perspective (and, therefore, what investors are uncertain

¹⁹The law of total variance implies that for random variables X and Y , $\text{var}(X) = \mathbb{E}[\text{var}(X|Y)] + \text{var}(\mathbb{E}[X|Y])$.

about) is firms' factor loadings with respect to the excess market return. Second, we use factor loadings estimated with (log) returns instead of (log) cash-flow growth. Although the two factor loadings are identical in our model, we recognize that these need not be the same more generally. However, since returns are observable at much a higher frequency than are cash flows, this allows us to estimate $\beta_{i,t+1}$ with greater precision. This also implies that our estimates of the effect of uncertainty on expected returns is conservative, since investors are likely to be *more* uncertain about the factor loadings of firms' cash-flow growth than they are about the factor loadings of firms' returns.

In the first stage, we estimate the average factor loading and the uncertainty about factor loadings by estimating firm-specific rolling window regressions of the excess (log) return on stock i on the excess return on the market—i.e.,

$$r_{i,t+1} - r_{f,t} = a_i + b_i (r_{m,t+1} - r_{f,t}) + e_{i,t+1}, \quad (61)$$

where $r_{i,t+1}$ and $r_{m,t+1}$ are the log return on stock i and the market, respectively, and $r_{f,t}$ is the log risk-free rate. We follow the standard approach in the literature and use monthly returns and a rolling window of sixty months to estimate b_i , and we define our proxy for investors' uncertainty about the factor loading as the squared standard error of this estimate—i.e., $V_{\beta,i} = (\text{std err}(b_i))^2$. In the second stage, we estimate monthly cross-sectional regressions of excess (log) returns on the estimated b_i , our proxy for uncertainty about factor loadings $V_{\beta,i}$, a convexity adjustment term b_i^2 , and a number of additional controls—i.e., for each date t , we estimate the following cross-sectional regression:

$$r_{i,t+1} - r_{f,t} = \alpha_{i,t} + \lambda_t b_{i,t} + \gamma_t V_{\beta,i,t} + \kappa_t b_{i,t}^2 + \text{controls}_{i,t} + \varepsilon_{i,t+1}. \quad (62)$$

Our model predicts that the coefficient on $V_{\beta,i,t}$ (i.e., γ_t) is negative. Since we are using log returns, it is important to include the $b_{i,t}^2$ term in the specification to properly account for the convexity adjustment. The controls include factor loadings on the (log) SMB, HML, and UMD portfolios (denoted by $b_{i,SMB}$, $b_{i,HML}$, and $b_{i,UMD}$, respectively), as well as the factor loading of returns on aggregate volatility risk (denoted by $b_{i,\Delta VIX}$).²⁰ The controls also include proxies for firm characteristics that have been shown to have explanatory power for the cross-section of expected returns, including (log) market value of equity, (log) turnover, market-to-book ratio, debt-to-equity ratio, volatility of earnings, return on assets, accruals, and bid-ask spread, (denoted by $\log(mve)$, $turn$, M/B , D/E , $vol(E)$, ROA , ACC , and $spread$, respectively). The variables are constructed using data from the CRSP and Compustat databases, as described in Appendix C. The t -statistics reported in the table are based on Newey-West standard errors (with sixty lags) that are adjusted by the Shanken correction for the errors-in-variables bias (e.g., Shanken 1992; Chordia, Goyal, and Shanken 2011).

²⁰The ΔVIX proxy is based on Ang, Hodrick, Xing, and Zhang (2006). We estimate rolling window regressions (window size of sixty) of monthly stock returns on innovations in the VIX index (i.e., ΔVIX_t) and on the market return (i.e., MKT_t) using the specification

$$R_{i,t} = a + b_{i,MKT}MKT_t + b_{i,\Delta VIX}\Delta VIX_t + \varepsilon_{i,t}. \quad (63)$$

We then include $b_{i,\Delta VIX}$ as a control for the factor loading on aggregate volatility risk. We follow Ang, Hodrick, Xing, and Zhang (2006) in using the VXO index, which is based on the S&P 500 and is available from January 1986.

Table 1 provides summary statistics for our sample, which consists of monthly returns on all stocks listed in the CRSP database matched with annual variables from the Compustat database. We restrict attention to stocks with positive market value of equity and require that each firm have at least sixty monthly observations. Our sample ranges from January 1964 to December 2011, but some variables in our analysis are available only for a subsample of this period (e.g., the VXO series used to compute $b_{i,\Delta VIX}$ begins in 1986, and the bid and ask series in CRSP that are used to compute *spread* are available beginning in 1983). The summary statistics imply that there is substantial cross-sectional variation in uncertainty about the market-factor loading (i.e., $V_{\beta,i}$), and that firms' market-factor loadings appear to be fairly persistent, which suggests that the effect of factor-loading uncertainty on the cross-section of average returns is likely to be economically significant.

Table 2 presents our main empirical results. Consistent with the model's primary prediction, the coefficient on $V_{\beta,i}$ is negative and statistically significant in every specification. In addition, the effect of factor-loading uncertainty on expected returns appears to be economically significant. The estimated coefficient on $V_{\beta,i}$ ranges from -0.0490 to -0.0646 , which implies that a firm with the median level of $V_{\beta,i}$ (of 0.0678) has annualized average returns of around 400 to 525 basis points lower than an otherwise identical firm with the same average factor loading (i.e., same b_i) but no uncertainty.²¹ Moreover, not only is the effect of $V_{\beta,i}$ on the cross-section of average returns robust to the inclusion of factor loadings on the SMB, HML, UMD, and aggregate volatility risk factors, but the coefficient on $V_{\beta,i}$ actually increases in magnitude and statistical significance when these additional factor loadings are included in the specifications. This suggests that the effect of $V_{\beta,i}$ on the cross-section of returns is not because it serves as a proxy for firms' loadings on aggregate volatility risk, or other risk factors.

While the monthly summary statistics reported in Table 1 are not directly applicable for calibrating the model's predictions at the annual frequency, the magnitude of the annual estimates from the regression analysis seems to be consistent with reasonable parameter values for the model. Figure 3 plots (annual) values of $V_{\beta,i}$, a_i , and ρ_i for which the model generates a spread of 450 basis points on annualized expected returns due to factor-loading uncertainty, which is comparable to the magnitude of the effect we estimate in Table 2. As is apparent from the expressions in Corollaries 3 and 4, the effect of uncertainty increases with $V_{\beta,i}$ and a_i , and so increasing one while decreasing the other can leave the annualized spread unchanged. To get a sense of specific parameter values that match our empirical estimates, note that in the full-information case for a conservative value for the price-dividend ratio (i.e., for $\rho_i = 0.8$, or a long-run mean PD of 4), the following pairs of factor-loading uncertainty and persistence all generate a spread of 450 basis points in annualized expected returns: (i) $a_i = 0.25$, $V_{\beta,i} = 1.5$, (ii) $a_i = 0.5$, $V_{\beta,i} = 0.97$, and (iii) $a_i = 0.9$, $V_{\beta,i} = 0.37$. Similarly, for the no-information case, the following pairs generate the same spread in annualized average returns: (i) $a_i = 0.25$, $V_{\beta,i} = 2.25$, (ii) $a_i = 0.9$, $V_{\beta,i} = 1.615$, and (iii) $a_i = 0.95$, $V_{\beta,i} = 1$. These parameter values appear plausible at the annual frequency, given the skewness in the distributions of serial correlation and uncertainty in factor loadings at the monthly frequency (i.e., $AC(b_i)$ and $V_{\beta,i}$, respectively, in Table 1), and as such, the magnitude of the effects estimated in the regression analysis appear consistent with our model's predictions.

²¹This is because $-0.049 \times 0.0678 \times 12 \approx -0.04$ and $-0.0646 \times 0.0678 \times 12 \approx -0.0525$.

As previously discussed, since we use log returns, the b_i^2 term in the specification (62) plays an important role. As the first two columns of Table 2 show, without the convexity adjustment, the estimated risk premium on the market factor (given by the coefficient of b_i in the regression) is negative. Even though the market risk premium is positive when the convexity adjustment is made, it is not statistically significant after the standard errors are adjusted for serial correlation and the errors-in-variables problem. Finally, comparing the second and third columns, note that the estimate of the market risk premium is significantly higher when we control for cross-sectional variation in $V_{\beta,i}$. Given the positive cross-sectional correlation between b_i and $V_{\beta,i}$ (see Table 1), our results suggest that failing to account for cross-sectional variation in factor-loading uncertainty can lead to estimates of the market risk-premia that are too low, or, alternatively, an estimated security market line (SML) that is too flat (as has been documented by Frazzini and Pedersen 2011 and others).

Table 3 provides evidence that the effect of $V_{\beta,i}$ is robust to the inclusion of a number of firm-level characteristics. As has been documented in the literature, these characteristics are important in capturing cross-sectional variation in firm-level returns. The importance of these characteristics is confirmed in our sample, since their coefficients are often statistically significant and stable across specifications, and their introduction generally leads to an increase in average and median R^2 's. Nevertheless, as in Table 2, the coefficient on $V_{\beta,i}$ remains negative and statistically significant across all specifications, despite the inclusion of these additional characteristics. This suggests that the cross-sectional effect of $V_{\beta,i}$ is distinct from the effects of firm size, market-to-book, leverage, liquidity (as proxied for by *turn* and *spread*), growth and uncertainty about earnings (*ROA* and *vol(E)*, respectively), and the level of accruals. Importantly, note that while factor-loading uncertainty and earnings volatility are positively correlated (with a correlation coefficient of over 0.4), they each appear to have incremental explanatory power for the cross-section of average returns.

Since idiosyncratic volatility is positively related to factor-loading uncertainty (as reported in Table 1, the correlation between these two variables in our sample is over 0.8), our model can generate a negative cross-sectional relation between expected returns and idiosyncratic volatility, as documented by Ang, Hodrick, Xing, and Zhang (2006), Ang, Hodrick, Xing, and Zhang (2009), and others. However, the distinction between uncertainty and information quality in the dynamic setting that we highlight in Section 3 provides a more subtle prediction about the joint relation between factor-loading uncertainty, idiosyncratic volatility, and expected returns that allows us to distinguish our model from other potential explanations for the negative expected return-idiosyncratic volatility relation. As Corollary 2 highlights, controlling for factor-loading uncertainty, idiosyncratic volatility increases in information quality. Moreover, as our discussion in Section 3.2 concludes (and the second and third panels of Figure 2 illustrate), controlling for factor-loading uncertainty, we expect a negative relation between information quality and expected returns for firms with low factor-loading persistence, but a positive relation for firms with high persistence. These results suggest that controlling for cross-sectional variation in factor-loading uncertainty, cross-sectional variation in information quality should generate a *negative* relation between expected returns and idiosyncratic volatility for firms with *low* factor-loading persistence, but a *positive* relation for firms with *high* factor-loading persistence.

In practice, testing this prediction directly is difficult, since it requires identifying distinct

firm-specific proxies for factor-loading uncertainty and information quality. Instead, we test this prediction indirectly, and report the results in Table 4. For each firm, we estimate the component of its idiosyncratic volatility that is uncorrelated with its factor-loading uncertainty. In particular, we estimate firm-specific regressions of idiosyncratic volatility (i.e., σ_i^2) on $V_{\beta,i}$, given by

$$\sigma_{i,t}^2 = c_i + d_i V_{\beta,i,t} + e_{i,t}, \quad (64)$$

and denote the conditional mean by $PIV_{i,t} = c_i + d_i V_{\beta,i,t}$ and the residuals by $RIV_i = e_{i,t}$. We then estimate the following second-stage regression:

$$r_{i,t+1} - r_{f,t} = \alpha_{i,t} + \lambda_t b_{i,t} + \gamma_{1,t} PIV_{i,t} + \gamma_{2,t} RIV_{i,t} + \gamma_{3,t} RIV_{i,t} 1_{HAC} + \kappa_t b_{i,t}^2 + \text{controls}_{i,t} + \varepsilon_{i,t+1}, \quad (65)$$

where 1_{HAC} is an indicator for firms whose factor-loading persistence (i.e., $AC(b_i)$) is in the top 25th percentile for our sample.²² Since the model predicts a negative relation between factor-loading uncertainty and expected returns, we expect $\gamma_{1,t}$ to be negative. Moreover, if cross-sectional variation in the residual component of idiosyncratic volatility (i.e., $RIV_{i,t}$) is driven by cross-sectional variation in information quality, then our model predicts that $\gamma_{2,t}$ is negative and $\gamma_{3,t}$ is positive.

The results in Table 4 are consistent with this prediction. Specifically, the coefficient on PIV is negative and statistically significant, suggesting a negative relation between expected returns and the component of idiosyncratic volatility that is related to factor-loading uncertainty. Also consistent with the model's prediction, the coefficient on RIV is negative (although not always statistically significant) and the coefficient on $RIV 1_{HAC}$ is positive and almost always statistically significant. Collectively, these results suggest that after controlling for the effect of factor-loading uncertainty, the relation between idiosyncratic volatility and expected returns is quite nuanced and depends on the persistence of firms' factor-loadings, as predicted by our model.²³

Finally, as a robustness test of our main prediction, we report results from repeating the analysis in Tables 2 and 3 using simple returns. Although our model is log-linear, and our benchmark empirical analysis in Table 2 using log returns tests this specification directly, the main intuition of Proposition 1 is that uncertainty about risk-factor loadings generates cross-sectional variation in expected returns that is not captured by the average level of firms' factor loadings. This intuition suggests that one should expect to find a cross-sectional effect of uncertainty about factor loadings even when using simple returns. To test this intuition, we estimate the factor loading on the excess market return (i.e., b_i), and the uncertainty about factor loadings $V_{\beta,i}$ using sixty month rolling window regressions of simple returns, and use these estimates in the following second-stage cross-sectional regression:

$$R_{i,t+1} - R_{f,t} = \alpha_{i,t} + \lambda_t b_{i,t} + \gamma_t V_{\beta,i,t} + \text{controls}_{i,t} + \varepsilon_{i,t+1}. \quad (66)$$

²²The results are qualitatively the same if we consider, instead, firms in the top 20th percentile or the top 10th percentile of factor-loading persistence.

²³Recall that the relation between idiosyncratic volatility and expected returns in our model depends not only on the persistence parameter a_i but also on the long-run mean of the log price-dividend ratio ρ_i . Given the substantial heterogeneity in ρ_i across firms, it is not feasible to precisely identify the cutoff level of a_i (in expression (37)) for each firm, and the noise this introduces may partially explain the lack of statistical significance of our estimates of $\gamma_{2,t}$ and $\gamma_{3,t}$.

The controls, as before, are loadings on other risk-factors and firm-level characteristics. Also note that there is no convexity adjustment term in this specification, since we are using simple, rather than log, returns. The results of this estimation are presented in Table 5. Consistent with the intuition of our main result, we find that the coefficient on $V_{\beta,i}$ is always negative, and is statistically significant across almost all specifications. Relative to Tables 2 and 3, the magnitude and statistical significance of the coefficient estimates are attenuated for some specifications but remain economically important.²⁴ As in Table 3, the estimates of the risk premiums are not statistically significant, but the coefficients on the firm-level characteristics often are, highlighting their importance in explaining the cross-section of firm-level returns.

Overall, the empirical evidence across our tests is both qualitatively and quantitatively consistent with the model’s main prediction, and suggests that cross-sectional variation in factor-loading uncertainty may be an important determinant of the cross-section of expected returns. Our results also suggest that the distinction between factor-loading uncertainty and information quality may be empirically relevant, and in particular, may be useful to consider in reconciling the seemingly conflicting empirical evidence about the cross-sectional relation between expected returns and idiosyncratic volatility that has been documented in the literature. For instance, while Ang, Hodrick, Xing, and Zhang (2006) and Ang, Hodrick, Xing, and Zhang (2009) document a negative relation between idiosyncratic volatility and expected returns, Lehmann (1990), Malkiel and Xu (2006), Boehme, Danielsen, Kumar, and Sorescu (2009), Fu (2009), Huang, Liu, Rhee, and Zhang (2010), and Spiegel and Wang (2010) document a positive relation. Our results show that the unconditional relation between expected returns and idiosyncratic volatility may be negative if cross-sectional variation in idiosyncratic volatility is dominated by cross-sectional variation in firms’ factor-loading uncertainty. However, after controlling for the cross-sectional variation in factor-loading uncertainty, the relation between expected returns and idiosyncratic volatility is driven by information quality, and hence may be negative for some firms and positive for others. More generally, our model may also provide insight into the mixed empirical evidence about the relation between expected returns and other firm-level proxies for uncertainty (e.g., analyst forecast dispersion), to the extent these proxies capture investor uncertainty and information quality about risk-factor loadings.²⁵

4.2 Additional implications

Although a more exhaustive empirical analysis of our model’s various predictions is beyond the scope of this paper, we summarize a number of additional implications to guide future empirical

²⁴Note that while the coefficient estimates on $V_{\beta,i}$ in Table 5 are smaller in magnitude than those in Tables 2 and 3, the distribution of $V_{\beta,i}$ across firms is also different when we use simple returns as opposed to log-returns. In particular, note that the mean and median estimates for $V_{\beta,i}$ when using simple returns are 0.1474 and 0.071, respectively, which suggests that the effect of uncertainty on expected returns is still economically significant.

²⁵Again, the empirical evidence for this relation is mixed. On one hand, Diether, Malloy, and Scherbina (2002), Johnson (2004), and others document a negative relation between analyst forecast dispersion and expected returns. On the other hand, Qu, Starks, and Yan (2004) and Banerjee (2011) document a positive relation between expected returns and analyst forecast dispersion, while Anderson, Ghysels, and Juergens (2005) find evidence of a negative relation between expected returns and short-term analyst forecast dispersion, but a positive relation between expected returns and long-term dispersion.

research. First, as Corollary 1 highlights, relative to a standard log-linear CAPM, our model predicts that alpha decreases with firm-specific factor-loading uncertainty (i.e., $V_{\beta,i}$). This basic relation generates additional cross-sectional predictions. As previously noted, a positive cross-sectional correlation between the level of factor loadings and factor-loading uncertainty implies that the slope of an estimated security market line (SML) in log-return space would be “too” flat, which is consistent with the evidence documented by Frazzini and Pedersen (2011) and others. Similarly, since investors are likely to face greater uncertainty about the risk-factor loadings of young firms that have recently issued stock, our model’s main prediction is also consistent with the relative underperformance of firms following their initial public offerings that has been documented in the literature (e.g., Ritter 1991; Teoh, Welch, and Wong 1998).

Our model with persistent factor loadings and learning provides a finer set of empirical predictions. For instance, the model implies that all else equal, the negative relation between expected returns and factor-loading uncertainty is stronger for firms with more persistent factor loadings (i.e., larger a_i) and higher long-run mean price-dividend ratios (i.e., higher ρ_i). As such, one should expect the strength of the relation between expected returns and firm-specific uncertainty to vary across industries and according to firm characteristics. The model also generates potentially testable implications for the cross-section of price-dividend ratios and idiosyncratic return volatility. In particular, the model suggests that price-dividend ratios are increasing in investors’ conditional uncertainty about factor loadings (i.e., $\Sigma_{i,t}$). Moreover, price-dividend ratios should be more sensitive to conditional expectations of future factor loadings (i.e., $b_{i,t}$) for firms with more persistent factor loadings (i.e., larger a_i) and higher long-run mean price-dividend ratios (i.e., higher ρ_i). The model also predicts that idiosyncratic volatility increases in investors’ uncertainty about factor loadings (i.e., $V_{\beta,i}$ or $\Sigma_{i,t}$), and all else equal, increases in both the persistence of factor loadings (i.e., a_i) and the long-run mean price-dividend ratio (i.e., in ρ_i).

Finally, the version of the model presented in Section 3.3 formalizes a very intuitive mechanism for how learning about factor loadings from realized dividends and returns can generate stochastic volatility in the idiosyncratic component of returns. Since the conditional uncertainty about factor loadings depends on past realizations of the aggregate risk factor, the model generates additional testable predictions that condition on past realizations of the factor-mimicking portfolio. For instance, the results suggest that the spread in expected returns across firms after controlling for the average level of their factor loadings (i.e., the spread in firms’ alphas relative to their factor loadings) should be smaller following large realizations (in either direction) of the aggregate risk factor (i.e., large m_t^2). This is because, all else equal, investors learn more about firms’ factor loadings from large realizations of m_t (in either direction), and as a result, their conditional uncertainty about factor loadings should be small. The model therefore predicts a factor structure to idiosyncratic volatility across firms, which is driven by past realizations of m_t^2 . Finally, the model suggests that in a more general model with multiple aggregate risk factors that affect the cross-section of stocks differently, one should expect to find a multifactor structure to the cross-section of idiosyncratic volatility. Accordingly, stocks that are exposed to the same set of risk factors (e.g., stocks in similar industries) should exhibit not only comovement in their expected returns, but also comovement in their idiosyncratic volatility.

5. Conclusions

We argue that firm-specific information can affect expected returns if it affects investors' uncertainty about firms' risk-factor loadings. We develop a simple, partial equilibrium model in which we relax the standard assumption that investors know firms' risk-factor loadings with certainty. We show that firms with greater uncertainty about their factor loadings have lower expected returns, even after controlling for the average level of their factor loadings. Moreover, when factor loadings are persistent and investors can learn about them, this introduces time-series variation in price-dividend ratios and idiosyncratic volatility. In particular, if investors use realized dividends and returns to learn about risk-factor loadings, returns exhibit stochastic volatility even if all fundamental shocks are homoscedastic.

Using Fama-MacBeth regressions of the cross-section of stock returns, we document empirical evidence that is consistent with our main prediction. Specifically, we find that all else equal, a firm with the median level of factor-loading uncertainty (in our sample) has annualized average returns that are roughly 400 to 525 basis points lower than an otherwise comparable firm with the same average factor loading (i.e., same b_i) but with no factor-loading uncertainty. This effect is robust to the inclusion of additional risk-factor loadings and firm characteristics. Finally, our model helps shed light on the seemingly ambiguous cross-sectional relation between expected returns and idiosyncratic volatility. Consistent with the model's prediction, we find that while the unconditional relation between expected returns and idiosyncratic volatility may be negative, after controlling for cross-sectional variation in factor-loading uncertainty, the relation depends on the persistence in factor loadings, and may be positive for some firms but negative for others.

An important limitation of our analysis is that it is done in a partial equilibrium setting. In particular, we do not impose the general equilibrium restriction that the aggregate dividends (or cash flows) in the economy must sum to aggregate output or consumption. This is purely to preserve tractability. As highlighted by Menzly, Santos, and Veronesi (2004), Santos and Veronesi (2006), Cochrane, Longstaff, and Santa-Clara (2008), Martin (2011), and others, imposing the market-clearing condition on dividends in a multiple-asset framework is generally not analytically tractable since it involves computing the sum of log-normal random variables. In our framework, this exercise is further complicated by the assumption that the risk-factor loading of a firm's dividend growth is stochastic. We leave the general equilibrium analysis for future work, but conjecture that our model's predictions should continue to be important if the number of firms increases and each firm's dividend becomes a negligible fraction of the aggregate dividend process.

Finally, although our model is very stylized, it provides a tractable benchmark for more sophisticated analysis. For instance, in a model with time-varying aggregate risk premia, one would expect to find that the effect of factor-loading uncertainty also depends on the level of aggregate uncertainty. In such a model, one would expect that if acquiring information is costly, investors are likely to condition their information acquisition and portfolio allocation decisions on the interaction between aggregate and firm-specific uncertainty. Similarly, since our model predicts that firm-specific information can affect firms' expected returns, we expect that optimal disclosure policy should vary over time and according to macroeconomic conditions. A complete

analysis of these and other extensions is beyond the scope of this paper and is left for future research.

Appendix A—Proofs

Proof of Proposition 1. Note that the price-dividend ratio at time t is given by:

$$PD_{i,t} = \mathbb{E}_t \left[\frac{M_{t+1}}{M_t} \frac{D_{i,t+1}}{D_{i,t}} (1 + PD_{i,t+1}) \right] = \mathbb{E}_t \left[\frac{M_{t+1}}{M_t} \frac{D_{i,t+1}}{D_{i,t}} \left(1 + \mathbb{E}_{t+1} \left[\frac{M_{t+2}}{M_{t+1}} \frac{D_{i,t+2}}{D_{i,t+1}} (1 + PD_{i,t+2}) \right] \right) \right]$$

and so on. Importantly, our assumption that $\beta_{i,t+1}$ is independent over time implies that

$$\frac{M_{t+1}}{M_t} \frac{D_{i,t+1}}{D_{i,t}} = \exp \left\{ \bar{d}_i + (\beta_{i,t+1} - 1) m_{t+1} + d_{i,t+1} - \frac{1}{2} V_{d,i} - r_f - \frac{1}{2} V_m - \frac{1}{2} \beta_{i,t+1}^2 V_m \right\} \quad (\text{A1})$$

is independent of

$$\frac{M_{t+2}}{M_{t+1}} \frac{D_{i,t+2}}{D_{i,t+1}} = \exp \left\{ \bar{d}_i + (\beta_{i,t+2} - 1) m_{t+2} + d_{i,t+2} - \frac{1}{2} V_{d,i} - r_f - \frac{1}{2} V_m - \frac{1}{2} \beta_{i,t+2}^2 V_m \right\}. \quad (\text{A2})$$

Moreover, the price-dividend ratio of next period's dividend is given by

$$p_t = \mathbb{E}_t \left[\frac{M_{t+1}}{M_t} \frac{D_{i,t+1}}{D_{i,t}} \right] = \mathbb{E}_t \left[\mathbb{E}_t \left[\frac{M_{t+1}}{M_t} \frac{D_{i,t+1}}{D_{i,t}} \mid \beta_{i,t+1} \right] \right] \quad (\text{A3})$$

$$= \mathbb{E}_t \left[\exp \left\{ \bar{d}_i - r_f + \frac{1}{2} (\beta_{i,t+1} - 1)^2 V_m - \frac{1}{2} V_m - \frac{1}{2} \beta_{i,t+1}^2 V_m \right\} \right] \quad (\text{A4})$$

$$= \mathbb{E}_t \left[\exp \left\{ \bar{d}_i - r_f - \beta_{i,t+1} V_m \right\} \right] = \exp \left\{ \bar{d}_i - r_f - b_i V_m + \frac{1}{2} V_m^2 V_{\beta,i} \right\} \equiv p, \quad (\text{A5})$$

and hence it is constant (i.e., $p_t = p_{t+1} = p$), and therefore

$$PD_{i,t} = p + p \mathbb{E}_t [PD_{i,t+1}] = p + p \mathbb{E}_t [p + p \mathbb{E}_{t+1} [PD_{i,t+2}]] = p + p^2 + \dots = \frac{p}{1-p}.$$

Since $R_{i,t+1} = \frac{D_{i,t+1}}{D_{i,t}} \frac{1+PD_{i,t+1}}{PD_{i,t}} = \frac{D_{i,t+1}}{D_{i,t}} \frac{1}{p}$, this implies

$$R_{i,t+1} = \exp \left\{ r_f + \beta_{i,t+1} m_{t+1} + d_{i,t+1} - \frac{1}{2} \beta_{i,t+1}^2 V_m - \frac{1}{2} V_{d,i} + b_i V_m - \frac{1}{2} V_m^2 V_{\beta,i} \right\} \quad (\text{A6})$$

$$\mathbb{E} [R_{i,t+1}] = \frac{1}{p} \mathbb{E} \left[\frac{D_{i,t+1}}{D_{i,t}} \right] = \exp \left\{ r_f + b_i V_m - \frac{1}{2} V_{\beta,i} V_m^2 \right\}, \quad (\text{A7})$$

which implies that the firm i 's expected return is decreasing in its factor-loading uncertainty $V_{\beta,i}$. ■

Proof of Proposition 2. Given persistence in $\beta_{i,t+1}$, in general it is not possible to derive exact expressions in closed form for our discrete time model. We shall instead rely on a log-linearization of the price-dividend ratio that is common in the literature to derive these expressions. Let $PD_{i,t}$ denote the price-dividend ratio of stock i at date t . Let $pd_{i,t} = \log(PD_{i,t})$, and note that a first-order Taylor expansion yields:

$$\log(1 + PD_{i,t+1}) = \log(1 + \exp\{pd_{i,t+1}\}) \approx C + \rho_i pd_{i,t+1} \quad (\text{A8})$$

where $\rho_i \equiv \frac{\exp\{\bar{pd}\}}{1 + \exp\{\bar{pd}\}}$, $C \equiv \log(1 + \exp\{\bar{pd}\}) - \frac{\exp\{\bar{pd}\}}{1 + \exp\{\bar{pd}\}} \bar{pd}$, and \bar{pd} is the long-run mean of

the log price-dividend ratio. Now conjecture that $pd_{i,t} = A_{i,t} + B_i b_{i,t}$. Then,

$$\begin{aligned}
PD_{i,t} &= \mathbb{E}_t \left[\frac{M_{t+1}}{M_t} \frac{D_{i,t+1}}{D_{i,t}} (1 + PD_{i,t+1}) \right] \\
&\approx \mathbb{E}_t \left[\exp \left\{ \begin{aligned} &\bar{d}_i - r_f + (\beta_{i,t+1} - 1) m_{t+1} + d_{i,t+1} - \frac{1}{2} \beta_{i,t+1}^2 V_m - \frac{1}{2} V_{d,i} - \frac{1}{2} V_m \\ &+ C + \rho_i (A_{i,t+1} + B_i b_{i,t+1}) \end{aligned} \right\} \right] \\
&= \mathbb{E}_t \left[\mathbb{E}_t \left[\exp \left\{ \begin{aligned} &\bar{d}_i - r_f + (\beta_{i,t+1} - 1) m_{t+1} + d_{i,t+1} - \frac{1}{2} \beta_{i,t+1}^2 V_m - \frac{1}{2} V_{d,i} \\ & - \frac{1}{2} V_m + C + \rho_i A_{i,t+1} + \rho_i B_i (1 - a_i) b_i \\ & + \rho_i B_i [a_i (b_{i,t} + \lambda_{i,t} (\beta_{i,t+1} + \varepsilon_{i,t+1} - b_{i,t})) + \phi_i y_{t+1}] \end{aligned} \right\} \middle| \beta_{i,t+1} \right] \right] \\
&= \mathbb{E}_t \left[\exp \left\{ \begin{aligned} &\bar{d}_i - r_f + \frac{1}{2} (\beta_{i,t+1} - 1)^2 V_m - \frac{1}{2} \beta_{i,t+1}^2 V_m - \frac{1}{2} V_m \\ & + C + \rho_i A_{i,t+1} + \rho_i B_i (1 - a_i) b_i + \frac{1}{2} (\rho_i B_i)^2 (a_i^2 \lambda_{i,t}^2 V_{\varepsilon,i,t} + \phi_i^2 V_y) \\ & + \rho_i B_i [a_i (b_{i,t} + \lambda_{i,t} (\beta_{i,t+1} - b_{i,t}))] \end{aligned} \right\} \right] \\
&= \mathbb{E}_t \left[\exp \left\{ \begin{aligned} &\bar{d}_i - r_f + \beta_{i,t+1} (\rho_i B_i a_i \lambda_{i,t} - V_m) + \rho_i B_i a_i b_{i,t} (1 - \lambda_{i,t}) \\ & + C + \rho_i A_{i,t+1} + \rho_i B_i (1 - a_i) b_i + \frac{1}{2} (\rho_i B_i)^2 (a_i^2 \lambda_{i,t}^2 V_{\varepsilon,i,t} + \phi_i^2 V_y) \end{aligned} \right\} \right] \\
&= \exp \left\{ \begin{aligned} &\bar{d}_i - r_f + (\rho_i B_i a_i - V_m) b_{i,t} + \frac{1}{2} (\rho_i B_i a_i \lambda_{i,t} - V_m)^2 \Sigma_{i,t} \\ & + C + \rho_i A_{i,t+1} + \rho_i B_i (1 - a_i) b_i + \frac{1}{2} (\rho_i B_i)^2 (a_i^2 \lambda_{i,t}^2 V_{\varepsilon,i,t} + \phi_i^2 V_y) \end{aligned} \right\}
\end{aligned}$$

Matching terms to our conjecture, we get

$$B_i = (\rho_i B_i a_i - V_m) \text{ or equivalently, } B_i = -\frac{1}{1 - \rho_i a_i} V_m \quad (\text{A9})$$

$$A_{i,t} = \begin{aligned} &\bar{d}_i - r_f + \frac{1}{2} (\rho_i B_i a_i \lambda_{i,t} - V_m)^2 \Sigma_{i,t} \\ & + C + \rho_i A_{i,t+1} + \rho_i B_i (1 - a_i) b_i + \frac{1}{2} (\rho_i B_i)^2 (a_i^2 \lambda_{i,t}^2 V_{\varepsilon,i,t} + \phi_i^2 V_y) \end{aligned} \quad (\text{A10})$$

Note that $A_{i,t}$ follows the recursion

$$A_{i,t} = \rho_i A_{i,t+1} + \chi_t, \quad (\text{A11})$$

where $\chi_s = \bar{d}_i - r_f + \frac{1}{2} (\rho_i B_i a_i \lambda_{i,s} - V_m)^2 \Sigma_{i,s} + C + \rho_i B_i (1 - a_i) b_i + \frac{1}{2} (\rho_i B_i)^2 (a_i^2 \lambda_{i,s}^2 V_{\varepsilon,i,s} + \phi_i^2 V_y)$ is bounded. This implies that $A_{i,t}$ has a solution of the form:

$$A_{i,t} = K \rho_i^{-t} + \sum_{s=t}^{\infty} \rho_i^{s-t} \chi_s. \quad (\text{A12})$$

The transversality condition implies that

$$\begin{aligned}
\lim_{T \rightarrow \infty} \mathbb{E}_t \left[\frac{M_T}{M_t} P_{t+T} \right] &= \lim_{T \rightarrow \infty} D_{i,t} \mathbb{E}_t \left[\frac{M_{t+T}}{M_t} \times \frac{D_{i,t+T}}{D_{i,t}} \times PD_{i,t+T} \right] \\
&= \lim_{T \rightarrow \infty} D_{i,t} \mathbb{E}_t \left[\exp \left\{ K \rho_i^{-(t+T)} + \sum_{s=t}^{\infty} \rho_i^{s-t} \chi_s - \frac{b_{i,t+T} V_m}{1 - \rho_i a_i} - T (r_f - \bar{d}_i) - \sum_{s=t}^{t+T} \beta_{i,s} V_m \right\} \right] = 0,
\end{aligned}$$

which implies that $K \leq 0$. Moreover, note that if $K < 0$, then $\log(PD_{i,T})$ decreases exponentially fast as $T \rightarrow \infty$ —that is,

$$\text{plim}_{T \rightarrow \infty} \log(PD_{i,T}) = \text{plim}_{T \rightarrow \infty} K \rho_i^{-T} + \sum_{s=T}^{\infty} \rho_i^{s-T} \chi_s - \frac{b_{i,T} V_m}{1 - \rho_i a_i} = -\infty. \quad (\text{A13})$$

But log dividend growth (i.e., $\log(\frac{D_{i,T}}{D_{i,t}})$) is linear in T (by assumption). This implies that the price converges to zero as $T \rightarrow \infty$ (in probability), since $\log(P_T) = \log(PD_{i,T}) + \log(\frac{D_{i,T}}{D_{i,t}}) + \log(D_{i,t})$. Since we want to rule out this degeneracy in price, we must have that $K = 0$, which implies $A_{i,t} = \sum_{s=t}^{\infty} \rho_i^{s-t} \chi_s$ for all t . If $V_{\varepsilon,i,t}$ is constant, the stationary equilibrium is characterized by $A_{i,t+1} = A_{i,t} = A_i$, where

$$A_i = \frac{1}{1-\rho_i} \left(\bar{d}_i - r_f + \frac{1}{2} (\rho_i B_i a_i \lambda_i - V_m)^2 \Sigma_i + C + \rho_i B_i (1 - a_i) b_i + \frac{1}{2} (\rho_i B_i)^2 (a_i^2 \lambda_i^2 V_{\varepsilon,i} + \phi_i^2 V_y) \right),$$

where λ_i and Σ_i are characterized in Lemma 1.

Next, note that

$$\begin{aligned} R_{i,t+1} &= \frac{D_{i,t+1}}{D_{i,t}} \frac{1+PD_{i,t+1}}{PD_{i,t}} \\ &\approx \exp \left\{ \bar{d}_i + \beta_{i,t+1} m_{t+1} + d_{i,t+1} - \frac{1}{2} \beta_{i,t+1}^2 V_m - \frac{1}{2} V_{d,i} + C + \rho_i (A_{i,t+1} + B_i b_{i,t+1}) - (A_{i,t} + B_i b_{i,t}) \right\} \\ &\equiv \exp \left\{ \alpha_{i,t} + \beta_{i,t+1} m_{t+1} + d_{i,t+1} - \frac{1}{2} (\beta_{i,t+1}^2 V_m + V_{d,i}) + \gamma_i b_{i,t+1} \right\}, \end{aligned}$$

where $\gamma_i = -\frac{\rho_i}{1-\rho_i a_i} V_m = \rho_i B_i$, and

$$\begin{aligned} \alpha_{i,t} &= \bar{d}_i + C + \rho_i A_{i,t+1} - A_{i,t} - B_i b_{i,t} \\ &= r_f - \frac{1}{2} (\rho_i B_i a_i \lambda_{i,t} - V_m)^2 \Sigma_{i,t} - B_i (\rho_i (1 - a_i) b_i + b_{i,t}) - \frac{1}{2} (\rho_i B_i)^2 (a_i^2 \lambda_{i,t}^2 V_{\varepsilon,i,t} + \phi_i^2 V_y). \end{aligned}$$

Finally, note that expected returns are given by

$$\begin{aligned} \mathbb{E}_t [R_{i,t+1}] &= \mathbb{E}_t \left[\exp \left\{ \alpha_{i,t} + \beta_{i,t+1} m_{t+1} + d_{i,t+1} - \frac{1}{2} (\beta_{i,t+1}^2 V_m + V_{d,i}) + \gamma_i b_{i,t+1} \right\} \right] \\ &= \mathbb{E}_t \left[\exp \left\{ \begin{aligned} &\alpha_{i,t} + \beta_{i,t+1} m_{t+1} + d_{i,t+1} - \frac{1}{2} (\beta_{i,t+1}^2 V_m + V_{d,i}) \\ &+ \gamma_i ((1 - a_i) b_i + a_i (b_{i,t} + \lambda_{i,t} (\beta_{i,t+1} + \varepsilon_{i,t+1} - b_{i,t})) + \phi_i y_{t+1}) \end{aligned} \right\} \right] \\ &= \exp \left\{ r_f - \frac{1}{2} (\rho_i B_i a_i \lambda_{i,t} - V_m)^2 \Sigma_{i,t} + \frac{1}{2} (\rho_i B_i a_i \lambda_{i,t})^2 \Sigma_{i,t} + (\rho_i a_i - 1) B_i b_{i,t} \right\} \\ &= \exp \left\{ r_f + b_{i,t} V_m - \frac{1}{2} \left(1 + 2\lambda_{i,t} \frac{\rho_i a_i}{1-\rho_i a_i} \right) V_m^2 \Sigma_{i,t} \right\}, \end{aligned}$$

which gives us the result. ■

Proof of Proposition 3. As in the proof of Proposition 2, we log-linearize the price-dividend around its long-run mean. However, given the nonlinearity in the filtering by investors, this is not sufficient to derive closed form expressions. As such, we derive an approximation to the log price-dividend ratio by using a first-order Taylor expansion in m_{t+1} and $d_{i,t+1}$. Specifically, conjecture a log-linear price-dividend ratio of the form $pd_{i,t} = A_t + B_i b_{i,t} + F_i \Sigma_{i,t}$. Plugging this conjecture into the definition, we have

$$PD_{i,t} = \mathbb{E}_t \left[\frac{M_{t+1}}{M_t} \frac{D_{i,t+1}}{D_{i,t}} (1 + PD_{i,t+1}) \right] \approx \mathbb{E}_t \left[\exp \{ G(m_{t+1}, d_{i,t+1}, \beta_{i,t+1}, y_{i,t+1}) \} \right]$$

where

$$\begin{aligned} G = & \bar{d}_i - r_f + (\beta_{i,t+1} - 1) m_{t+1} + d_{i,t+1} - \frac{1}{2} \beta_{i,t+1}^2 V_m - \frac{1}{2} V_{d,i} - \frac{1}{2} V_m \\ & + C + \rho_i (A_{i,t+1} + B_i b_{i,t+1} + F_i \Sigma_{i,t+1}) \end{aligned} \quad (\text{A14})$$

and where $b_{i,t+1}$, $\Sigma_{i,t+1}$ and $\lambda_{i,t+1}$ are given by Lemma 2. A first-order Taylor expansion of G in m_{t+1} and $d_{i,t+1}$ around $(0, 0)$ gives:

$$G = \bar{d}_i - r_f + C - \frac{1}{2}V_{d,i} - \frac{1}{2}V_m - \frac{1}{2}\beta_{i,t+1}^2 V_m + \rho_i A_{i,t+1} + \rho_i B_i (b_i + a_i (b_{i,t} - b_i)) + a_i^2 \rho_i F_i \Sigma_{i,t} + \rho_i F_i V_{\beta,i} + \rho_i \phi_i B_i y_{i,t+1} + d_{i,t+1} + (\beta_{i,t+1} - 1) m_{t+1} + O(x) \quad (\text{A15})$$

where $x \in \{m_{t+1}^2, d_{i,t+1}^2, m_{t+1}d_{i,t+1}\}$. Intuitively, we eliminate the higher order terms of m_{t+1} and $d_{i,t+1}$ since we expect them to be small. However, we do not eliminate $\beta_{i,t+1}m_{t+1}$ terms since $\beta_{i,t+1}$ are not likely to be as small as m_{t+1} or $d_{i,t+1}$. This implies

$$\begin{aligned} PD_{i,t} &\approx \mathbb{E}_t \left[\exp \left\{ \bar{d}_i - r_f + C - \frac{1}{2}V_{d,i} - \frac{1}{2}V_m - \frac{1}{2}\beta_{i,t+1}^2 V_m + \rho_i A_{i,t+1} + \rho_i B_i (b_i + a_i (b_{i,t} - b_i)) + a_i^2 \rho_i F_i \Sigma_{i,t} + \rho_i F_i V_{\beta,i} + \rho_i \phi_i B_i y_{i,t+1} + d_{i,t+1} + (\beta_{i,t+1} - 1) m_{t+1} + O(x) \right\} \right] \\ &= \mathbb{E}_t \left[\exp \left\{ \bar{d}_i - r_f + C - \frac{1}{2}V_{d,i} - \frac{1}{2}V_m - \frac{1}{2}\beta_{i,t+1}^2 V_m + \rho_i A_{i,t+1} + \rho_i B_i (b_i + a_i (b_{i,t} - b_i)) + a_i^2 \rho_i F_i \Sigma_{i,t} + \rho_i F_i V_{\beta,i} + \frac{1}{2}(\rho_i \phi_i B_i)^2 V_y + \frac{1}{2}V_{d,i} + \frac{1}{2}(\beta_{i,t+1} - 1)^2 V_m \right\} \right] \\ &= \exp \left\{ \bar{d}_i - r_f + C + \rho_i A_{i,t+1} + \rho_i B_i (b_i + a_i (b_{i,t} - b_i)) + \rho_i F_i V_{\beta,i} + a_i^2 \rho_i F_i \Sigma_{i,t} + \frac{1}{2}(\rho_i \phi_i B_i)^2 V_y - b_{i,t} V_m + \frac{1}{2}V_m^2 \Sigma_{i,t} \right\} \end{aligned} \quad (\text{A16})$$

Matching terms with the conjecture, we have

$$A_{i,t} = \bar{d}_i - r_f + C + \rho_i A_{i,t+1} + \rho_i B_i b_i (1 - a_i) + \rho_i F_i V_{\beta,i} + \frac{1}{2}(\rho_i \phi_i B_i)^2 V_y \quad (\text{A17})$$

$$B_i = \rho_i B_i a_i - V_m \Rightarrow B_i = -\frac{1}{1-a_i \rho_i} V_m \quad (\text{A18})$$

$$F_i = \frac{1}{2}V_m^2 + a_i^2 \rho_i F_i \Rightarrow F_i = \frac{1}{2(1-a_i^2 \rho_i)} V_m^2 \quad (\text{A19})$$

Using arguments analogous to those in the previous proposition, the transversality condition implies that $A_{i,t} = \sum_{s=t}^{\infty} \rho_i^{s-t} \chi_s$, for $\chi_s = \bar{d}_i - r_f + C + \rho_i B_i b_i (1 - a_i) + \rho_i F_i V_{\beta,i} + \frac{1}{2}(\rho_i \phi_i B_i)^2 V_y$, since χ_s is bounded. Given the approximation above, returns are given by

$$\begin{aligned} R_{i,t+1} &= \exp \left\{ \bar{d}_i + \beta_{i,t+1} m_{t+1} + d_{i,t+1} - \frac{1}{2}\beta_{i,t+1}^2 V_m - \frac{1}{2}V_{d,i} + C + \rho_i (A_{i,t+1} + B_{i,t+1} b_{i,t+1} + F_i \Sigma_{i,t+1}) - (A_{i,t} + B_{i,t} b_{i,t} + F_i \Sigma_{i,t}) \right\} \\ &= \exp \left\{ \alpha_{i,t} + \beta_{i,t+1} m_{t+1} + d_{i,t+1} - \frac{1}{2}\beta_{i,t+1}^2 V_m - \frac{1}{2}V_{d,i} + \gamma_i b_{i,t+1} + \rho_i F_i \Sigma_{i,t+1} \right\} \end{aligned}$$

where $\gamma_i = -\frac{\rho_i}{1-\rho_i a_i} V_m = \rho_i B_i$, and

$$\alpha_{i,t} = r_f - \frac{1}{2}(\rho_i B_i \phi_i)^2 V_y - \rho_i B_i b_i (1 - a_i) - \rho_i F_i V_{\beta,i} - B_i b_{i,t} - F_i \Sigma_{i,t}. \quad (\text{A20})$$

Finally, using the same first-order Taylor expansion as above, we have that expected returns are given by

$$\begin{aligned} \mathbb{E}_t [R_{i,t+1}] &= \mathbb{E}_t \left[\exp \left\{ \alpha_{i,t} + \beta_{i,t+1} m_{t+1} + d_{i,t+1} - \frac{1}{2}\beta_{i,t+1}^2 V_m - \frac{1}{2}V_{d,i} + \gamma_i b_{i,t+1} + \rho_i F_i \Sigma_{i,t+1} \right\} \right] \\ &\approx \mathbb{E}_t \left[\exp \left\{ \alpha_{i,t} + \beta_{i,t+1} m_{t+1} + d_{i,t+1} - \frac{1}{2}\beta_{i,t+1}^2 V_m - \frac{1}{2}V_{d,i} + \rho_i F_i (V_{\beta,i} + a_i^2 \Sigma_{i,t}) + \frac{1}{2}(\rho_i B_i \phi_i)^2 V_y + \rho_i B_i b_i (1 - a_i) + \rho_i a_i B_i b_{i,t} \right\} \right] \\ &= \exp \left\{ \alpha_{i,t} + \rho_i F_i (V_{\beta,i} + a_i^2 \Sigma_{i,t}) + \frac{1}{2}(\rho_i B_i \phi_i)^2 V_y + \rho_i B_i b_i (1 - a_i) + \rho_i a_i B_i b_{i,t} \right\} \\ &= \exp \left\{ r_f + b_{i,t} V_m - \frac{1}{2}V_m^2 \Sigma_{i,t} \right\}. \end{aligned}$$

This concludes the proof. ■

Proof of Corollary 5. The log-linearization and Taylor expansion implies that for $x \in \{m_{t+1}^2, d_{i,t+1}^2, m_{t+1}d_{i,t+1}\}$, the return is given by:

$$R_{i,t+1} = \exp \left\{ \alpha_{i,t} + \beta_{i,t+1}m_{t+1} + d_{i,t+1} - \frac{1}{2}\beta_{i,t+1}^2V_m - \frac{1}{2}V_{d,i} + \gamma_i b_{i,t+1} + \rho_i F_i \Sigma_{i,t+1} + O(x) \right\}$$

and so the idiosyncratic volatility of (log) returns can be approximated as follows:

$$\begin{aligned} \sigma_{i,t}^2 &= \text{var}_t \left(\log(R_{i,t+1}) - \left(\beta_{i,t+1}m_{i,t+1} - \frac{1}{2}\beta_{i,t+1}^2V_m \right) \right) \\ &\approx \text{var}_t (d_{i,t+1} + \gamma_i b_{i,t+1} + \rho_i F_i \Sigma_{i,t+1}) \end{aligned} \quad (\text{A21})$$

$$= \text{var}_t \left(d_{i,t+1} + \gamma_i \left(a_i \lambda_{i,t+1} \left(\beta_{i,t+1} - b_{i,t} + \frac{d_{i,t+1}}{m_{t+1}} \right) + \phi_i y_{t+1} \right) + \rho_i F_i \Sigma_{i,t+1} \right) \quad (\text{A22})$$

$$\begin{aligned} &= \mathbb{E}_t \left[\text{var}_t \left(d_{i,t+1} + \gamma_i \left(a_i \lambda_{i,t+1} \left(\beta_{i,t+1} - b_{i,t} + \frac{d_{i,t+1}}{m_{t+1}} \right) + \phi_i y_{t+1} \right) + \rho_i F_i \Sigma_{i,t+1} \mid m_{t+1} \right) \right] \\ &\quad + \text{var}_t \left(\mathbb{E}_t \left[d_{i,t+1} + \gamma_i \left(a_i \lambda_{i,t+1} \left(\beta_{i,t+1} - b_{i,t} + \frac{d_{i,t+1}}{m_{t+1}} \right) + \phi_i y_{t+1} \right) + \rho_i F_i \Sigma_{i,t+1} \mid m_{t+1} \right] \right) \end{aligned} \quad (\text{A23})$$

$$= \mathbb{E}_t \left[\gamma_i^2 a_i^2 \lambda_{i,t+1}^2 \Sigma_{i,t} + \gamma_i^2 \phi_i^2 V_y + \left(1 + \frac{\lambda_{i,t+1}}{m_{t+1}} \gamma_i a_i \right)^2 V_{d,i} \right] + \text{var}_t (\rho_i F_i \Sigma_{i,t+1}) \quad (\text{A24})$$

$$= \gamma_i^2 \phi_i^2 V_y + \gamma_i^2 a_i^2 \Sigma_{i,t} \mathbb{E}_t [\lambda_{i,t+1}^2] + V_{d,i} \mathbb{E}_t \left[\left(1 + \frac{\lambda_{i,t+1}}{m_{t+1}} \gamma_i a_i \right)^2 \right] + \rho_i^2 F_i^2 \text{var}_t (\Sigma_{i,t+1}) \quad (\text{A25})$$

which gives us the result. ■

Appendix B—Alternate specification of the convexity adjustment

We shall repeatedly use the result that for a normally distributed random variable $X \sim N(\mu, \sigma^2)$, the expected value of the exponential quadratic form is given by

$$\mathbb{E} \left[\exp \{ aX^2 + bX + c \} \right] = \frac{1}{\sqrt{1-2a\sigma^2}} \exp \left\{ c - \frac{1}{2} \frac{\mu^2}{\sigma^2} + \frac{1}{2} \frac{(b\sigma^2 + \mu)^2}{\sigma^2(1-2a\sigma^2)} \right\} \quad (\text{A26})$$

Suppose, as in the benchmark case, the pricing kernel is given by

$$M_{t+1} = M_t \exp \left\{ -r_f - \frac{1}{2} V_m - m_{t+1} \right\} \quad \text{where } m_{t+1} \sim \mathcal{N}(0, V_m), \quad (\text{A27})$$

and that investors' beliefs about firm i 's dividends (or cash flows), at date $t+1$ are given by:

$$D_{i,t+1} = D_{i,t} \exp \left\{ \bar{d}_i + \beta_{i,t+1} m_{t+1} + d_{i,t+1} - \frac{1}{2} \left(V_{d,i} + \frac{b_i^2 V_m}{1-V_m V_{\beta,i}} \right) - \log \left(\frac{1}{\sqrt{1-V_m V_{\beta,i}}} \right) \right\}, \quad (\text{A28})$$

where

$$d_{i,t+1} \sim \mathcal{N}(0, V_{d,i}), \quad \text{and } \beta_{i,t+1} \sim \mathcal{N}(b_i, V_{\beta,i}), \quad (\text{A29})$$

and we assume that

$$1 - V_{\beta,i} V_m > 0. \quad (\text{A30})$$

This condition ensures that the expectation of the exponential of the product of $\beta_{i,t+1}$ and m_{t+1} (i.e., $\mathbb{E}[\exp\{\beta_{i,t+1} m_{t+1}\}]$) is well defined. Moreover, note that

$$\mathbb{E} \left[\exp \{ \bar{d}_i + \beta_{i,t+1} m_{t+1} + d_{i,t+1} \} \right] = \mathbb{E} \left[\mathbb{E} \left[\exp \{ \bar{d}_i + \beta_{i,t+1} m_{t+1} + d_{i,t+1} \} \mid \beta_{i,t+1} \right] \right] \quad (\text{A31})$$

$$= \mathbb{E} \left[\exp \left\{ \bar{d}_i + \frac{1}{2} V_{d,i} + \frac{1}{2} \beta_{i,t+1}^2 V_m \right\} \right] \quad (\text{A32})$$

$$= \frac{1}{\sqrt{1 - V_m V_{\beta,i}}} \exp \left\{ \bar{d}_i + \frac{1}{2} \left(V_{d,i} + \frac{b_i^2 V_m}{1 - V_m V_{\beta,i}} \right) \right\}, \quad (\text{A33})$$

and so a convexity adjustment of

$$\begin{aligned} & \log \mathbb{E} \left[\exp \{ \bar{d}_i + \beta_{i,t+1} m_{t+1} + d_{i,t+1} \} \right] - \mathbb{E} \left[\log \left(\exp \{ \bar{d}_i + \beta_{i,t+1} m_{t+1} + d_{i,t+1} \} \right) \right] \\ &= \frac{1}{2} \left(V_{d,i} + \frac{b_i^2 V_m}{1 - V_m V_{\beta,i}} \right) + \log \left(\frac{1}{\sqrt{1 - V_m V_{\beta,i}}} \right) \end{aligned} \quad (\text{A34})$$

ensures that the log expected dividend growth (i.e., $\log \left(\mathbb{E} \left[\frac{D_{i,t+1}}{D_{i,t}} \right] \right)$) is given by \bar{d}_i . As in the benchmark model, $\frac{M_{t+1}}{M_t} \frac{D_{i,t+1}}{D_{i,t}}$ and $\frac{M_{t+2}}{M_{t+1}} \frac{D_{i,t+2}}{D_{i,t+1}}$ are independent, and the price-dividend ratio of next period's dividend is given by

$$p_t = \mathbb{E}_t \left[\frac{M_{t+1}}{M_t} \frac{D_{i,t+1}}{D_{i,t}} \right] = \mathbb{E}_t \left[\mathbb{E}_t \left[\frac{M_{t+1}}{M_t} \frac{D_{i,t+1}}{D_{i,t}} \mid \beta_{i,t+1} \right] \right] \quad (\text{A35})$$

$$= \mathbb{E}_t \left[\exp \left\{ \bar{d}_i - r_f + \frac{1}{2} (\beta_{i,t+1} - 1)^2 V_m - \frac{1}{2} \frac{b_i^2 V_m}{1 - V_m V_{\beta,i}} - \log \left(\frac{1}{\sqrt{1 - V_m V_{\beta,i}}} \right) \right\} \right] \quad (\text{A36})$$

$$= \exp \left\{ \bar{d}_i - r_f - \frac{b_i V_m}{1 - V_{\beta,i} V_m} + \frac{1}{2} \frac{V_{\beta,i} V_m^2}{1 - V_{\beta,i} V_m} \right\} \equiv p. \quad (\text{A37})$$

This implies that $R_{i,t+1} = \frac{D_{i,t+1}}{D_{i,t}} \frac{1+PD_{i,t+1}}{PD_{i,t}} = \frac{D_{i,t+1}}{D_{i,t}} \frac{1}{p}$, which, in turn, implies

$$\mathbb{E}[R_{i,t+1}] = \frac{1}{p} \mathbb{E}\left[\frac{D_{i,t+1}}{D_{i,t}}\right] = \exp\left\{r_f + \frac{b_i V_m}{1-V_{\beta,i} V_m} - \frac{1}{2} \frac{V_{\beta,i} V_m^2}{1-V_{\beta,i} V_m}\right\}. \quad (\text{A38})$$

Note that when $1 - V_{\beta,i} V_m \approx 1$, the above expression for expected returns corresponds to the one described in Proposition 1. In general, when the product of the variances is not small, the effect of $V_{\beta,i}$ on the expected return is larger than in the benchmark specification (since $\frac{1}{1-V_{\beta,i} V_m} > 1$).

Appendix C—Variable Definitions

The control variables used in the empirical analysis of Section 4.1 are computed using the monthly CRSP (lower-case variables) and the annual Compustat (upper-case variables) databases as follows:

1. Factor loadings on the excess market return, the SMB, HML, and UMD portfolios, and aggregate volatility risk (denoted by b_i , $b_{i,SMB}$, $b_{i,HML}$, $b_{i,UMD}$ and $b_{i,\Delta VIX}$, respectively) are estimated using monthly returns over sixty month rolling window regressions on the relevant factors. Idiosyncratic volatility (denoted by σ_i^2) is calculated relative to the four-factor model (i.e., relative to excess market returns, SMB, HML, and UMD) using sixty month rolling window regressions.
2. Log market value of equity, turnover, and bid-ask spread (denoted by $\log(mve)$, $turn$, and $spread$, respectively) are calculated using the CRSP database, as follows: $\log(mve) = \log(shrout \times abs(prc))$, $turn = \log(vol/shrout)$, and $spread = \frac{bid-ask}{(bid+ask)/2}$. The market value of equity and spread measures are lagged by one month when included in the regressions to avoid a mechanical effect on expected returns.
3. Market-to-book ratio, debt-to-equity ratio, volatility of earnings, return on assets, and accruals (denoted by M/B , D/E , $vol(E)$, ROA , and ACC , respectively) are calculated using CRSP and Compustat data as follows: $M/B = \frac{shrout \times abs(prc)}{AT-LT}$, $D/E = \frac{DLTT+DLC}{shrout \times abs(prc)}$, $ROA = \frac{OIADP}{(AT+lag(AT))/2}$, $ACC = \frac{(\Delta ACT - \Delta CHE) - (\Delta LCT - \Delta DLC - \Delta TXP) - DP}{(AT+lag(AT))/2}$, and $vol(E)$ is the volatility of earnings (i.e., $\frac{OIADP}{(AT+lag(AT))/2}$) over the previous five years.

References

- Adrian, T., and F. Franzoni. 2009. Learning about beta: time-varying factor loadings, expected returns, and the conditional CAPM. *Journal of Empirical Finance* 16(4):537–56.
- Andersen, T., T. Bollerslev, F. Diebold, and J. Wu. 2005. A framework for exploring the macroeconomic determinants of systematic risk. *American Economic Review* 95(2):398–404.
- Anderson, E., E. Ghysels, and J. Juergens. 2005. Do heterogeneous beliefs matter for asset pricing? *Review of Financial Studies* 18(3):875–924.
- Ang, A., and J. Chen. 2007. CAPM over the long run: 1926–2001. *Journal of Empirical Finance* 14(1):1–40.
- Ang, A., R. Hodrick, Y. Xing, and X. Zhang. 2006. The cross-section of volatility and expected returns. *Journal of Finance* 61(1):259–99.
- Ang, A., R. Hodrick, Y. Xing, and X. Zhang. 2009. High idiosyncratic volatility and low returns: international and further U.S. evidence. *Journal of Financial Economics* 91(1):1–23.
- Baiman, S., and R. Verrecchia. 1996. The relation among capital markets, financial disclosure, production efficiency, and insider trading. *Journal of Accounting Research* 34(1):1–22.
- Banerjee, S. 2011. Learning from prices and the dispersion in beliefs. *Review of Financial Studies* 24(9):3025–68.
- Barth, M., Y. Konchitchki, and W. Landsman. 2011. Cost of capital and financial statements transparency. Working Paper, Stanford University.
- Beaver, W., P. Kettler, and M. Scholes. 1970. The association between market determined and accounting determined risk measures. *Accounting Review* 45(4):654–82.
- Ben-Rephael, A., O. Kadan, and A. Wohl. 2010. The diminishing liquidity premium. Working Paper.
- Boehme, R., B. Danielsen, P. Kumar, and S. Sorescu. 2009. Idiosyncratic risk and the cross-section of stock returns: Merton (1987) meets Miller (1977). *Journal of Financial Markets* 12(3):438–68.
- Botosan, C.. 1997. Disclosure level and the cost of equity capital. *Accounting Review* 72(3):323–49.
- Botosan, C., and M. Plumlee. 2002. A re-examination of disclosure level and the expected cost of equity capital. *Journal of Accounting Research* 40(1):21–40.
- Bowman, R.. 1979. The theoretical relationship between systematic risk and financial (accounting) variables. *Journal of Finance* 34(3):617–30.
- Campbell, J., M. Lettau, B. Malkiel, and Y. Xu. 2001. Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *Journal of Finance* 56(1):1–43.
- Campbell, J., and R. Shiller. 1988. The dividend-price ratio and expectations of future divi-

- dends and discount factors. *Review of Financial Studies* 1(3):195–228.
- Chordia, T., A. Goyal, and J. Shanken. 2011. Cross-sectional asset pricing with individual stocks. Working Paper, Emory University.
- Cochrane, J., F. Longstaff, and P. Santa-Clara. 2008. Two trees. *Review of Financial Studies* 21(1):347–385.
- Cooley, T., and E. Prescott. 1976. Estimation in the presence of stochastic parameter variation. *Econometrica* 44(1):167–84.
- Core, J., W. Guay, and R. Verdi. 2008. Is accruals quality a priced risk factor? *Journal of Accounting and Economics* 46(1):2–22.
- David, A. 1997. Fluctuating confidence in stock markets: Implications for returns and volatility. *Journal of Financial and Quantitative Analysis* 32(4):427–62.
- Diamond, D., and R. Verrecchia. 1991. Disclosure, liquidity, and the cost of capital. *Journal of Finance* 46(4):1325–59.
- Diether, K., C. Malloy, and A. Scherbina. 2002. Differences of opinion and the cross section of stock returns. *Journal of Finance* 57(5):2113–41.
- Duarte, J., A. Kamara, S. Siegel, and C. Sun. 2011. The common component of idiosyncratic volatility. Working Paper, University of Washington.
- Duarte, J., and L. Young. 2009. Why is PIN priced? *Journal of Financial Economics* 91(2):119–38.
- Easley, D., and M. O’Hara. 2004. Information and the cost of capital. *Journal of Finance* 59(4):1553–83.
- Fama, E., and K. French. 1997. Industry costs of equity. *Journal of Financial Economics* 43(2):153–93.
- Francis, J., R. LaFond, P. Olsson, and K. Schipper. 2004. Costs of equity and earnings attributes. *Accounting Review* 79(4):967–1010.
- Francis, J., R. LaFond, P. Olsson, and K. Schipper. 2005. The market pricing of accruals quality. *Journal of Accounting and Economics* 39(2):295–327.
- Francis, J., D. Nanda, and P. Olsson. 2008. Voluntary disclosure, information quality, and costs of capital. *Journal of Accounting Research* 46(1):53–99.
- Frazzini, A., and L. Pedersen. 2011. Betting against beta. NBER Working Paper No. 16601.
- Fu, F. 2009. Idiosyncratic risk and the cross-section of expected stock returns. *Journal of Financial Economics* 91(1):24–37.
- Hellwig, M. 1980. On the aggregation of information in competitive markets. *Journal of Economic Theory* 22(3):477–98.
- Huang, W., Q. Liu, S. Rhee, and L. Zhang. 2010. Return reversals, idiosyncratic risk, and expected returns. *Review of Financial Studies* 23(1):147–68.

- Hughes, J., J. Liu, and J. Liu. 2007. Information asymmetry, diversification, and cost of capital. *Accounting Review* 82(3):705–29.
- Jagannathan, R., and Z. Wang. 1996. The conditional CAPM and the cross-section of expected returns. *Journal of Finance* 51(1):3–53.
- Johnson, T. 2004. Forecast dispersion and the cross section of expected returns. *Journal of Finance* 59(5):1957–78.
- Jostova, G., and A. Philipov. 2005. Bayesian analysis of stochastic betas. *Journal of Financial and Quantitative Analysis* 40(4):747–78.
- Kumar, P., S. Sorescu, R. Boehme, and B. Danielsen. 2008. Estimation risk, information, and the conditional CAPM: theory and evidence. *Review of Financial Studies* 21(3): 1037–85.
- Lambert, R., C. Leuz, and R. Verrecchia. 2007. Accounting information, disclosure, and the cost of capital. *Journal of Accounting Research* 45(2):385–420.
- Lehmann, B. 1990. Residual risk revisited. *Journal of Econometrics* 45(1–2):71–97.
- Lettau, M., and S. Ludvigson. 2001. Resurrecting the (C) CAPM: a cross-sectional test when risk premia are time-varying. *Journal of Political Economy* 109(6):1238–87.
- Malkiel, B., and Y. Xu. 2006. Idiosyncratic risk and security returns. Working Paper, University of Texas at Dallas.
- Martin, I.. 2011. The Lucas orchard. NBER Working Paper No. 17563.
- Menzly, L., T. Santos, and P. Veronesi. 2004. Understanding predictability. *Journal of Political Economy* 112(1):1–47.
- Pástor, L., and P. Veronesi. 2003. Stock valuation and learning about profitability. *Journal of Finance* 58(5):1749–89.
- Qu, S., L. Starks, and H. Yan. 2004. Risk, dispersion of analyst forecasts and stock returns. Working Paper, University of Texas at Austin.
- Ritter, J. 1991. The long-run performance of initial public offerings. *Journal of Finance* 46(1):3–27.
- Santos, T., and P. Veronesi. 2006. Labor income and predictable stock returns. *Review of Financial Studies* 19(1):1–44.
- Shanken, J.. 1992. On the estimation of beta-pricing models. *Review of Financial Studies* 5(1):1–55.
- Spiegel, M., and X. Wang. 2010. Cross-sectional variation in stock returns: liquidity and idiosyncratic risk. Working Paper, Yale University.
- Teoh, S., I. Welch, and T. Wong. 1998. Earnings management and the long-run market performance of initial public offerings. *Journal of Finance* 53(6):1935–74.
- Veronesi, P. 1999. Stock market overreactions to bad news in good times: a rational expectations equilibrium model. *Review of Financial Studies* 12(5):975–1007.

Xia, Y. 2001. Learning about predictability: The effects of parameter uncertainty on dynamic asset allocation. *Journal of Finance* 56(1):205–46.

Figure 1

The effect of a mean-preserving spread in $\beta_{i,t+1}$ across states of the world

This figure plots the pricing kernel M_{t+1}/M_t and the present value of the dividend $\frac{M_{t+1}}{M_t}D_{i,t+1}$ as functions of the state of the world (i.e., m_{t+1}) for different values of $\beta_{i,t+1}$. For the bottom panel, the solid line corresponds to $\beta_{i,t+1} = 1$, the dashed line corresponds to $\beta_{i,t+1} = 0.9$, and the dotted line corresponds to $\beta_{i,t+1} = 1.1$. The mean dividend growth rate, \bar{d}_i , and the idiosyncratic shocks, $d_{i,t+1}$, are both set to zero; the risk-free rate r_f is set to 2%; the variance in the pricing kernel V_m is set to 0.2; and $D_{i,t}$ is normalized to 1.

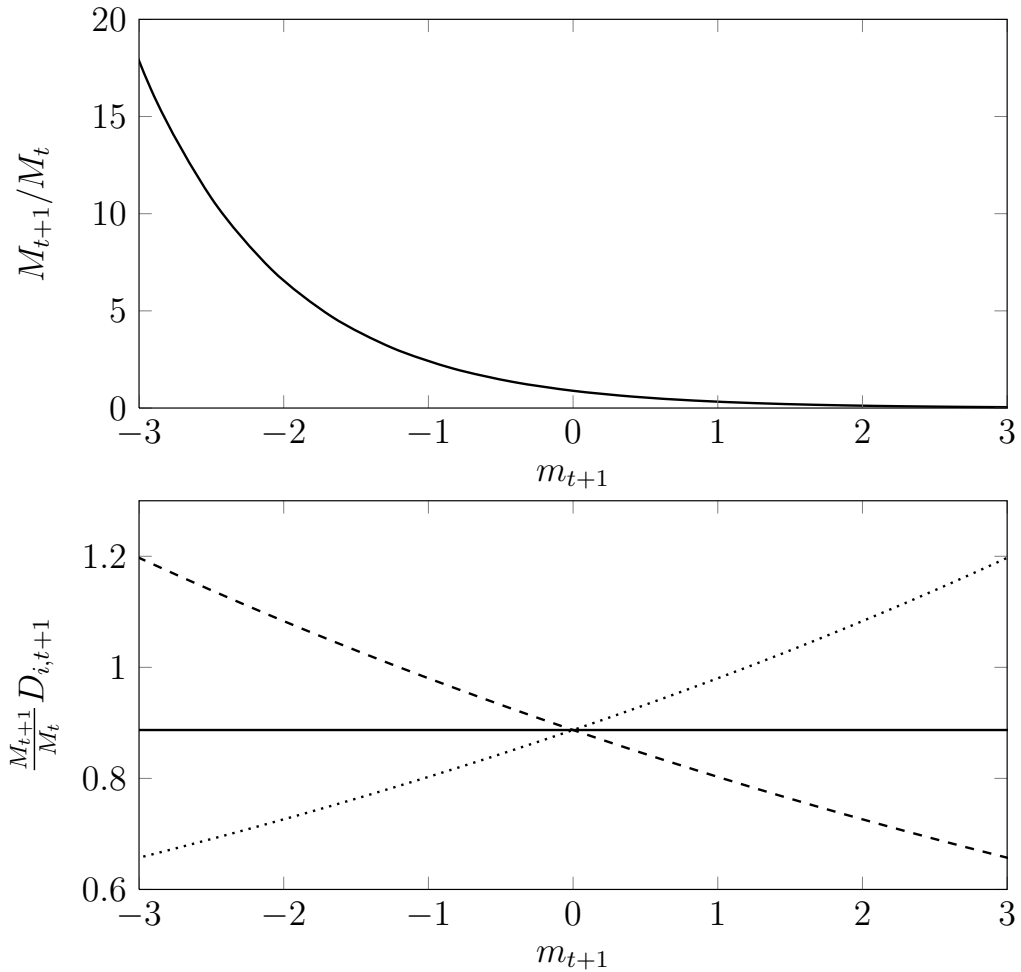


Figure 2

Log expected returns as a function of uncertainty about factor loadings $V_{\beta,i}$

This figure plots the (log) expected return on stock i (i.e., $\log(\mathbb{E}[R_{i,t+1}])$) as a function of the uncertainty about its factor loadings, $V_{\beta,i}$, for three different levels of persistence, a_i , in factor loadings. The dotted line plots the i.i.d. benchmark presented in Proposition 1, the dashed line plots the no-information benchmark of Corollary 3, and the solid line plots the full-information benchmark of Corollary 4. The risk-free rate r_f is set to 2%, the variance in the pricing kernel, V_m , is set to 0.2, the unconditional mean factor loading, b_i , is set to 1, and the long-run mean price-dividend ratio is set to $PD_{i,t} = 4$, which implies $\rho_i = 0.80$.

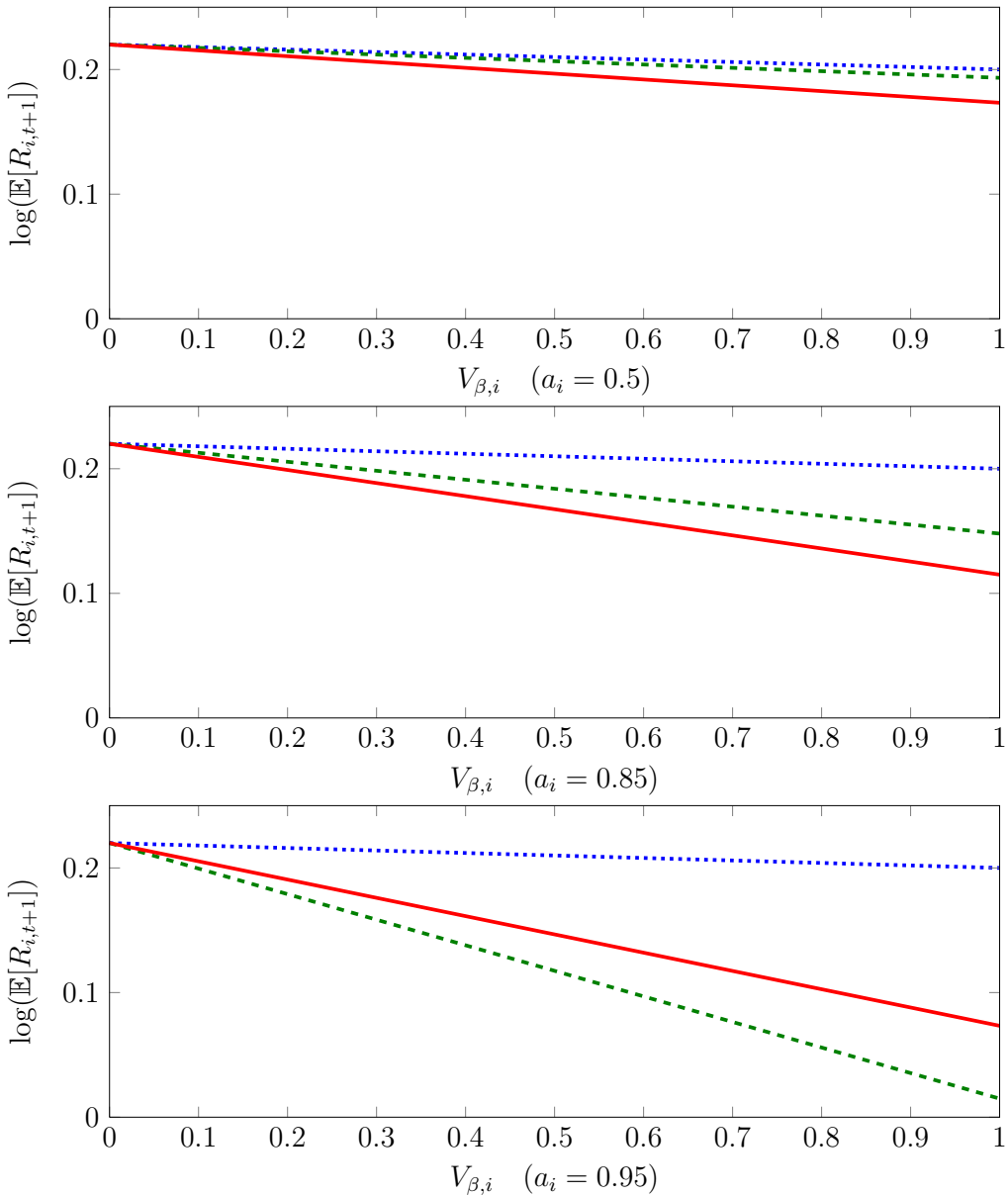


Figure 3
Factor-loading uncertainty and persistence pairs

This figure plots parameter values of $V_{\beta,i}$ and a_i such that the model-implied effect of factor-loading uncertainty on annualized expected returns is 450 basis points. The solid line corresponds to the no-information benchmark of Corollary 3, and the dash-dotted, dashed, and dotted lines correspond to the full-information benchmark of Corollary 4, for ρ_i equal to 0.65, 0.8, and 0.95, respectively. The aggregate risk premium (i.e., V_m) is set to 0.2.

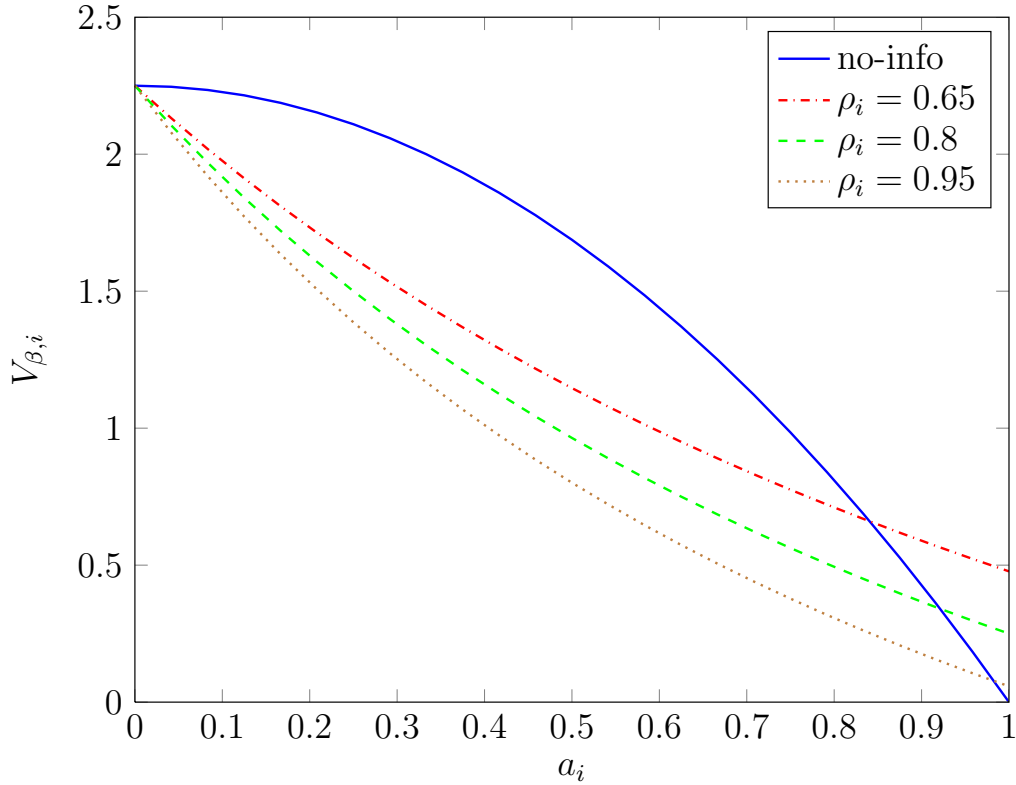


Table 1
Summary statistics

This table provides summary statistics for our sample, which consists of monthly observations from January 1964 to December 2011. The variables include the return on stocks from the CRSP database (i.e., $R_{i,t+1}$); excess log returns (i.e., $r_{i,t+1} - r_{f,t}$); estimates of factor loadings on the excess market return, the SMB, HML, and UMD portfolios, and aggregate volatility risk (denoted by b_i , $b_{i,SMB}$, $b_{i,HML}$, $b_{i,UMD}$, and $b_{i,\Delta VIX}$, respectively); and our proxy for uncertainty about the factor loading on excess market returns (i.e., $V_{\beta,i}$). The table also reports estimates of the following firm-level characteristics computed from the CRSP and Compustat databases: log market value of equity, turnover, market-to-book ratio, debt-to-equity ratio, volatility of earnings, return on assets, accruals, bid-ask spread, and idiosyncratic volatility relative to the four-factor model (denoted by $\log(mve)$, $turn$, M/B , D/E , $vol(E)$, ROA , ACC , $spread$, and σ_i^2 , respectively). We also report summary statistics for the cross-section of autocorrelation coefficients of estimated factor loadings (i.e., $AC(b_i)$). Some variables are available only for a subsample of our full sample period (e.g., the VXO series used to compute $b_{i,\Delta VIX}$ begins in 1986, and the bid and ask series in CRSP used to compute $spread$ begins in 1983). The table also reports the pairwise correlation coefficients for b_i , $V_{\beta,i}$, $vol(E)$, and σ_i^2 in our sample.

Variable	Nobs	Mean	Median	Std. Dev	Min	5th Perc	95th Perc	Max
$R_{i,t+1}$	1077532	0.0094	0.0057	0.122	-0.6765	-0.1765	0.2035	1.933
$r_{i,t+1} - r_{f,t}$	1077532	-0.0014	0.0022	0.1212	-1.134	-0.1977	0.1819	1.071
b_i	1077532	1.133	1.062	0.7005	-5.334	0.1491	2.375	6.613
$V_{\beta,i}$	1077532	0.1268	0.0678	0.1687	0.0001	0.012	0.4426	4.274
$b_{i,SMB}$	1077532	0.5895	0.4493	0.8492	-5.365	-0.5374	2.168	7.962
$b_{i,HML}$	1077532	0.1825	0.2391	0.9179	-9.476	-1.366	1.547	8.729
$b_{i,UMD}$	1077532	-0.1128	-0.0832	0.5117	-6.475	-0.9666	0.6555	4.933
$b_{i,\Delta VIX}$	802243	-0.024	-0.0514	0.5704	-1.767	-0.9296	1.01	2.196
$\log(mve)$	1067498	13.04	12.98	1.816	4.8	10.17	16.18	20.22
$turn$	1077384	-0.4992	-0.5207	1.141	-6.35	-2.326	1.37	6.37
M/B	941577	0.8839	0.5912	1.235	-0.1828	0.119	2.34	16.09
D/E	941021	0.8816	0.3015	1.948	0.0	0.0	3.518	22.59
$vol(E)$	748015	0.0797	0.0893	0.1171	-0.6083	-0.1217	0.2376	0.382
ROA	447692	0.0389	0.025	0.0528	0.0	0.005	0.1159	2.497
ACC	375597	-0.0348	-0.0354	0.0792	-1.721	-0.1436	0.0838	1.155
$spread$	698170	0.013	0.0051	0.0251	-0.1684	0.0002	0.05	1.98
σ_i^2	1077532	0.0122	0.0069	0.015	0.0	0.0012	0.042	0.3216
$AC(b_i)$	24294	0.5302	0.6322	0.3683	-0.7558	-0.1402	0.9576	0.9975

	$V_{\beta,i}$	$vol(E)$	σ_i^2
b_i	0.295	0.207	0.345
$V_{\beta,i}$		0.419	0.804
$vol(E)$			0.454

Table 2

The cross-section of expected returns controlling for factor loadings

This table presents the results from a Fama-MacBeth estimation of the following cross-sectional regression:

$$r_{i,t+1} - r_{f,t} = \alpha_{i,t} + \lambda_t b_{i,t} + \gamma_t V_{\beta,i,t} + \kappa_t b_{i,t}^2 + \text{controls}_{i,t} + \varepsilon_{i,t+1},$$

where $b_{i,t}$ and $V_{\beta,i}$ are estimates from firm-specific first-stage, sixty month rolling window regressions of excess log returns $r_{i,t+1} - r_{f,t}$ on market excess (log) returns $r_{m,t+1} - r_{f,t}$. The controls include factor loadings on the SMB, HML, and UMD portfolios and aggregate volatility risk (denoted by $b_{i,SMB}$, $b_{i,HML}$, $b_{i,UMD}$, and $b_{i,\Delta VIX}$, respectively). The t -statistics (reported below the coefficient estimates) are computed using Newey-West standard errors using sixty lags, and are adjusted for the Shanken correction. The sample consists of a maximum 575 monthly cross-sections from January 1964 through December 2011 and the mean and median adjusted R^2 's, and the average number of firms in each cross-section are reported. The estimates of $b_{i,\Delta VIX}$ are available for a shorter sample since the data on the VXO index is only available beginning in January 1986.

Intercept	0.00578	0.00299	0.00461	0.00513	0.00484	0.00445	0.00599
	4.05	1.90	2.86	3.97	3.96	3.63	9.79
b_i	-0.00559	0.00004	0.00111	0.00068	0.00073	0.00116	0.00039
	-1.97	0.01	0.41	0.25	0.28	0.45	0.11
b_i^2		-0.00226	-0.00118	-0.00121	-0.00102	-0.00112	-0.00100
		-1.09	-0.59	-0.60	-0.50	-0.55	-0.36
$V_{\beta,i}$			-0.04897	-0.06220	-0.06455	-0.06313	-0.06305
			-3.11	-6.24	-6.31	-6.38	-4.67
$b_{i,SMB}$				0.00169	0.00160	0.00151	0.00016
				0.86	0.83	0.77	0.06
$b_{i,HML}$					-0.00002	0.00002	-0.00173
					-0.01	0.01	-0.69
$b_{i,UMD}$						0.00070	0.00048
						0.31	0.15
$b_{i,\Delta VIX}$							-0.00305
							-0.74
Mean R^2	6.1%	6.5%	8.5%	11.1%	13.6%	16.0%	17.6%
Med. R^2	3.0%	3.4%	5.5%	7.4%	9.9%	11.1%	12.1%
# months	575	575	575	575	575	575	310
# firms	1874	1874	1874	1874	1874	1874	2588

Table 3

The cross-section of expected returns controlling for factor loadings and characteristics

This table presents the results from a Fama-MacBeth estimation of the cross-sectional regression in equation (62) where $b_{i,t}$ and $V_{\beta,i}$ are estimates from firm-specific first-stage, sixty month rolling window regressions of excess log returns $r_{i,t+1} - r_{f,t}$ on market excess (log) returns $r_{m,t+1} - r_{f,t}$. The controls include factor loadings on the SMB, HML, and UMD portfolios (denoted by $b_{i,SMB}$, $b_{i,HML}$, and $b_{i,UMD}$, respectively) and firm characteristics $\log(mve)$, M/B , D/E , $vol(E)$, ROA , ACC , $spread$ (as described in Table 1). The t -statistics (reported below the coefficient estimates) are computed using Newey-West standard errors using sixty lags, and adjusted for the Shanken correction. The sample consists of a maximum of 575 monthly cross-sections from January 1964 through December 2011 and the mean and median adjusted R^2 's, and the average number of firms in each cross-section are reported.

Intercept	-0.00041 -0.09	0.02401 3.11	0.05734 6.13	0.05854 6.30	0.06174 6.90	0.05686 6.53	0.05860 6.40	0.05226 4.55
b_i	0.00038 0.15	-0.00738 -2.77	-0.00703 -2.43	-0.00735 -2.55	-0.00834 -2.80	-0.00844 -2.76	-0.01241 -2.78	-0.01202 -1.86
b_i^2	-0.00092 -0.47	0.00015 0.07	0.00008 0.04	0.00020 0.10	0.00055 0.27	0.00096 0.47	0.00264 1.07	0.00255 0.88
$V_{\beta,i}$	-0.05963 -6.43	-0.09157 -6.59	-0.09794 -6.42	-0.09641 -6.36	-0.09476 -5.89	-0.08347 -5.19	-0.08051 -3.59	-0.06551 -4.05
$b_{i,SMB}$	0.00160 0.85	0.00073 0.41	-0.00043 -0.25	-0.00052 -0.30	-0.00041 -0.24	-0.00055 -0.30	0.00011 0.05	-0.00097 -0.43
$b_{i,HML}$	-0.00008 -0.05	0.00062 0.34	0.00285 1.47	0.00305 1.58	0.00261 1.39	0.00258 1.37	0.00218 1.11	-0.00001 -0.01
$b_{i,UMD}$	0.00064 0.30	-0.00037 -0.17	-0.00237 -1.13	-0.00259 -1.22	-0.00240 -1.15	-0.00187 -0.89	-0.00266 -1.18	-0.00079 -0.26
$\log(mve)$	0.00037 1.01	0.00018 0.39	-0.00169 -3.37	-0.00175 -3.51	-0.00196 -4.03	-0.00171 -3.79	-0.00184 -3.85	-0.00223 -5.31
$turn$		0.01079 6.90	0.01134 7.65	0.01145 7.77	0.01150 8.02	0.01003 5.98	0.01001 5.99	0.00810 3.05
M/B			-0.00918 -9.62	-0.00873 -9.15	-0.00847 -8.40	-0.00695 -5.37	-0.00597 -4.91	-0.00652 -2.26
D/E				-0.00109 -3.96	-0.00140 -4.75	-0.00131 -3.75	-0.00188 -4.25	-0.00144 -1.90
ROA					0.00406 0.62	0.00886 1.09	0.01459 1.78	0.02606 2.74
$vol(E)$						-0.04130 -4.62	-0.03244 -3.61	-0.03791 -3.63
ACC							-0.01407 -1.85	-0.02185 -4.65
$spread$								0.01294 0.14
Mean R^2	16.3%	18.0%	18.7%	18.8%	18.9%	19.9%	20.2%	19.2%
Med. R^2	11.6%	13.6%	14.7%	14.8%	14.9%	16.0%	16.8%	14.3%
# months	575	575	575	575	575	575	572	290
# firms	1857	1856	1623	1600	743	732	605	412

Table 4

The relation between expected returns and idiosyncratic volatility

This table presents the results from a Fama-MacBeth estimation of the cross-sectional regression in equation (65). The variables PIV and RIV are the conditional expectation and residual, respectively, from firm-specific regressions of idiosyncratic volatility on $V_{\beta,i}$ and a constant (see equation (64)), and 1_{HAC} is an indicator variable for firms whose autocorrelation in estimated factor loadings is in the top 25th percentile. All other variables are defined as in Table 3. The t -statistics (reported below the coefficient estimates) are computed using Newey-West standard errors using sixty lags, and adjusted for the Shanken correction. The sample consists of a maximum of 575 monthly cross-sections from January 1964 through December 2011 and the mean and median adjusted R^2 's, and the average number of firms in each cross-section are reported.

Intercept	0.00452	0.00454	0.00262	0.02783	0.06315	0.06454	0.06727	0.06258	0.06751	0.06138
	3.51	3.52	0.61	3.56	6.59	6.79	7.33	7.07	7.32	4.56
b_i	0.00130	0.00125	0.00065	-0.00692	-0.00662	-0.00698	-0.00793	-0.00823	-0.01498	-0.01147
	0.52	0.50	0.26	-2.62	-2.32	-2.45	-2.69	-2.70	-2.42	-1.85
b_i^2	-0.00111	-0.00109	-0.00095	0.00008	0.00005	0.00020	0.00054	0.001015	0.003829	0.002599
	-0.56	-0.55	-0.48	0.04	0.03	0.10	0.26	0.49	1.26	0.89
PIV	-0.61140	-0.61240	-0.59470	-0.93130	-1.03500	-1.02500	-1.00000	-0.9436	-1.005	-0.7829
	-7.66	-7.69	-7.87	-8.24	-8.08	-8.21	-7.24	-6.94	-4.98	-5.43
RIV	-0.03299	-0.17250	-0.16480	-0.36820	-0.39500	-0.37370	-0.37650	-0.3185	-0.5068	-0.2526
	-0.23	-1.31	-1.25	-2.24	-2.07	-1.95	-1.81	-1.39	-1.40	-0.97
$RIV1_{HAC}$		0.23360	0.25340	0.17900	0.20100	0.22100	0.22750	0.3457	0.6213	0.4986
		2.48	2.54	2.30	2.09	2.19	2.20	2.35	1.85	1.80
$b_{i,SMB}$	0.00122	0.00123	0.00115	-0.00002	-0.00126	-0.00133	-0.00124	-0.0008744	-0.000397	-0.001098
	0.65	0.66	0.63	-0.01	-0.75	-0.79	-0.73	-0.50	-0.22	-0.50
$b_{i,HML}$	-0.00014	-0.00014	-0.00026	0.00038	0.00257	0.00277	0.00237	0.002162	0.001826	-0.0006651
	-0.08	-0.08	-0.15	0.21	1.35	1.46	1.30	1.16	0.92	-0.26
$b_{i,UMD}$	0.00086	0.00085	0.00087	0.00001	-0.00200	-0.00223	-0.00205	-0.001682	-0.001716	-0.0003245
	0.41	0.40	0.41	0.00	-0.95	-1.06	-0.98	-0.80	-0.82	-0.11
$\log(mve)$			0.00015	-0.00008	-0.00208	-0.00215	-0.00232	-0.002058	-0.002288	-0.002777
			0.42	-0.17	-3.88	-4.04	-4.47	-4.28	-4.69	-5.10
$turn$				0.01082	0.01140	0.01151	0.01156	0.0101	0.009918	0.007921
				6.91	7.71	7.83	8.09	6.00	5.82	3.06
M/B					-0.00926	-0.00879	-0.00852	-0.006977	-0.006012	-0.006719
					-9.74	-9.17	-8.42	-5.39	-4.87	-2.25
D/E						-0.00113	-0.00143	-0.001323	-0.001826	-0.001488
						-4.10	-4.97	-3.87	-4.44	-2.01
ROA							0.00316	0.006372	0.01063	0.02073
							0.50	0.83	1.37	2.11
$vol(E)$								-0.03703	-0.02732	-0.0302
								-4.17	-3.15	-3.61
ACC									-0.01781	-0.0215
									-4.53	-4.30
$spread$										0.02665
										0.36
Mean R^2	16.2%	16.2%	16.5%	18.2%	18.8%	19.0%	19.0%	20.3%	20.8%	19.9%
Med. R^2	11.3%	11.2%	11.9%	13.8%	14.7%	14.9%	15.2%	16.4%	17.2%	15.9%
# months	575	575	575	575	575	575	575	575	571	290
# firms	1874	1874	1857	1856 ⁵⁰	1623	1600	1249	732.3	605	412.2

Table 5
Robustness: Using simple returns

This table presents the results from a Fama-MacBeth estimation of the following cross-sectional regression:

$$R_{i,t+1} - R_{f,t} = \alpha_{i,t} + \lambda_t b_{i,t} + \gamma_t V_{\beta,i,t} + \text{controls}_{i,t} + \varepsilon_{i,t+1},$$

where $b_{i,t}$ and $V_{\beta,i}$ are estimates from firm-specific first-stage, sixty month rolling window regressions of excess returns $R_{i,t+1} - R_{f,t}$ on market excess returns $R_{m,t+1} - R_{f,t}$. The controls include factor loadings on the SMB, HML, and UMD portfolios (denoted by $b_{i,SMB}$, $b_{i,HML}$, and $b_{i,UMD}$ respectively), and firm characteristics $\log(mve)$, M/B , D/E , $vol(E)$, ROA , ACC , $spread$ (as described in Table 1). The t -statistics (reported below the coefficient estimates) are computed using Newey-West standard errors using sixty lags, and adjusted for the Shanken correction. The sample consists of a maximum of 575 monthly cross-sections from January 1964 through December 2011 and the mean and median adjusted R^2 's, and the average number of firms in each cross-section are reported.

Intercept	0.00411	-0.00135	0.02246	0.05509	0.05639	0.06020	0.05618	0.05507	0.04620
	3.10	-0.27	2.92	5.81	6.00	6.79	6.46	5.80	4.79
b_i	0.00191	0.00167	-0.00434	-0.00413	-0.00415	-0.00432	-0.00329	-0.00311	-0.00220
	0.85	0.73	-1.92	-1.81	-1.82	-1.89	-1.41	-1.34	-0.65
$V_{\beta,i}$	-0.02108	-0.01878	-0.03838	-0.04429	-0.04331	-0.04216	-0.03239	-0.02652	-0.01833
	-3.31	-3.21	-3.88	-3.71	-3.64	-3.42	-2.82	-1.36	-2.37
$b_{i,SMB}$	0.00179	0.00198	0.00054	-0.00060	-0.00067	-0.00059	-0.00057	0.00031	-0.00036
	0.92	1.04	0.31	-0.35	-0.38	-0.33	-0.31	0.15	-0.15
$b_{i,HML}$	0.00042	0.00032	0.00117	0.00330	0.00345	0.00303	0.00290	0.00255	0.00055
	0.25	0.19	0.68	1.79	1.88	1.71	1.61	1.40	0.24
$b_{i,UMD}$	0.00025	0.00012	-0.00098	-0.00278	-0.00296	-0.00291	-0.00245	-0.00317	-0.00088
	0.12	0.06	-0.45	-1.28	-1.35	-1.35	-1.13	-1.38	-0.30
$\log(mve)$		0.00040	0.00027	-0.00157	-0.00164	-0.00192	-0.00173	-0.00184	-0.00212
		1.06	0.51	-2.78	-2.95	-3.48	-3.38	-3.51	-4.82
$turn$			0.01213	0.01281	0.01291	0.01302	0.01151	0.01144	0.01016
			7.92	8.97	9.11	9.70	7.13	7.34	3.74
M/B				-0.00903	-0.00867	-0.00849	-0.00707	-0.00613	-0.00639
				-9.60	-9.23	-8.54	-5.58	-5.24	-2.39
D/E					-0.00098	-0.00122	-0.00116	-0.00168	-0.00115
					-3.43	-4.17	-3.41	-3.75	-1.79
ROA						0.00363	0.00819	0.01384	0.02463
						0.56	1.02	1.73	2.59
$vol(E)$							-0.05077	-0.04216	-0.04877
							-5.27	-5.23	-4.31
ACC								-0.01491	-0.02238
								-2.21	-4.96
$spread$									0.06360
									0.72
Mean R^2	15.1%	15.5%	17.3%	18.0%	18.1%	18.3%	19.3%	19.5%	18.7%
Med. R^2	10.4%	11.0%	13.3%	13.9%	14.2%	14.6%	15.7%	15.7%	14.4%
# months	575	575	575	575	575	575	575	572	290
# firms	1874	1857	1856	1623	1600	1249	732.3	605	412.2