Extended Appendix

Companion to forthcoming article in M̂anagement Science titled "Can Noise Create Size and Value Effects?" based on an earlier version.

1 Noise

In this section, we discuss key assumptions and technical assumptions of the paper.

The following is the key assumption of the paper.

Assumption 1 *Every stock has a value V_t, which is determined by economic theory. The price P_t of a stock deviates from its value V_t by a noise Δ_t. Specifically,*

$$P_t = V_t \frac{e^{\Delta_t}}{\mathrm{E}[e^{\Delta_t}]}, \tag{1}$$

where Δ_t is independent of V_s for all t and s and $\mathrm{E}[e^{\Delta_t}]$ is the unconditional expectation of e^{Δ_t}. The dividend D_t of the stock is also independent of Δ_s, for all t and s.

In assumption 1, the theory that determines the value V_t is unspecified and can be consumption-based asset pricing models, CAPM, or APT, just to name a few. The value V_t is the price if there were no noise and has all the "nice" properties, for example, the expected return computed using V_t is determined by risk and thus the cross section of expected returns computed using V_t is determined by beta only if the asset pricing model is APT. For our purpose, it is not necessary to define how the market arrives at this value V_t. However, it may be convenient to think of the discounted cashflow valuation equation where $V_t = \mathrm{E}_t[\sum_{s=t}^{\infty} e^{-\mu(s-t)} D_s]$, where μ is the discount rate and D_s is the dividend at time s.

Assumption 1 implies that

$$\mathrm{E}[P_t|V_t] = V_t. \tag{2}$$

That is, the price for a stock is a noisy proxy for its value, which we assume is unobservable, and the price is, on average, right. The assumption on dividend D_t is necessary for drawing conclusion on returns since dividend D_{t+1} is part of the cashflow for $t+1$, in addition to the price P_{t+1}. Without loss of generality, we will assume that $\mathrm{E}[\Delta_t] = 0$.

Black (1986) also argues that there might be a difference between the price and the fair value of a stock but he does not present a form analysis. Summers (1986) assumes an additive form, $P_t = V_t + \Delta_t$. Summers asserts, "[This assumption of pricing noise] clearly captures Keynes's notion that markets are sometimes driven by animal spirits unrelated to economic activities. It, also, is consistent with the experimental evidence of Tversky and Kahneman that subjects overreact to new information in making probabilistic judgements. The formulation considered here [also] captures Robert Shiller's suggestion that financial markets display excess volatility and overreact to new information." We remark that the noise in Assumption 1 is specified in multiplicative form, which is used in Blume and Stambaugh (1983) and Fama and French (1988) (see also Hsu (2006)). The additive form of Summers (1986) implies that the noise becomes negligible over time as V_t grows, if Δ_t is stationary as Summers assumes. Aboody, Hughes, and Liu (2002) also assume an additive form. Campbell and Kyle (1993) recognize this problem and use an additive form with de-trended dividends. Such a problem does not arise from the multiplicative form.

Many of the qualitative results of the paper follows from this assumption. We will make more technical assumptions for quantitative results.

Assumption 2 *The noise satisfies,*

$$\Delta_{t+1} = \rho \Delta_t + \sigma_{\epsilon_\Delta} \epsilon_{\Delta t+1}, \tag{3}$$

where ϵ_t are independent standard normals.

When $\rho < 1$, Δ_t is mean-reverting and stationary. This implies that a noise Δ_{t+1} at time $t+1$, on average, will lead to smaller noise Δ_t at time t. The mean reversion of Δ_t towards zero captures the intuition that information is slowly impounded into prices. When $\rho = 0$, the noise is independent and identically distributed (IID). If $\rho = 1$, the noise Δ_{t+1} will be equal to Δ_t on average. In this case, the noise is infinitely persistent and price levels do not predict returns $E[R_{t+1}|P_0...P_t] = E[R_{t+1}]$.

Whether noise Δ_t is mean reverting or not is an empirical question. To avoid cubersome notations, the rest of the paper will assume that $\rho < 1$. Presumably, the market sets price P_t to be its best estimate of V_t, therefore P_t should revert towards value V_t, as new information becomes known. However, most of the derivation in the paper goes through with minor changes if $\rho = 1$.

Table 1: Summary of Parameters

μ	σ_r	σ_{ϵ_Δ}	ρ	$\bar{x}_v$	ρ_x	σ_{ϵ_x}
3%	30%	6%	0.5	4	0.9	10%

The calibration of these parameters are described in Section 1.

We assume that σ_{ϵ_Δ} is a constant. This assumption may be a little restrictive since σ_{ϵ_Δ} could be state dependent. For example, noise during economic expansions may have a different volatility from noise during recessions.

Similar specifications of the noise follow from Blume and Stambaugh (1983), Summers (1986), Fama and French (1988), Aboody, Hughes, and Liu (2002), Arnott and Hsu (2006), Hsu (2006), and Brennan and Wang (2006).

For ease of exposition, we denote the logarithm of V_t by v_t and logarithm of P_t by p_t,

$$V_t = e^{v_t}; \quad P_t = e^{p_t}. \tag{4}$$

Equation (1) can then be written as

$$p_t = v_t + \Delta_t - \ln(\mathrm{E}[e^{\Delta_t}]). \tag{5}$$

We call $\frac{V_{t+1}+D_{t+1}}{V_t}$ the *value* return R^v_{t+1}, which is dictated by some asset pricing model. We call $\frac{P_{t+1}+D_{t+1}}{P_t}$ the return R_{t+1}. We will use $d_t = \ln D_t$ to denote the logarithm of the time t dividend D_t. We make the following assumption on the value and the value-dividend ratio.

Assumption 3 *The value v_t is a random walk,*

$$v_{t+1} = \mu + v_t + \sigma_r \epsilon_{rt+1}. \tag{6}$$

The value-dividend ratio satisfies

$$v_{t+1} - d_{t+1} = (1-\rho_x)\bar{x}_v + \rho_x(v_t - d_t) + \sigma_{\epsilon_x}\epsilon_{xt+1}. \tag{7}$$

Furthermore, v_t is independent of $v_s - d_s$ for all t and s.

Assumption 2 implies that, if there is no dividend, μ is the mean of the log-value-return $(v_{t+1} - v_t)$ and σ_r is the volatility. According to Assumption 3, the value-to-dividend ratio $v_t - d_t$ has a mean of $\bar{x}_v$ and conditional volatility of σ_{ϵ_x}, and is mean reverting with coefficient ρ_x. Equations (6) and (7) in Assumption 3 are used in the literature on predictive regressions, see for example, Stambaugh (1999) and Valkanov and Torous (2005).[1]

Asset pricing models typically determine the value-to-dividend ratio from preferences of the investors. For example, in the consumption-based asset pricing model where the representative agent has constant relative risk aversion coefficient and the dividend growth is independent and identically distributed (IID) over time, the value-to-dividend ratio is constant. However, in most models, the value-to-dividend ratio is stochastic and stationary. The above specification is an approximation and a simplification to a stationary value-to-dividend ratio. With the value process and value-dividend ratio process specified as above, the dividend growth process is implicitly determined. See Ang and Liu (2006) for a discussion on related issues.

Assumptions 2 and 3 are needed to obtain closed-form inference on noise Δ_t from prices and price ratios. With other non-gaussian specifications, it is not easy to compute in closed form the inference about the noise, but the same intuition applies. The independence assumption between v_t and $v_t - d_t$ is made to simplify the expression. Closed-form inference still obtains if the correlation is a non-zero constant.

When there are multiple stocks, the shocks $\epsilon_{\Delta t+1}$, ϵ_{rt+1}, and ϵ_{xt+1} could all have systematic components as well as idiosyncratic components. As we will show later, our results in later sections still apply with a reinterpretation of parameters when the correlation between stocks are introduced through common systematic factors.

We calibrate the above specification as follows, with all the parameters summarized in Table 1. The parameter μ only affects the overall magnitude of the expected return. We take μ to be 10%. Since the mean and volatility of the price-dividend ratio are small, the volatility of the stock return is largely due to price fluctuations. Note that, from Assumptions 1, 2, and 3,

$$p_{t+1} - p_t = v_{t+1} - v_t + e_{t+1} - e_t = \mu + (1-\rho)\Delta_t + \sigma_r\epsilon_{rt+1} + \sigma_{\epsilon_\Delta}\epsilon_{\Delta t+1},$$

[1] Note that there is no price noise in these studies, thus the value-dividend ratio is the price-dividend ratio.

thus, the variance of the return is the sum of the variance σ_r^2 of the value return $v_{t+1} - v_t$ and the conditional variance $\sigma_{\epsilon_\Delta}^2$ of the noise Δ_{t+1}. We will take $\sigma_r = 15\%$ and $\sigma_{\epsilon_\Delta} = \sigma_r/3 \approx 5\%$. The ratio of $\sigma_r/\sigma_{\epsilon_\Delta} = 3$ gives a ratio between variance of the noise and total variance of the stock return of 10%. French and Roll (1986) suggest that "between 4% and 12% of the daily return variances is caused by noise." Fama and French (1988) estimate that predictable variation due to mean reversion is about 35 percent of 3-5 year variances and they suggest, following Summers (1986), that the mean-reversion may be due to market inefficiency. In his calibration exercises, Summers (1986) uses the values of σ_r^2 that is of the same order of magnitude as $\sigma_{\epsilon_\Delta}^2$.

The value of ρ can be inferred from mean-reversion in prices, assuming the mean reversion is due to noise. Fama and French (1988) shows that there are significant mean-reversion in prices for holding-period horizons larger than 1 year. Summers (1986) uses values of ρ between 0.75 to 0.995 and Poterba and Summers (1988) use values between 0 and 0.70. We will consider a range of ρ, as Summers and Poterba and Summers. However, the value and size effect is not overly sensitive to ρ , as long as $0 < \rho < 1$.

The calibration of parameters for value-dividend ratio are based on the studies of Stambaugh (1999) and Valkanov and Torous (2005) on the predictive regression of the market portfolio. They found that the mean dividend ratio is about 3%, the AR(1) coefficient is above 0.9 and the conditional volatility is less than 1%. Because noise largely averages out in the market portfolio,[2] we expected the mean and AR(1) coefficient for the value-ratio process should be in the neighborhood of their estimates for for the market, thus we set $\bar{x}_v = 4$ and $\rho_x = 0.9$. We will set $\sigma_{\epsilon_x} = 10\%$.

2 Unconditional Expected Returns

In this section, we study the implications of noise on unconditional expected returns. We show that noise can generate cross-sectional variations in unconditional expected stock returns.

From equation (1) and by the stationarity of Δ_t, we have

$$\frac{P_{t+1}}{P_t} = \frac{V_{t+1}}{V_t}\frac{\mathrm{E}\left[e^{\Delta_t}\right]}{\mathrm{E}\left[e^{\Delta_{t+1}}\right]}e^{\Delta_{t+1}-\Delta_t} = \frac{V_{t+1}}{V_t}e^{\Delta_{t+1}-\Delta_t}. \tag{8}$$

Let D_t denote the dividend of the stock at time t. We assume that it is independent of the noise Δ_t. Then

$$\frac{D_{t+1}}{P_t} = \frac{D_{t+1}}{V_t}\mathrm{E}\left[e^{\Delta_t}\right]e^{-\Delta_t}. \tag{9}$$

The unconditional expected return is,

$$\mathrm{E}\left[\frac{P_{t+1}+D_{t+1}}{P_t}\right] = \mathrm{E}\left[\frac{V_{t+1}}{V_t}\right]\mathrm{E}\left[e^{\Delta_{t+1}-\Delta_t}\right] + \frac{D_{t+1}}{V_t}\mathrm{E}\left[e^{\Delta_t}\right]e^{-\Delta_t}. \tag{10}$$

Proposition 1 *If Assumption 1 holds, the expected return is higher than the expected value return.*

(Proof) By stationarity,

$$\mathrm{E}\left[\Delta_{t+1}\right] = \mathrm{E}\left[\Delta_t\right], \tag{11}$$

therefore,

$$\mathrm{E}\left[\Delta_{t+1}-\Delta_t\right] = 0. \tag{12}$$

By Jensen's inequality,

$$\mathrm{E}\left[e^{\Delta_{t+1}-\Delta_t}\right] \geq e^{\mathrm{E}\left[\Delta_{t+1}-\Delta_t\right]} = 1. \tag{13}$$

Equation (10) then gives,

$$\mathrm{E}\left[\frac{P_{t+1}}{P_t}\right] = \mathrm{E}\left[\frac{V_{t+1}}{V_t}\right]\mathrm{E}\left[e^{\Delta_{t+1}-\Delta_t}\right] \geq \mathrm{E}\left[\frac{V_{t+1}}{V_t}\right]. \tag{14}$$

Furthermore,

$$\mathrm{E}\left[\frac{D_{t+1}}{P_t}\right] = \mathrm{E}\left[\frac{D_{t+1}\mathrm{E}[e^{\Delta_t}]}{V_t e^{\Delta_t}}\right] = \mathrm{E}\left[\frac{D_{t+1}}{V_t}\right]\mathrm{E}[e^{\Delta_t}]\mathrm{E}\left[\frac{1}{e^{\Delta_t}}\right] \geq \mathrm{E}\left[\frac{D_{t+1}}{V_t}\right]. \tag{15}$$

Combining inequalities in (14) and (15), we conclude that

$$\mathrm{E}\left[\frac{P_{t+1}+D_{t+1}}{P_t}\right] \geq \mathrm{E}\left[\frac{V_{t+1}+D_{t+1}}{V_t}\right]. \tag{16}$$

[2]Campbell and Kyle (1993) study price noise of the market portfolio. Their paper suggest that there are systematic components in the price noise of individual stocks.

Blume and Stambaugh (1983) suggest that bid-ask spreads lead to a noise of the form $\Delta_t = 1 + \epsilon_{\Delta t}$, where $\epsilon_{\Delta t}$ is mean zero and independent across the time. They show that the noise increases the unconditional expected return for $\rho = 0$ and $D = 0$ case of the above Proposition.

Proposition 1 only requires that the noise is independent of the value and the dividend. With the additional assumption that the noise is an AR(1) process, we can established an exact relationship between the unconditional expected return and unconditional expected value return.

Proposition 2 *If Assumptions 1 and 2 hold, the expected return is given by*

$$\mathrm{E}\left[\frac{P_{t+1}+D_{t+1}}{P_t}\right] = \mathrm{E}\left[\frac{V_{t+1}}{V_t}\right] e^{\frac{\sigma^2_{\epsilon_\Delta}}{1+\rho}} + \mathrm{E}\left[\frac{D_{t+1}}{V_t}\right] e^{\frac{\sigma^2_{\epsilon_\Delta}}{1-\rho^2}}, \tag{17}$$

which is higher than the expected value return $\mathrm{E}\left[\frac{V_{t+1}}{V_t}\right] + \mathrm{E}\left[\frac{D_{t+1}}{V_t}\right]$. *Furthermore, if Assumption 3 also holds, then*

$$\mathrm{E}\left[\frac{P_{t+1}+D_{t+1}}{P_t}\right] = e^{\mu+\frac{1}{2}\sigma_r^2}\left(e^{\frac{\sigma^2_{\epsilon_\Delta}}{1+\rho}} + e^{-\bar{x}_v + \frac{\sigma^2_{\epsilon_x}}{2(1-\rho_x^2)} + \frac{\sigma^2_{\epsilon_\Delta}}{1-\rho^2}}\right). \tag{18}$$

(Proof) When equation (3) holds, we have

$$\begin{aligned}
\mathrm{E}[e^{\Delta_{t+1}-\Delta_t}] &= \mathrm{E}[e^{(1-\rho)\Delta_t}]\mathrm{E}[e^{\sigma_{\epsilon_\Delta}\epsilon_{t+1}}] = e^{\frac{(1-\rho)^2\sigma^2_{\epsilon_\Delta}}{2(1-\rho^2)}} e^{\frac{\sigma^2_{\epsilon_\Delta}}{2}} = e^{\frac{\sigma^2_{\epsilon_\Delta}}{1+\rho}};\\
\mathrm{E}[e^{\Delta_t}]\mathrm{E}\left[\frac{1}{e^{\Delta_t}}\right] &= e^{\frac{\sigma^2_{\epsilon_\Delta}}{1-\rho^2}},
\end{aligned}$$

noting $\frac{\sigma^2_{\epsilon_\Delta}}{1-\rho^2}$ is the unconditional variance of Δ_t. Since $e^{\frac{\sigma^2_{\epsilon_\Delta}}{1+\rho}} \geq 1$ and $e^{\frac{\sigma^2_{\epsilon_\Delta}}{1-\rho^2}} \geq 1$, we conclude that $\mathrm{E}\left[\frac{P_{t+1}+D_{t+1}}{P_t}\right] \geq \mathrm{E}\left[\frac{V_{t+1}+D_{t+1}}{V_t}\right]$. When Assumption 3 holds, equation (18) is proved by noting that

$$\begin{aligned}
\mathrm{E}\left[\frac{V_{t+1}}{V_t}\right] &= e^{\mu+\frac{1}{2}\sigma_r^2};\\
\mathrm{E}\left[\frac{D_{t+1}}{V_t}\right] &= \mathrm{E}\left[\frac{V_{t+1}}{V_t}\right]\mathrm{E}\left[\frac{D_{t+1}}{V_{t+1}}\right] = e^{\mu+\frac{1}{2}\sigma_r^2} e^{-\bar{x}_v + \frac{\sigma^2_{\epsilon_x}}{2(1-\rho_x^2)}}.
\end{aligned}$$

The unconditional expected return in the absence of noise is

$$e^{\mu+\frac{1}{2}\sigma_r^2}\left(1 + e^{-\bar{x}_v + \frac{\sigma^2_{\epsilon_x}}{2(1-\rho_x^2)}}\right),$$

which should be determined by asset pricing theories thus should depend only on beta under CAPM or APT. Proposition 1 and 2 hold without any specifications of asset pricing theory and thus are valid quite generally.

Cross-section variations in unconditional expected returns can be generated by noise, according to Proposition 2. With noise, the unconditional expected return given in equation (18) depends also on idiosyncratic volatility, the volatility σ_{ϵ_Δ} and AR(1) coefficient ρ of noise Δ_t and the parameters $(\bar{x}_v, \sigma_x, \rho_x)$ of the price-dividend ratio, in addition to beta. That is, given two stocks with either different noise variance or mean price-dividend ratio, the unconditional expected returns can be different, even if they have the same (systematic) risk. In other words, cross-sectional variations can be generated by variations in these parameters. It is not very satisfactory that the cross-sectional variation has to be exogenously specified (through specification of parameter variations). On the other hand, it is not true that one can always generate cross-sectional variations in expected returns with parameter variations. For example, in standard asset pricing models such as CAPM and APT, variations in idiosyncratic volatilities do not generate cross-sectional variations in expected returns.

From the above Proposition, the effect of noise on unconditional expected returns is at the order of $\sigma^2_{\epsilon_\Delta}$. With a value of 6% for $\sigma^2_{\epsilon_\Delta}$, given in Table 1, the change in unconditional expected returns is about 36 basis point. However, if $\sigma^2_{\epsilon_\Delta} = 10\%$, which is not unreasonable for some stocks, the change will be 1%.

The difference between the unconditional expected return and unconditional expected value return is due to Jensen's inequality, which is driven by the variance of the random variable. Therefore it is only natural that the difference between the expected return and value return increases with $\sigma^2_{\epsilon_\Delta}$ for $\rho < 1$. Proposition 1 and 2 are more generalized versions of the result presented in Hsu (2006). Brennan and Wang (2006) also derive similar results.

Blume and Stambaugh (1983) compute the unconditional expected return for $\rho = 0$ and $D = 0$ case of Proposition 2. They show that the size effect observed in daily returns can be explained by the noise they suggested.

Berk (1997) computes unconditional cross-section correlation between price and the return. As in our model, the cross sectional variation in unconditional expected returns in Berk (1997) needs to be generated from variations in parameters.

One implication of our paper is that, ceteris paribus, a less transparent stock (one that is more likely to be mispriced and therefore has a higher σ_{ϵ_Δ}) will have a higher unconditional expected return. This is consistent with recent empirical findings where the cost of capital for a firm, controlling for beta, is higher when the firm is less transparent. Hughes, Liu, and Liu (2006) argue that these empirical findings may not be explained by risk. The propositions suggest that noise could provide a potential explanation for this empirical finding.

Shiller (1981) points out that the return variance for a stock, in a world with IID dividend growth and CRRA representative preference, should be equal to the variance of its dividend growth. However, empirically, the variance in stock dividend growth is lower than the variance in return, giving rise to Shiller's excess-volatility puzzle. In our model, the variance of the return is the sum of the variance of the value return and the variance of the noise. This potentially offers a perhaps indelicate explanation for the excess-volatility puzzle, as suggested in Campbell and Kyle (1993).

In later sections, the conditional expected return will be compared with equation (18).

3 The Intuition for Conditional Expected Returns

In this section, we present the intuition for why expected returns depend on price or price ratios when there is noise in price. Let us first assume that the noise Δ_t is observed. In this section and this section only, for the notational simplicity, we will use the additive form of noise:

$$P_t = V_t + \Delta_t.$$

It then follows that

$$\begin{aligned}\frac{P_{t+1} + D_{t+1}}{P_t} &= \frac{V_{t+1} + D_{t+1}}{V_t + \Delta_t} + \frac{V_{t+1} + \Delta_{t+1} - V_{t+1}}{V_t + \Delta_t} \\ &= \frac{V_t}{V_t + \Delta_t}\frac{V_{t+1} + D_{t+1}}{V_t} + \frac{\Delta_{t+1}}{V_t + \Delta_t}. \end{aligned} \tag{19}$$

The factor $\frac{V_t}{V_t+\Delta_t}$ is the relative mispricing at time t, $\frac{V_{t+1}+D_{t+1}}{V_t} + \frac{\Delta_{t+1}}{V_t+\Delta_t}$ is the value return, which is the return without noise, and $\frac{\Delta_{t+1}}{V_t+\Delta_t}$ is due to noise at time $t+1$. To be specific, we will assume that the value return satisfies the following relation

$$\frac{V_{t+1} + D_{t+1}}{V_t} = R_f + \beta\lambda + \beta F_{t+1} + \sigma_r \epsilon_{rt+1},$$

which is true under either CAPM or APT. The gross risk-free rate is R_f, the factor is F_{t+1}, the factor risk premium is λ, idiosyncratic risk is given by ϵ_{rt+1}, and the idiosyncratic volatility is σ_r. We can write

$$\frac{V_t}{V_t + \Delta_t}\frac{V_{t+1} + D_{t+1}}{V_t} = -\frac{\Delta_t}{V_t + \Delta_t}R_f + R_f + \frac{V_t}{V_t + \Delta_t}(\beta\lambda + \beta F_{t+1} + \sigma_r \epsilon_{rt+1}).$$

This equation implies that the beta and volatility of the return is scaled by a factor of $\frac{V_t}{V_t+\Delta_t}$. The risk premium is also scaled by the same factor. Thus, $R_f + \frac{V_t}{V_t+\Delta_t}(\beta\lambda + \beta F_{t+1} + \sigma_r \epsilon_{rt+1})$ is a fair return with theoretically correct compensation. The term $-\frac{\Delta_t}{V_t+\Delta_t}R_f$ represents the extra return spread that is not associated with risk but is associated with mispricing generated by noise. When $\Delta_t < 0$, the stock is under-valued and the spread is positive. Note that in this case, both systematic risk and idiosyncratic risk are higher.

Furthermore,

$$\frac{\Delta_{t+1}}{V_t + \Delta_t} = \frac{\rho\Delta_t + \sigma_{\epsilon_\Delta}\epsilon_{\Delta t+1}}{V_t + \Delta_t}.$$

When the AR(1) coefficient ρ of the noise is not zero, the pricing error produce by noise Δ_t at time t will be persistent and lead to an average pricing error of $\rho\Delta_{t+1}$ at time $t+1$, thus leading to an extra term $\frac{\rho\Delta_t}{V_t+\Delta_t}$ in expected return. Putting all terms together, the return is

$$\begin{aligned}\frac{P_{t+1} + D_{t+1}}{P_t} &= -\frac{R_f - \rho}{V_t + \Delta_t}\Delta_t + R_f + \frac{V_t}{V_t + \Delta_t}(\beta\lambda + \beta F + \sigma_r \epsilon_{rt+1} + \sigma_{\epsilon_\Delta}\epsilon_{\Delta t+1}) \\ &= -\frac{R_f - \rho}{P_t}\Delta_t + R_f + \frac{P_t - \Delta_t}{P_t}(\beta\lambda + \beta F + \sigma_r \epsilon_{rt+1} + \sigma_{\epsilon_\Delta}\epsilon_{\Delta t+1}).\end{aligned}$$

Accordingly, suppose that there is a negative pricing error at time t, $\Delta_t < 0$, the idiosyncratic risk will be higher because both $\frac{P_t - \Delta_t}{P_t}\sigma_r > \sigma_r$ and there is an extra risk associated with noise at time $t+1$, the beta thus the risk premium associated with the factor risk will be higher. In addition, there is an alpha term, $-\frac{R_f - \rho}{P_t}\Delta_t$, which is due to the fact that the stock is under-valued.

In reality, we do not observed the noise Δ_t. However, we can still infer Δ_t from the price P_t or price ratios. The lower the price or the price ratios, the more likely Δ_t is negative and the stock is under-valued. Under the Gaussian setting specified in Assumptions 1-3, the inference can be precisely computed. In the rest of the paper, we will compute the average Δ_t given P_t or price ratios and thus the expected return conditional on P_t or price ratios.

Note that in Berk (1995, 1997), higher expected returns for low-priced stocks are due to higher systematic risks, which is different from ours.

4 The Size Effect

In this section, we study the expected return, conditional on the current price P_t. We show that the conditional expected return decreases with P_t. We also compute the expected return conditional on price deciles.

Note that the return is,

$$\frac{P_{t+1} + D_{t+1}}{P_t} = \frac{V_{t+1}}{V_t} e^{\Delta_{t+1} - \Delta_t} + \frac{D_{t+1}}{V_t} \mathrm{E}[e^{\Delta_t}] e^{-\Delta_t}. \tag{20}$$

We are interested in the expected return, conditional on the current price P_t,

$$\mathrm{E}\left[\frac{P_{t+1} + D_{t+1}}{P_t} | P_t\right].$$

As we noted previously, the value return $\frac{V_{t+1} + D_{t+1}}{V_t}$ is determined by pricing models and may have systematic as well as idiosyncratic component; for our purpose, it is not necessary to specify this. Similarly, Δ_t may also have systematic components, as in Campbell and Kyle (1993). The systematic components will not affect the inferences on individual noise in an economy with a large number of stocks, as we shown in the appendix.

Note that $p_t = v_t + \Delta_t - \ln(\mathrm{E}[e^{\Delta_t}])$. To draw inference of noise Δ_t from price p_t, we need to know the joint distribution of v_t and Δ_t. It is natural to assume that the distribution of Δ_t is its stationary distribution, which has mean of 0 and variance of $\frac{\sigma^2_{\epsilon_\Delta}}{1-\rho^2}$. Since v_t is not stationary, there is no natural choice of distribution for v_t. We will assume that v_t is normal with mean $\bar{v}_t$ and variance σ^2_{vt}. From Assumptions 1, 2, 3, v_t and Δ_t are independent.

Proposition 3 *Suppose Assumptions 1, 2, and 3 hold. Furthermore, assume that the distribution of Δ_t is its unconditional distribution and the distribution of v_t is normal with mean $\bar{v}_t$ and variance σ^2_{vt}. Then the expected return conditional on P_t is*

$$\begin{aligned} &\mathrm{E}\left[\frac{P_{t+1} + D_{t+1}}{P_t} | p_t\right] \\ = \;& e^{\mu + \frac{1}{2}\sigma_r^2} \left(e^{\frac{\sigma^2_{\epsilon_\Delta}}{1+\rho}} \frac{P_t^{-(1-\rho)\gamma_1}}{\mathrm{E}\left[P_t^{-(1-\rho)\gamma_1}\right]} + e^{-\bar{x}_v + \frac{\sigma^2_{\epsilon_x}}{2(1-\rho_x^2)} + \frac{\sigma^2_{\epsilon_\Delta}}{1-\rho^2}} \frac{P_t^{-\gamma_1}}{\mathrm{E}\left[P_t^{-\gamma_1}\right]} \right), \end{aligned} \tag{21}$$

where $\gamma_1 = \frac{\sigma^2_{\epsilon_\Delta}}{(1-\rho^2)\sigma^2_{vt} + \sigma^2_{\epsilon_\Delta}}$.

The proof is given in the appendix. It is clear that the expected return, conditional on P_t, decreases with P_t. The results from the proposition is intuitive. Consider the case where the noise is independent over time ($\rho = 0$). In this case,

$$\frac{P_{t+1}}{P_t} = \frac{V_{t+1}}{V_t} e^{\Delta_{t+1} - \Delta_t}. \tag{22}$$

The expectation of $e^{\Delta_{t+1}}$ conditional on Δ_t is independent of Δ_t when $\rho = 0$. Thus, the expected return will be decreasing in Δ_t. If there is a negative noise, the stock is under-valued, so that the subsequent return is high on average. Clearly, we do not observe Δ_t; however, we can infer information on Δ_t from observing P_t. That is, the price can be a noisy signal for the noise. Recall,

$$p_t = v_t + \Delta_t - \ln(\mathrm{E}[e^{\Delta_t}]). \tag{23}$$

Therefore, the higher the p_t, the higher the probable pricing error, on average, and the lower the next period return.

In this paper, we do not assume that $\rho = 0$, thus Δ_{t+1} need not be independent of Δ_t. This is plausible since some forms of pricing error may require months or years to be identified and corrected by the market. When $0 < \rho < 1$, the effect of noise on return should be reduced. In this case, a positive realization of noise at time t implies on average a positive realiation at $t+1$, although the it will be smaller. Suppose, for example, the noise is persistent; in this case, ρ approaches 1, and Δ_t is a random walk. If this is the case, although the noise affects the market price, it does not affect the return because the error does not correct over time; an under-valued stock remains under-valued.

We should remark that in Proposition 3, the parameter μ is assumed to be a constant. This implies that the expected value return is independent of value v_t, which is true in many asset pricing theories, such as Capital Asset Pricing Model (CAPM) and the Arbitrage Pricing Theory (APT) and can be obtained more or less under homothetic preference. However, this assumption does not always hold. For example, Black and Litterman (1992) assume that the risk premium of a stock should be proportional to its market cap (which is price), which is an easy way to clear the market. In this case, μ depends linearly on v_t. Depending on relative magnitude of the coefficient of this linear dependence and the γ_1, the conditional expected return may decrease or increase with P_t.

Fama and French (1992) provide an informative illustration of the size effect as follows. Stocks are classified into deciles according to their market capitalization and the average return for each decile is computed. We will term these averages the expected return conditional on deciles. These expected returns demonstrate the cross-sectional variations in expected return conditional on size. The size spread is defined to be the difference between the expected return conditional on the 10th decile and 1st decile, which is more coarse measure of size effect. Both expected return conditional on deciles and size premium can be computed in our model. Let δ_i by the following equation

$$N(\delta_i) = \frac{i}{10}, \quad i = 1, ..., 9,$$

where $N(\cdot)$ is the cumulative probability distribution function of the standard normal random variable, $\delta_0 = -\infty$, and $\delta_{10} = +\infty$. At time t, p_t is normally distributed with mean $\bar{p}_t$ and variance $\sigma^2_{p_t} = \sigma^2_{vt} + \frac{\sigma^2_{\epsilon_\Delta}}{1-\rho^2}$. Therefore, $p_{ti} = \sigma_{pt}\delta_i + \bar{p}_t, \quad i = 0,\ 1,\ ...,\ 9,\ 10$, divide p_t-space into deciles.

Proposition 4 (Size Effect) *Suppose that the assumptions of Proposition 3 hold, then the expected return conditional on decile is*

$$e^{\mu+\frac{1}{2}\sigma_r^2}\left(e^{\frac{\sigma^2_{\epsilon_\Delta}}{1+\rho}}\frac{N(\hat{p}_{ti}) - N(\hat{p}_{ti-1})}{0.1} + e^{-\bar{x}_v + \frac{\sigma^2_{\epsilon_x}}{2(1-\rho_x^2)} + \frac{\sigma^2_{\epsilon_\Delta}}{1-\rho^2}}\frac{N(\check{p}_{ti}) - N(\check{p}_{t_{i-1}})}{0.1}\right), \tag{24}$$

where $\hat{p}_{ti} \equiv \delta_i + (1-\rho)\gamma_1\sigma_{pt}$ *and* $\check{p}_{ti} \equiv \delta_i + \gamma_1\sigma_{pt}$, $i = 1, ..., 9$. *The size spread is given by*

$$e^{\mu+\frac{1}{2}\sigma_r^2}\left(e^{\frac{\sigma^2_{\epsilon_\Delta}}{1+\rho}}\frac{N(\hat{p}_{t9}) + N(\hat{p}_{t1}) - 1}{0.1} + e^{-\bar{x}_v + \frac{\sigma^2_{\epsilon_x}}{2(1-\rho_x^2)} + \frac{\sigma^2_{\epsilon_\Delta}}{1-\rho^2}}\frac{N(\check{p}_{t9}) + N(\check{p}_{t1}) - 1}{0.1}\right). \tag{25}$$

The proposition can be proved from Proposition 3 by integration.

When $\sigma_{\epsilon_\Delta} = 0$, the conditional expected return is independent of P_t, and the return spreads between two price deciles portfolios are zero. Similarly, as σ_{vt} increases, the spread decreases, because a higher σ_{vt} is equivalent to a lower σ_{ϵ_Δ}.

For calibration, we use parameters given in Table 1. In addition, we need to specify σ^2_{vt}. Since v_t is not stationary, there is no natural choice for $\bar{v}_t$ and σ^2_{vt}. Fortunately, $\bar{v}_t$ does not affect the p_t dependence. We choose σ^2_{vt} to be at the same order of magnitude σ^2_r. With these parameters, the size spread is about 3%. The more persistence the noise exhibits, the less effect it has on the spread. Thus, the spread decreases with ρ for small ρ. However, for a given Σ_{ϵ_Δ}, the higher ρ leads to a higher unconditional variance of Δ, which is assumed to be the prior distribution of Δ_t, thus higher spread. This effects dominates for ρ near 1. Thus, the spread has an U-shaped dependence and thus a minimum, this feature makes it relatively easier to generate higher spreads than lower spreads.

So far, we have examined a single stock; we have not consider noise in a multi-asset framework. If there are multiple assets, we need to consider the correlations between the value returns and the correlations between noise. We argue in the appendix that our results on price dependence still hold. Specifically, we can still examine the price dependence of expected returns on a stock-by-stock basis, if the correlations are introduced through a factor structure and the number of asset is large.[3] Roughly speaking, in this case, we have infinitely many signals on a few factors. As such, the factors will be completely revealed and the inference problem reduces to that without systematic factors.

[3] Note that this is the assumption needed for APT to hold.

Arnott, Hsu, and Moore (2005) and Arnott (2005a) propose noise as a likely source for size and value effects. Hsu (2006) shows that mispricing premium may exist because there are investors with liquidity needs. Berk (1997) and Arnott (2005b) suggest that size and value are highly interrelated and may be proxies for a shared risk. Arnott and Hsu (2006) show that mean-reverting mispricing can lead to small cap and value stock outperformance; however, they predict that size and value might subsume each other. Brennan and Wang (2006) also use a similar model to explore asset pricing implication associated with mispricing. Similar to Hsu (2006), they derive a return premium associated with mispricing. Specifically they argue that common liquidity measures in finance may be proxies for mispricing and that estimated liquidity premium is likely mispricing premium.

5 The Value Effect

Many empirical studies analyze expected returns conditional on price-fundamental ratios, such as price-dividend ratio, price-book ratio, and price-earning ratios. In this section, we examine the price-dividend ratio dependence of expected returns when noise is present. Conceptually, the analysis applies in the same way to any price-fundamental ratio dependence. Since we have to specify dividend-price ratio for computing return already, we choose the price-to-dividend ratio instead of other ratios to avoid additional parameters.

In this section, we use the price-dividend ratio $X_t \equiv \frac{P_t}{D_t} = e^{p_t - d_t}$ to draw inference on the noise Δ_t. We will use x_t to denote $\ln X_t = p_t - d_t$. Recall, when there is noise,

$$p_t = v_t + \Delta_t - \ln(\mathrm{E}[e^{\Delta_t}]). \tag{26}$$

The error also works itself into the price-dividend ratio,

$$p_t - d_t = v_t - d_t + \Delta_t - \ln(\mathrm{E}[e^{\Delta_t}]). \tag{27}$$

Thus, a high price-dividend ratio can be a signal for a high noise. This same logic applies equally for price-book, price-earnings, and other price-fundamental ratios.

The specification of value-dividend ratio given in equation (7) implies the following relationship for the price-dividend ratio,

$$\begin{aligned} & p_{t+1} - d_{t+1} \\ = \; & (1-\rho_x)\bar{x}_v - (1-\rho_x)\ln(\mathrm{E}[e^{\Delta_t}]) + \rho_x(p_t - d_t) + (\rho - \rho_x)\Delta_t + \sigma_{\epsilon_x}\epsilon_{xt+1} + \sigma_{\epsilon_\Delta}\epsilon_{t+1}. \end{aligned} \tag{28}$$

Denoting $x_t = p_t - d_t$, we have,

$$x_{t+1} = (1-\rho_x)\bar{x} + \rho_x x_t + (\rho - \rho_x)\Delta_t + \sigma_{\epsilon_x}\epsilon_{xt+1} + \sigma_{\epsilon_\Delta}\epsilon_{\Delta t+1}, \tag{29}$$

where $\bar{x} = \bar{x}_v - \ln(\mathrm{E}[e^{\Delta_t}])$ is the mean of x_t. We make the standard assumption that value-dividend ratio is stationary, which means that x_{t+1} is stationary, thus $\rho_x < 1$. The above equation implies that the (*log*) price-dividend ratio x_t is a signal on the noise Δ_t. This implies that price-dividend ratio and other price-fundamental ratios could provide inference on the noise. Since x_t is stationary, we can use its unconditional distribution as the prior distribution for inference.

Proposition 5 *Suppose that Assumptions 1, 2, and 3 hold. Furthermore, assume that the distribution of (Δ_t, x_t) is their unconditional distribution. Then the expected return conditional on x_t is*

$$\begin{aligned} & \mathrm{E}\left[\frac{P_{t+1} + D_{t+1}}{P_t} | x_t\right] \\ = \; & e^{\mu + \frac{1}{2}\sigma_r^2}\left(e^{\frac{\sigma_{\epsilon_\Delta}^2}{1+\rho}}\frac{X_t^{-(1-\rho)\gamma_2}}{\mathrm{E}\left[X_t^{-(1-\rho)\gamma_2}\right]} + e^{-\bar{x}_v + \frac{\sigma_{\epsilon_x}^2}{2(1-\rho_x^2)} + \frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}\frac{X_t^{-(1-\rho_x)\gamma_2 - \rho_x}}{\mathrm{E}\left[X_t^{-(1-\rho_x)\gamma_2 - \rho_x}\right]}\right), \end{aligned}$$

where $\gamma_2 = \frac{(1-\rho_x^2)\sigma_{\epsilon_\Delta}^2}{(1-\rho_x^2)\sigma_{\epsilon_\Delta}^2 + (1-\rho^2)\sigma_{\epsilon_x}^2}$.

The proof is given in the Appendix. The intuition for the x_t dependence is the same as the intuition for the p_t dependence explored in in Section 4. A high price-dividend ratio implies a high noise Δ_t, on average, thus a low expected return.

Proposition 5 also implies that the return is predicted by the dividend yield even though the value return is not. This is not surprising because there is a one-to-one correspondence between excess volatility and dividend yield predictability. That is, while return exhibits excess volatility relative to dividend variation, value return does not, and while dividend yield predicts return, it does not predict value return. Note that both the excess volatility and dividend yield predictability puzzle results from noise instead of a rational equilibrium.

We can also compute the expected return conditional on value deciles, following Fama and French (1992). At time t, x_t is normally distributed with mean $\bar{x}$ and variance $\frac{\sigma^2_{\epsilon_x}}{1-\rho_x^2}+\frac{\sigma^2_{\epsilon_\Delta}}{1-\rho^2}$. Therefore, $x_i=\sqrt{\frac{\sigma^2_{\epsilon_x}}{1-\rho_x^2}+\frac{\sigma^2_{\epsilon_\Delta}}{1-\rho^2}}\delta_i+\bar{x}, \quad i=0,1,...,9,10$, divides x_t-space into deciles. We will term the difference in the expected returns between 1st and 10th decile the value spread.

Proposition 6 (Value Effect) *Suppose assumptions in Proposition 5 hold. Then the expected return conditional on value decile is*

$$e^{\mu+\frac{1}{2}\sigma_r^2}\left(e^{\frac{\sigma^2_{\epsilon_\Delta}}{1+\rho}}\frac{N(\hat{x}_i)-N(\hat{x}_{i-1})}{0.1}+e^{-\bar{x}_v+\frac{\sigma^2_{\epsilon_x}}{2(1-\rho_x^2)}+\frac{\sigma^2_{\epsilon_\Delta}}{1-\rho^2}}\frac{N(\check{x}_i)-N(\check{x}_{i-1})}{0.1}\right), \tag{30}$$

where $\hat{x}_i\equiv\delta_i+(1-\rho)\gamma_2\sqrt{\frac{\sigma^2_{\epsilon_x}}{1-\rho_x^2}+\frac{\sigma^2_{\epsilon_\Delta}}{1-\rho^2}}$ *and* $\check{x}_i=\delta_i+\Big((1-\rho_x)\gamma_2+\rho_x\Big)\sqrt{\frac{\sigma^2_{\epsilon_x}}{1-\rho_x^2}+\frac{\sigma^2_{\epsilon_\Delta}}{1-\rho^2}}$, $i=1,...,9$. *The value spread is given by*

$$e^{\mu+\frac{1}{2}\sigma_r^2}\left(e^{\frac{\sigma^2_{\epsilon_\Delta}}{1+\rho}}\frac{N(\hat{x}_9)+N(\hat{x}_1)-1}{0.1}+e^{-\bar{x}_v+\frac{\sigma^2_{\epsilon_x}}{2(1-\rho_x^2)}+\frac{\sigma^2_{\epsilon_\Delta}}{1-\rho^2}}\frac{N(\check{x}_9)+N(\check{x}_1)-1}{0.1}\right). \tag{31}$$

The proposition can be proved from Proposition 5 by integration.

For the parameters given in Table 1, the value spread is about 6%. The dependence on ρ is more sensitive for the value spread, primarily due to the fact that the volatility σ_x of price-dividend ratio x_t is much smaller than that of the volatility σ_{vt} of the value v_t.

6 The Size-Value Effect

So far, we have studied the expected return conditional on either the price or the price-dividend ratio alone. We now compute the expected return conditional on the price and price-dividend ratio simultaneously.

In our model, the size and value effects are both driven by the same source: the noise in the price. Conversely, both price p_t and price-dividend ratio p_t-d_t are noisy signals of Δ_t. We assume that the correlation between v_t and v_t-d_t is zero, however, there is an imperfect correlation between p_t and p_t-d_t induced by the noise Δ_t. When p_t is low, it is likely that Δ_t is negative, but we are not sure, because the value v_t is not observed. When both p_t and p_t-d_t are low, it is more likely that Δ is negative. Thus p_t and p_t-d_t are correlated but not a substitute of each other. Using both of them simultaneously gives us more precise information about Δ_t.

Proposition 7 *Suppose Assumptions 1, 2, and 3 hold. Furthermore, assume that the distribution of (Δ_t, x_t) is their unconditional distribution and the distribution of v_t is normal with mean $\bar{v}_t$ and variance σ^2_{vt}. Then the expected return conditional on p_t and x_t is,*

$$\begin{aligned}&\mathrm{E}\left[\frac{P_{t+1}+D_{t+1}}{P_t}|x_t,p_t\right]\\ =\;& e^{\mu+\frac{1}{2}\sigma_r^2}\left(e^{\frac{\sigma^2_{\epsilon_\Delta}}{1+\rho}}\frac{P_t^{-(1-\rho)\gamma_3}X_t^{-(1-\rho)\gamma_4}}{\mathrm{E}\Big[P_t^{-(1-\rho)\gamma_3}X_t^{-(1-\rho)\gamma_4}\Big]}+e^{-\bar{x}_v+\frac{\sigma^2_{\epsilon_x}}{2(1-\rho_x^2)}+\frac{\sigma^2_{\epsilon_\Delta}}{1-\rho^2}}\frac{P_t^{-(1-\rho_x)\gamma_3}X_t^{-(1-\rho_x)\gamma_4-\rho_x}}{\mathrm{E}\Big[P_t^{-(1-\rho_x)\gamma_3}X_t^{-(1-\rho_x)\gamma_4-\rho_x}\Big]}\right),\end{aligned}$$

where $\gamma_3=\frac{\frac{1}{\sigma^2_{vt}}}{\frac{1}{\sigma^2_{vt}}+\frac{1-\rho_x^2}{\sigma^2_{\epsilon_x}}+\frac{1-\rho^2}{\sigma^2_{\epsilon_\Delta}}}$ *and* $\gamma_4=\frac{\frac{(1-\rho_x^2)}{\sigma^2_{\epsilon_x}}}{\frac{1}{\sigma^2_{vt}}+\frac{1-\rho_x^2}{\sigma^2_{\epsilon_x}}+\frac{1-\rho^2}{\sigma^2_{\epsilon_\Delta}}}$.

The proof is given in the Appendix. We assume that the correlation between v_t and v_t-d_t is zero for notational simplicity. Incorporation of a non-zero correlation is straightforward.

Fama and French (1992) use the matrix of expected return conditional on size and value deciles to demonstrate the size and value effects. Next we compute these conditional expected returns using our model. We first divide (p_t,x_t) space into cells of 10 deciles by 10 deciles. Note that p_t and x_t are joint normal with variances $\sqrt{\sigma^2_{vt}+\frac{\sigma^2_{\epsilon}}{1-\rho^2}}$

and $\sqrt{\frac{\sigma_x^2}{1-\rho_x^2}+\frac{\sigma_e^2}{1-\rho^2}}$ and correlation $\hat{\rho}=\frac{\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}{\sigma_{pt}\sqrt{\frac{\sigma_{\epsilon_x}^2}{1-\rho_x^2}+\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}}$. Following Fama and French, we will first use p_{ti} to divided p_t space into 10 deciles. For i-th size decile, we further divide x_t space into 10 deciles, using $x_{i,j}=\sqrt{\frac{\sigma_x^2}{1-\rho_x^2}+\frac{\sigma_e^2}{1-\rho^2}}\delta_{i,j}+\bar{x}$, where $\delta_{i,j}$ can be solved numerically. Let $\mathrm{E}\left[f(z)|_{\underline{z}}^{\overline{z}}\right]$ denote the expectation of $f(z)$ for z between $\underline{z}$ and $\overline{z}$ for a standard normal random variable z .

Proposition 8 (Size-Value Effect) *Suppose that assumptions in Proposition 7 hold. Then the expected return conditional on (i,j) decile of (p_t,x_t) space is,*

$$e^{\mu+\frac{1}{2}\sigma_r^2}\left(e^{\frac{\sigma_{\epsilon_\Delta}^2}{1+\rho}}\frac{\mathrm{E}\left[\left(N\left(\frac{\hat{\hat{p}}_{i+1}-\hat{\rho}z}{\sqrt{1-\hat{\rho}^2}}\right)-N\left(\frac{\hat{\hat{p}}_i-\hat{\rho}z}{\sqrt{1-\hat{\rho}^2}}\right)\right)\Big|_{\hat{\hat{x}}_{i,j}}^{\hat{\hat{x}}_{i,j+1}}\right]}{0.01}\right.$$

$$\left.+\quad e^{-\bar{x}_v+\frac{\sigma_{\epsilon_x}^2}{2(1-\rho_x^2)}+\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}\frac{\mathrm{E}\left[\left(N\left(\frac{\check{\check{p}}_{i+1}-\hat{\rho}z}{\sqrt{1-\hat{\rho}^2}}\right)-N\left(\frac{\check{\check{p}}_i-\hat{\rho}z}{\sqrt{1-\hat{\rho}^2}}\right)\right)\Big|_{\check{\check{x}}_{i,j}}^{\check{\check{x}}_{i,j+1}}\right]}{0.01}\right),$$

where $\hat{\hat{p}}_{ti}\equiv\delta_i+(1-\rho)\left(\gamma_3\sigma_{pt}+\hat{\rho}\gamma_4\sqrt{\frac{\sigma_{\epsilon_x}^2}{1-\rho_x^2}+\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}\right)$, $\hat{\hat{x}}_{i,j}\equiv\delta_{i,j}+(1-\rho)\left(\gamma_4\sqrt{\frac{\sigma_{\epsilon_x}^2}{1-\rho_x^2}+\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}+\hat{\rho}\gamma_3\sigma_{pt}\right)$, $\check{\check{p}}_{ti}\equiv\delta_i+(1-\rho_x)\left(\gamma_3\sigma_{pt}+\hat{\rho}\gamma_4\sqrt{\frac{\sigma_{\epsilon_x}^2}{1-\rho_x^2}+\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}\right)$, *and* $\check{\check{x}}_{i,}=\delta_{i,j}+\Big((1-\rho_x)\gamma_4+\rho_x\Big)\sqrt{\frac{\sigma_{\epsilon_x}^2}{1-\rho_x^2}+\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}+(1-\rho_x)\hat{\rho}\gamma_3\sigma_{pt}$, $i=1,...,9$, *and z is a standard normal random variable.*

The proof is given in the appendix.

Let us consider the case where there are many stocks with correlations between stock returns. We show that, in the appendix, if the correlations in the returns as well as noise is introduced through a factor model, the inference on Δ_t is the same as if there is no factor. This means that, Propositions 3–8 hold when the correlations are through factors, provided we replace the variance parameters by their idiosyncratic components.

Suppose the returns of all stocks are given by a factor model and all have the same beta and same idiosyncratic volatility. Then the cross-section average are the same as population average, thus can be computed using Propositions 3-8. So, these proposition imply cross-sectional variations in conditional expected returns, even in the absence of parameter variation. The variation in this case is generated by random realization of the price noise. Of course, parameter variations in reality, such as variations in betas and idiosyncratic volatility, lead to additional cross-sectional variations in expected returns. Next we will show that these variations are consistent with those observed in the US data, with plausible parameters.

For the calibration exercise, we use parameters specified in Table 1. We present expected returns conditional on both size and value in Table 2. The intuition for the table is simple. Decile expected returns are really expected returns conditional on price intervals or price-ratio intervals, which decreases with price and/or price-ratios, as shown in the table. We assume that stocks are independent draws from the same distribution.

It is interesting to compare Table 2 with Table V of Fama-French (1992), which are sample average of returns conditional on size and price-to-book deciles. As we pointed out earlier, we choose price-dividend deciles mainly to avoid extra parameters. We expect the difference in using price-dividend ratio and price-book ratio to be small. The expected returns our Table 2 are similar to those of Table V of Fama and French (1992), when annualized. The expected returns are monotonic as a functions of deciles while the monotonicity is not strict in Table V of Fama and French (1992), presumably because of measurement errors in the sample averages.

It is important to determine whether small and value stocks have higher expected returns because they have higher systematic risks. In Table 3, we present the beta matrix for size-value deciles. Assuming that beta in the absence of noise is 1, small and value stocks have a slightly higher beta. Stocks in the smallest decile have a beta of 1.02 while those in the largest decile has a beta of 0.99. Similarly, Stocks in the lowest dividend-price ratio decile have a beta of 0.98 while those in the highest decile has a beta of 1.03. This finding is consistent Lakonishok, Shleifer, and Vishny (1994) who find that "the betas of value portfolios with respect to the value-weighted index tend to be about 0.1 higher than the betas of the glamour portfolios."

Table 2: Expected Annual Returns Conditional on Size and Value Deciles

	Dividend-to-Price Ratio										
	All	1	2	3	4	5	6	7	8	9	10
All	10.08	7.52	8.50	9.03	9.45	9.84	10.22	10.62	11.08	11.68	12.89
Small-ME	11.63	9.08	10.04	10.56	10.98	11.36	11.73	12.13	12.58	13.17	14.36
ME-2	11.00	8.49	9.44	9.95	10.37	10.74	11.11	11.51	11.95	12.53	13.71
ME-3	10.67	8.18	9.13	9.64	10.05	10.43	10.80	11.19	11.63	12.21	13.39
ME-4	10.42	7.94	8.88	9.39	9.80	10.18	10.55	10.94	11.38	11.95	13.13
ME-5	10.19	7.72	8.66	9.17	9.58	9.95	10.32	10.71	11.15	11.73	12.90
ME-6	9.97	7.51	8.45	8.95	9.36	9.74	10.11	10.49	10.93	11.51	12.68
ME-7	9.74	7.29	8.23	8.73	9.14	9.51	9.88	10.27	10.71	11.28	12.45
ME-8	9.49	7.04	7.98	8.49	8.89	9.27	9.63	10.02	10.46	11.03	12.20
ME-9	9.17	6.74	7.67	8.18	8.58	8.95	9.32	9.71	10.14	10.71	11.87
Large-ME	8.56	6.13	7.07	7.57	7.98	8.35	8.72	9.10	9.54	10.11	11.27

This table presents annual expected returns, in percentage, conditional on price (ME) and dividend-to-price deciles. These expected returns are computed using Proposition 8 with the parameters given by Table 1. The beta in the absence of noise is assumed to be 1.

Table 3: Beta Conditional on Size and Value Deciles

	Dividend-to-Price Ratio										
	All	1	2	3	4	5	6	7	8	9	10
All	1.005	0.971	0.984	0.991	0.997	1.002	1.007	1.012	1.018	1.025	1.040
Small-ME	1.019	0.984	0.998	1.005	1.011	1.016	1.021	1.026	1.032	1.040	1.054
ME-2	1.013	0.979	0.992	1.000	1.005	1.010	1.015	1.021	1.026	1.034	1.048
ME-3	1.010	0.976	0.990	0.997	1.002	1.007	1.012	1.018	1.023	1.031	1.045
ME-4	1.008	0.974	0.987	0.994	1.000	1.005	1.010	1.015	1.021	1.028	1.043
ME-5	1.006	0.972	0.985	0.992	0.998	1.003	1.008	1.013	1.019	1.026	1.041
ME-6	1.004	0.970	0.983	0.990	0.996	1.001	1.006	1.011	1.017	1.024	1.039
ME-7	1.002	0.968	0.981	0.988	0.994	0.999	1.004	1.009	1.015	1.022	1.037
ME-8	1.000	0.966	0.979	0.986	0.992	0.997	1.002	1.007	1.013	1.020	1.034
ME-9	0.997	0.963	0.976	0.983	0.989	0.994	0.999	1.004	1.010	1.017	1.031
Large-ME	0.991	0.957	0.971	0.978	0.984	0.989	0.993	0.999	1.004	1.012	1.026

This table presents beta of price (ME) and dividend-to-price deciles. The parameters are given by Table 1.

Assuming an annual riskfree return of 1.04, we can compute the abnormal return alpha, that is, the risk-adjusted excess expected return for each size and value decile with betas given in Table 3. We present alpha in Table 4. Small and value stocks have positive alpha while the large and glamor stocks have negative alpha. Stocks in the smallest decile have an alpha of 1.67% while those in the largest decile have an alpha of -1.18%. Similarly, stocks in the lowest dividend-price-ratio decile have an alpha of -0.98% while those in the highest dividend-price-ratio decile have an alpha of 1.47%. These two tables show that, in our model, small and value stocks have higher expected returns because they are under-valued due to negative price noise, not because there have higher betas.

One might wonder if these alphas persist over time. On the one hand, it is possible that alphas may be eliminated over time. On the other hand, it is possible that they will persist over time because of limits to arbitrage, associated with either transaction costs or risks in the strategies to explore these alphas.

As a model for the cross section of expected return, our paper is different from Berk (1995, 1997). The heterogeneity of expected return is mainly driven by the random realization of the noise, while it is specified in terms of the heterogeneity of the beta. Suppose that stock returns are identically distributed but correlated through systematic factors. In this case, there is no cross-section variation in expected returns and the correlation between price and the expected return will be zero, under Berk. By contrast, under our framework, a stock with a lower price still has a higher expected return. On the other hand, one can have an example where there is correlation between price and return but no conditional spreads.

The expected returns conditional on the price deciles in Propositions 4, 6, and 8 are state independent. It is possible that the size and value effects may be state dependent, for example, there are empirical studies documenting that the size and value spreads are different between booms and recessions. The most natural way to introduce the state dependence in our model is through the state-dependence of the conditional variance of noise. This can be potentially used to accommodate the dependence on business cycles of size and value effects.

Summers (1986) argues that "the data in conjunction with current methods provide no evidence against the view

Table 4: Alpha Conditional on Size and Value Deciles

	Dividend-to-Price Ratio										
	All	1	2	3	4	5	6	7	8	9	10
All	0.24	-1.43	-0.76	-0.41	-0.13	0.12	0.36	0.61	0.89	1.24	1.93
Small-ME	1.67	-0.03	0.65	1.01	1.29	1.55	1.79	2.05	2.33	2.69	3.40
ME-2	1.09	-0.59	0.08	0.44	0.71	0.96	1.21	1.45	1.74	2.10	2.79
ME-3	0.79	-0.89	-0.21	0.14	0.41	0.66	0.91	1.16	1.44	1.79	2.48
ME-4	0.55	-1.12	-0.45	-0.10	0.18	0.43	0.67	0.92	1.20	1.55	2.24
ME-5	0.34	-1.33	-0.66	-0.31	-0.03	0.22	0.46	0.71	0.99	1.34	2.03
ME-6	0.14	-1.53	-0.86	-0.51	-0.24	0.01	0.25	0.50	0.78	1.13	1.82
ME-7	-0.08	-1.74	-1.07	-0.72	-0.45	-0.20	0.04	0.29	0.57	0.92	1.60
ME-8	-0.31	-1.97	-1.30	-0.96	-0.68	-0.44	-0.20	0.05	0.33	0.68	1.37
ME-9	-0.61	-2.26	-1.60	-1.25	-0.98	-0.73	-0.49	-0.24	0.03	0.38	1.07
Large-ME	-1.18	-2.84	-2.17	-1.82	-1.55	-1.30	-1.06	-0.81	-0.54	-0.18	0.50

This table presents annual alpha, in percentage, of price (ME) and dividend-to-price deciles. The parameters are given by Table 1 and the gross riskfree return is assumed to be 1.04.

that financial market prices deviate widely and frequently from rational valuations." We would like to argue that the size and value effects are evidence for the view that financial market prices deviate from values.

7 Conditioning on Past Prices

In previous sections, we have studied the expected return, conditional on current prices and/or price ratios. In this section, we will study the expected returns conditional on both current and past prices. We can also compute the expected return conditional on past price-ratios as well; we choose prices to be the conditioning variables for notational simplicity.

We first consider the expected return conditional on past return e^{r_t}. That is, we are interested in the mean of $\frac{P_{t+1}}{P_t}$ conditional on the previous period return[4] $\frac{P_t}{P_{t-1}} = e^{r_t}$. A high return r_t implies a high Δ_t and low Δ_{t-1} on average, thus lower expected return for $t+1$. This is the return reversal effect.

Proposition 9 (Conditioning on Return) *If Assumptions 1, 2, and 3 hold, the expected return at time $t+1$ conditional on return R_t is,*

$$\mathrm{E}\left[\frac{P_{t+1}+D_{t+1}}{P_t}|r_t\right] = e^{\mu+\frac{1}{2}\sigma_r^2}\left(e^{\frac{\sigma^2_{\epsilon_\Delta}}{1+\rho}}\frac{R_t^{-(1-\rho)\gamma_5}}{\mathrm{E}\left[R_t^{-(1-\rho)\gamma_5}\right]} + e^{-\bar{x}_v+\frac{\sigma^2_{\epsilon_x}}{2(1-\rho_x^2)}+\frac{\sigma^2_{\epsilon_\Delta}}{1-\rho^2}}\frac{R_t^{-\gamma_5}}{\mathrm{E}\left[R_t^{-\gamma_5}\right]}\right),$$

where $\gamma_5 = \frac{-\frac{1}{1+\rho}\sigma^2_{\epsilon_\Delta} r_t}{\sigma^2_{\epsilon_\Delta}+\sigma_r^2+(1-\rho)^2\frac{\sigma^2_{\epsilon_\Delta}}{1-\rho^2}}$. The conditional expected return decreases with r_t for $\rho < 1$.

The proof is given in the Appendix.

According to Proposition 9, a mean-reverting noise lead to return reversal. That is, the expected return, conditional on past return, decreases with the past return. In the US market data, return reversal is observed for horizons greater than 2 years (DeBondt and Thaler (1985, 1987) and Chopra, Lakonishok and Ritter (1992)). However, return momentum, which means that the expected return increases with the past return, is observed for horizons less than 1 year (Jegadeesh and Titman, (1993, 2001)). Thus the observed expected return conditional past return cannot be explained by mean-reverting noise, at least for horizon less than 1 year.

Note that conditioning on return P_t/P_{t-1} is different from conditioning on past prices P_t and P_{t-1} separately, which we turn to next.

So far, we have conditioned on current prices or price ratios to produce size and value effects and on past return $\frac{P_t}{P_{t-1}}$ to produce momentum and reversal effects. However, it is obvious that one should use the full price history. We now consider the time $t+1$ expected return conditional on past market prices, P_s, for $s = t, t-1, ..., t_0$. Our analysis can be extended to include past price-ratios. We only present the case for past prices for ease of exposition.

[4] Strictly speaking, the previous-period return should be $\frac{P_t+D_t}{P_{t-1}}$. However, we do not have the closed form solution for the inference of Δ_t. Nevertheless, the intuition still applies.

We would like to compute,

$$\mathrm{E}\left[\frac{P_{t+1}+D_{t+1}}{P_t}|\{p_s\}_{t_0}^t\right], \tag{32}$$

where $t_0 \leq t$. That is, the expected return from time t to $t+1$ conditional on prices from t_0 to t. We will need to additionally specify the prior distribution for v_{t_0} and Δ_{t_0}. We assume that Δ_{t_0} is drawn from the unconditional distribution of Δ_t, which has a mean of 0 and variance of $\frac{\sigma^2_{\epsilon_\Delta}}{1-\rho^2}$. We assume that v_{t_0} is drawn from a normal distribution with a mean $\bar{v}_{t_0}$ and $\sigma^2_{vt_0}$. We assume that v_{t_0} and Δ_{t_0} are independent in the prior distribution.

Proposition 10 (Conditioning on Current and Past Prices) *Suppose Assumptions 1, 2, and 3 hold. Furthermore assume that*

$$\frac{1}{\sigma^2_{vt_0}} = \frac{1}{\bar{\sigma}^2} - \frac{1-\rho^2}{\sigma^2_{\epsilon_\Delta}},$$

where $\bar{\sigma}^2 = \sqrt{\sigma^2_{\epsilon_\Delta} + \frac{(1+\rho)^2}{4}\sigma_r^2 \frac{\sigma_r}{1-\rho}} - \frac{1+\rho}{2(1-\rho)}\sigma_r^2$. *Then the expected return at time t conditional the prices from $t_0 \leq t$ to t is*

$$\begin{aligned}
& \mathrm{E}\left[\frac{P_{t+1}+D_{t+1}}{P_t}|r_t\right] \\
= \quad & e^{\mu+\frac{1}{2}\sigma_r^2}\left(e^{\frac{\sigma^2_{\epsilon_\Delta}}{1+\rho}}\frac{\left(P_t^{h_p}\left(\prod_{s=t_0+1}^{t-1} P_s^{-h_e^{t-s-1}(1-h_e)h_p}\right)P_{t_0}^{-h_e^{t-t_0-1}h_{t_0}}\right)^{-(1-\rho)}}{\mathrm{E}\left[\left(P_t^{h_p}\left(\prod_{s=t_0+1}^{t-1} P_s^{-h_e^{t-s-1}(1-h_e)h_p}\right)P_{t_0}^{-h_e^{t-t_0-1}h_{t_0}}\right)^{-(1-\rho)}\right]}\right. \\
+ \quad & \left. e^{-\bar{x}_v+\frac{\sigma^2_{\epsilon_x}}{2(1-\rho_x^2)}+\frac{\sigma^2_{\epsilon_\Delta}}{1-\rho^2}}\frac{P_t^{-h_p}\left(\prod_{s=t_0+1}^{t-1} P_s^{h_e^{t-s-1}(1-h_e)h_p}\right)P_{t_0}^{h_e^{t-t_0-1}h_{t_0}}}{\mathrm{E}\left[P_t^{-h_p}\left(\prod_{s=t_0+1}^{t-1} P_s^{h_e^{t-s-1}(1-h_e)h_p}\right)P_{t_0}^{h_e^{t-t_0-1}h_{t_0}}\right]}\right),
\end{aligned} \tag{33}$$

where $h_p = \frac{\sigma^2_{\epsilon_\Delta}-\rho(1-\rho)\bar{\sigma}^2}{\sigma_r^2+\sigma^2_{\epsilon_\Delta}+(1-\rho)^2\bar{\sigma}^2}$, $h_e = \frac{\rho\sigma_r^2+\sigma^2_{\epsilon_\Delta}}{\sigma_r^2+\sigma^2_{\epsilon_\Delta}+(1-\rho)^2\bar{\sigma}^2}$, *and* $h_{t_0} = h_e\left(1-\frac{1-\rho^2}{\sigma^2_{\epsilon_\Delta}}\bar{\sigma}^2 - h_p\right)$.

Again, we include the proof in the Appendix. We use the convention for the product operator $\prod_{s=i}^j$ that the product is 1 if the upper index j is smaller than the lower index i.

According to Equation (33), the conditional expected return decreases with the current price but increases with past prices. Note that $h_e < 1$; past prices are discounted by powers of h_e in the conditional expected returns, the further away in the past, the higher the discount and the lower the relevance to next period return.

In general, the variance of Δ_t conditional on past prices depends on t_0. However, when $t - t_0 \to \infty$, this variance goes to a constant, which can be shown to be $\bar{\sigma}^2$. The technical condition at the beginning of the proposition implies that the conditional variance reaches $\bar{\sigma}^2$ at time t_0 and is assumed only to simplify the notation. In the Appendix, we show results for the general case.

Appendix

The following lemma is special case studied in Liptser and Shiryaev (1977).

Lemma 1 *Suppose that θ is a vector of normal random variables with the mean vector $\bar{\theta}$ and the variance-covariance matrix Σ_θ. Furthermore, a vector of random variables ξ satisfies*

$$\xi = A_0 + A_1\theta + B\epsilon,$$

where ϵ is a vector of standard normal random variables that are independent of θ. Assuming that $A_1\Sigma_\theta A_1' + BB'$ is invertible. Then mean vector $\mathrm{E}[\theta|\xi]$ of θ conditional on ξ and the variance-covariance matrix $\Sigma_{\theta|\xi}$ conditional on ξ are

$$\mathrm{E}[\theta|\xi] = \bar{\theta} + \Sigma_\theta A_1'(A_1\Sigma_\theta A_1' + BB')^{-1}(\xi - A_0 - A_1\bar{\theta}).$$

and

$$\Sigma_{\theta|\xi} = \Sigma_\theta - \Sigma_\theta A_1'(A_1\Sigma_\theta A_1' + BB')^{-1}A_1\Sigma_\theta.$$

We will apply this lemma repeatedly. In our applications, θ will be the noise Δ_t, ξ will be the price p_t or the price-dividend ratio $p_t - d_t$, and ϵ will be the other random variables such as ϵ_{rt} (or F_t later in the Appendix).

Proof of Proposition 3

Proof. Note that

$$p_t = v_t + \Delta_t - \ln(\mathrm{E}[e^{\Delta_t}]).$$

We will assume that without information, v_t is normal with mean of $\bar{v}_t$ and variance σ_{vt}^2, the distribution of Δ_t is its unconditional distribution of mean 0 and variance $\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}$. v_t and Δ_t is independent, as assumed. Lemma 1 in the appendix implies that conditional on p_t, the mean of Δ_t is

$$\mathrm{E}[\Delta_t|p_t] = \frac{\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}\left(p_t - \bar{v}_t + \ln(\mathrm{E}[e^{\Delta_t}])\right)}{\sigma_{vt}^2 + \frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}};$$

and the variance is

$$\frac{\sigma_{vt}^2\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}{\sigma_{vt}^2 + \frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}.$$

Therefore, we get

$$\begin{aligned}
\mathrm{E}[e^{-(1-\rho)\Delta_t}|p_t] &= e^{-(1-\rho)\frac{\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}\left(p_t-\bar{v}_t+\ln(\mathrm{E}[e^{\Delta_t}])\right)}{\sigma_{vt}^2+\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}+(1-\rho)^2\frac{\sigma_{vt}^2\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}{2\left(\sigma_{vt}^2+\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}\right)}} \\
&= e^{\frac{(1-\rho)\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}\left(\bar{v}_t-\ln(\mathrm{E}[e^{\Delta_t}])\right)+(1-\rho)^2\frac{\sigma_{vt}^2\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}{2}}{\sigma_{vt}^2+\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}} e^{-\frac{(1-\rho)\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}{\sigma_{vt}^2+\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}p_t} \\
&= e^{\frac{(1-\rho)\sigma_{\epsilon_\Delta}^2\left(\bar{v}_t-\ln(\mathrm{E}[e^{\Delta_t}])\right)+(1-\rho)^2\frac{\sigma_{vt}^2\sigma_{\epsilon_\Delta}^2}{2}}{(1-\rho^2)\sigma_{vt}^2+\sigma_{\epsilon_\Delta}^2}} e^{-\frac{(1-\rho)\sigma_{\epsilon_\Delta}^2}{(1-\rho^2)\sigma_{vt}^2+\sigma_{\epsilon_\Delta}^2}p_t}.
\end{aligned}$$

Finally,

$$\mathrm{E}\left[\frac{P_{t+1}}{P_t}|p_t\right] = e^{\mu+\frac{1}{2}(\sigma_r^2+\sigma_{\epsilon_\Delta}^2)}\mathrm{E}\left[e^{-(1-\rho)\Delta_t}|p_t\right] = e^{\mu+\frac{1}{2}(\sigma_r^2+\sigma_{\epsilon_\Delta}^2)} e^{\frac{(1-\rho)^2\frac{\sigma_{vt}^2\sigma_{\epsilon_\Delta}^2}{2}}{(1-\rho^2)\sigma_{vt}^2+\sigma_{\epsilon_\Delta}^2}} e^{-\frac{(1-\rho)\sigma_{\epsilon_\Delta}^2}{(1-\rho^2)\sigma_v^2+\sigma_{\epsilon_\Delta}^2}(p_t-\bar{p}_t)}.$$

From

$$v_{t+1} - d_{t+1} = (1-\rho_x)\bar{x}_v + \rho_x(v_t - d_t) + \sigma_{\epsilon_x}\epsilon_{xt+1},$$

we get

$$\mathrm{E}\left[\frac{D_{t+1}}{V_{t+1}}\right] = \mathrm{E}\left[e^{-(v_{t+1}-d_{t+1})}\right] = e^{-\bar{x}_v+\frac{\sigma_{\epsilon_x}^2}{2(1-\rho_x^2)}}.$$

From the assumption that v_t, $v_t - d_t$, and Δ_t are independent, we get

$$\begin{aligned}
&\mathrm{E}\left[\frac{D_{t+1}}{P_t}|p_t\right] = \mathrm{E}\left[\frac{V_{t+1}}{P_t}|p_t\right]\mathrm{E}\left[\frac{D_{t+1}}{V_{t+1}}\right] = \mathrm{E}\left[\frac{V_{t+1}}{V_t}e^{-\Delta_t}|p_t\right]\mathrm{E}\left[\frac{D_{t+1}}{V_{t+1}}\right] \\
&= \mathrm{E}\left[\frac{V_{t+1}}{V_t}\right]\mathrm{E}\left[\frac{D_{t+1}}{V_{t+1}}\right]\mathrm{E}\left[e^{-\Delta_t}|p_t\right] \\
&= e^{\mu+\frac{1}{2}\sigma_r^2}e^{-\bar{x}_v+\frac{\sigma_{\epsilon_x}^2}{2(1-\rho_x^2)}}e^{\frac{\frac{\sigma_{vt}^2\sigma_{\epsilon_\Delta}^2}{2}}{(1-\rho^2)\sigma_{vt}^2+\sigma_{\epsilon_\Delta}^2}}e^{-\frac{\sigma_{\epsilon_\Delta}^2}{(1-\rho^2)\sigma_v^2+\sigma_{\epsilon_\Delta}^2}(p_t-\bar{p}_t)}.
\end{aligned}$$

We get

$$\begin{aligned}
&\mathrm{E}\left[\frac{P_{t+1}+D_{t+1}}{P_t}|p_t\right] = \mathrm{E}\left[\frac{P_{t+1}}{P_t}|p_t\right] + \mathrm{E}\left[\frac{D_{t+1}}{P_t}|p_t\right] \\
&= e^{\mu+\frac{1}{2}(\sigma_r^2+\sigma_{\epsilon_\Delta}^2)}e^{\frac{(1-\rho)^2\frac{\sigma_{vt}^2\sigma_{\epsilon_\Delta}^2}{2}}{(1-\rho^2)\sigma_{vt}^2+\sigma_{\epsilon_\Delta}^2}}e^{-\frac{(1-\rho)\sigma_{\epsilon_\Delta}^2}{(1-\rho^2)\sigma_v^2+\sigma_{\epsilon_\Delta}^2}(p_t-\bar{p}_t)} \\
&+e^{\mu+\frac{1}{2}\sigma_r^2}e^{-\bar{x}_v+\frac{\sigma_{\epsilon_x}^2}{2(1-\rho_x^2)}}e^{\frac{\frac{\sigma_{vt}^2\sigma_{\epsilon_\Delta}^2}{2}}{(1-\rho^2)\sigma_{vt}^2+\sigma_{\epsilon_\Delta}^2}}e^{-\frac{\sigma_{\epsilon_\Delta}^2}{(1-\rho^2)\sigma_v^2+\sigma_{\epsilon_\Delta}^2}(p_t-\bar{p}_t)}.
\end{aligned}$$

The above equation can be expressed in terms of P_t using the definition of $P_t = e^{p_t}$. It is straightforward to evaluate $\mathrm{E}\left[e^{-\frac{(1-\rho)\sigma_{\epsilon_\Delta}^2}{(1-\rho^2)\sigma_v^2+\sigma_{\epsilon_\Delta}^2}p_t}\right]$ and $\mathrm{E}\left[e^{-\frac{\sigma_{\epsilon_\Delta}^2}{(1-\rho^2)\sigma_v^2+\sigma_{\epsilon_\Delta}^2}p_t}\right]$ and prove the equivalence between the above equation implies the equation given in the proposition.

Proof of Proposition 5

At time t,

$$x_t = (v_t - d_t) + \Delta_t - \ln(\mathrm{E}[e^{\Delta_t}]).$$

We assume that $v_t - d_t$ and Δ_t are both drawn from the stationary distribution, under which $v_t - d_t$ is normal with a mean of $\bar{x}_v$ and a variance of $\frac{\sigma_{\epsilon_x}^2}{1-\rho_x^2}$ and is independent of Δ_t and Δ_t is normal with a mean of 0 and a variance of $\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho_e^2}$.

Therefore, conditional on x_t, the mean of Δ_t is

$$\frac{(1-\rho_x^2)\sigma_{\epsilon_\Delta}^2}{(1-\rho_x^2)\sigma_{\epsilon_\Delta}^2+(1-\rho^2)\sigma_{\epsilon_x}^2}\Big(x_t - \bar{x}\Big),$$

where $\bar{x} = \bar{x}_v - \ln(\mathrm{E}[e^{\Delta_t}])$ is the unconditional mean of x, and the variance is

$$\frac{\sigma_{\epsilon_x}^2\sigma_{\epsilon_\Delta}^2}{(1-\rho_x^2)\sigma_{\epsilon_\Delta}^2+(1-\rho^2)\sigma_{\epsilon_x}^2}.$$

Thus, we get

$$\begin{aligned}
\mathrm{E}\left[e^{-(1-\rho)\Delta_t}|x_t\right] &= e^{-(1-\rho)\frac{(1-\rho_x^2)\sigma_{\epsilon_\Delta}^2}{(1-\rho_x^2)\sigma_{\epsilon_\Delta}^2+(1-\rho^2)\sigma_{\epsilon_x}^2}(x_t-\bar{x})}e^{\frac{(1-\rho)^2}{2}\frac{\sigma_{\epsilon_x}^2\sigma_{\epsilon_\Delta}^2}{(1-\rho_x^2)\sigma_{\epsilon_\Delta}^2+(1-\rho^2)\sigma_{\epsilon_x}^2}} \\
&= e^{\frac{(1-\rho)(1-\rho_x^2)\sigma_{\epsilon_\Delta}^2\bar{x}+\frac{(1-\rho)^2}{2}\sigma_{\epsilon_x}^2\sigma_{\epsilon_\Delta}^2}{(1-\rho_x^2)\sigma_{\epsilon_\Delta}^2+(1-\rho^2)\sigma_{\epsilon_x}^2}}e^{-\frac{(1-\rho)(1-\rho_x^2)\sigma_{\epsilon_\Delta}^2}{(1-\rho_x^2)\sigma_{\epsilon_\Delta}^2+(1-\rho^2)\sigma_{\epsilon_x}^2}x_t}.
\end{aligned}$$

The first equality of the equation in the proposition obtains by noting that

$$\mathrm{E}\left[\frac{P_{t+1}}{P_t}|x_t\right] = e^{\mu+\frac{1}{2}(\sigma_r^2+\sigma_{\epsilon_\Delta}^2)}\mathrm{E}\left[e^{-(1-\rho)\Delta_t}|x_t\right].$$

The second equality follows from the definition of $X_t = e^{x_t}$. From

$$v_{t+1} - d_{t+1} = (1-\rho_x)\bar{x}_v + \rho_x(v_t - d_t) + \sigma_{\epsilon_x}\epsilon_{t+1}^x,$$

we get

$$
\begin{aligned}
\mathrm{E}\left[\frac{D_{t+1}}{P_t}|x_t\right] &= \mathrm{E}\left[\frac{D_{t+1}}{V_t e^{\Delta_t-\ln(\mathrm{E}[e^{\Delta_t}])}}|x_t\right] = \mathrm{E}\left[\frac{V_{t+1}}{V_t}\frac{D_{t+1}}{V_{t+1}}e^{-\Delta_t+\ln(\mathrm{E}[e^{\Delta_t}])}|x_t\right] \\
&= \mathrm{E}\left[e^{\mu+\sigma_r\epsilon_{rt+1}-(1-\rho_x)\bar{x}_v-\rho_x(v_t-d_t)-\sigma_{\epsilon_x}\epsilon^x_{t+1}-\Delta_t+\ln(\mathrm{E}[e^{\Delta_t}])}|x_t\right] \\
&= e^{\mu+\frac{1}{2}(\sigma_r^2+\sigma_{\epsilon_x}^2)-(1-\rho_x)\bar{x}_v+\ln(\mathrm{E}[e^{\Delta_t}])}\mathrm{E}\left[e^{-\rho_x(v_t-d_t)-\Delta_t}|x_t\right] \\
&= e^{\mu+\frac{1}{2}(\sigma_r^2+\sigma_{\epsilon_x}^2)-(1-\rho_x)\bar{x}_v+\ln(\mathrm{E}[e^{\Delta_t}])}\mathrm{E}\left[e^{-\rho_x(x_t-\Delta_t+\ln(\mathrm{E}[e^{\Delta_t}]))-\Delta_t}|x_t\right] \\
&= e^{\mu+\frac{1}{2}(\sigma_r^2+\sigma_{\epsilon_x}^2)-\rho_x x_t-(1-\rho_x)\bar{x}}\mathrm{E}\left[e^{-(1-\rho_x)\Delta_t}|x_t\right].
\end{aligned}
$$

Finally,

$$
\mathrm{E}\left[\frac{P_{t+1}+D_{t+1}}{P_t}|x_t\right] = e^{\mu+\frac{1}{2}\sigma_r^2}\left(e^{\frac{1}{2}\sigma_{\epsilon_\Delta}^2}\mathrm{E}\left[e^{-(1-\rho)\Delta_t}|x_t\right] + e^{\frac{1}{2}\sigma_{\epsilon_x}^2-\rho_x(x_t+\ln(\mathrm{E}[e^{\Delta_t}]))}\mathrm{E}\left[e^{-(1-\rho_x)\Delta_t}|x_t\right]\right)
$$

$$
\begin{aligned}
= \quad & e^{\mu+\frac{1}{2}(\sigma_r^2+\sigma_{\epsilon_\Delta}^2)}e^{\frac{(1-\rho)(1-\rho_x^2)\sigma_{\epsilon_\Delta}^2\bar{x}+\frac{(1-\rho)^2}{2}\sigma_{\epsilon_x}^2\sigma_{\epsilon_\Delta}^2}{(1-\rho_x^2)\sigma_{\epsilon_\Delta}^2+(1-\rho^2)\sigma_{\epsilon_x}^2}}e^{-\frac{(1-\rho)(1-\rho_x^2)\sigma_{\epsilon_\Delta}^2}{(1-\rho_x^2)\sigma_{\epsilon_\Delta}^2+(1-\rho^2)\sigma_{\epsilon_x}^2}x_t} \\
& +e^{\mu+\frac{1}{2}(\sigma_r^2+\sigma_{\epsilon_x}^2)-(1-\rho_x)\bar{x}}e^{\frac{(1-\rho_x)(1-\rho_x^2)\sigma_{\epsilon_\Delta}^2\bar{x}+\frac{(1-\rho_x)^2}{2}\sigma_{\epsilon_x}^2\sigma_{\epsilon_\Delta}^2}{(1-\rho_x^2)\sigma_{\epsilon_\Delta}^2+(1-\rho^2)\sigma_{\epsilon_x}^2}}e^{-\frac{(1-\rho_x)(1-\rho_x^2)\sigma_{\epsilon_\Delta}^2}{(1-\rho_x^2)\sigma_{\epsilon_\Delta}^2+(1-\rho^2)\sigma_{\epsilon_x}^2}x_t-\rho_x x_t}.
\end{aligned}
$$

It is straightforward to evaluate the expectations in the proposition and prove the equivalence between the above equation implies the equation given in the proposition.

Proof of Proposition 7

At time t, we have two signals on Δ_t,

$$
\begin{aligned}
p_t &= v_t + \Delta_t - \ln(\mathrm{E}[e^{\Delta_t}]); \\
x_t &= (v_t - d_t) + \Delta_t - \ln(\mathrm{E}[e^{\Delta_t}]).
\end{aligned}
$$

Note that v_t, $v_t - d_t$, and Δ_t are have a distribution of normal with mean $(\bar{v}_t, \bar{x}_v, 0)$ and a diagonal covariance matrix with diagonal covariance matrix element of $\left(\sigma_{vt}^2, \frac{\sigma_{\epsilon_x}^2}{1-\rho_x^2}, \frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}\right)$. We can express the above equation as

$$
\begin{aligned}
p_t - \bar{v}_t + \ln(\mathrm{E}[e^{\Delta_t}]) &= (v_t - \bar{v}_t) + \Delta_t; \\
x_t - \bar{x} &= (v_t - d_t - \bar{x}_v) + \Delta_t.
\end{aligned}
$$

Therefore, conditional on p_t and x_t, the mean of Δ_t is

$$
\frac{\frac{1}{\sigma_{vt}^2}(p_t - \bar{p}_t) + \frac{1-\rho_x^2}{\sigma_{\epsilon_x}^2}(x_t - \bar{x})}{\frac{1}{\sigma_{vt}^2} + \frac{1-\rho_x^2}{\sigma_{\epsilon_x}^2} + \frac{1-\rho^2}{\sigma_{\epsilon_\Delta}^2}}.
$$

and the variance is

$$
\frac{1}{\frac{1}{\sigma_{vt}^2} + \frac{1-\rho_x^2}{\sigma_{\epsilon_x}^2} + \frac{1-\rho^2}{\sigma_{\epsilon_\Delta}^2}}.
$$

Thus

$$
\begin{aligned}
\mathrm{E}\left[e^{-(1-\rho)\Delta_t}|p_t, x_t\right] &= e^{-(1-\rho)\frac{\frac{1}{\sigma_{vt}^2}(p_t-\bar{p}_t)+\frac{1-\rho_x^2}{\sigma_{\epsilon_x}^2}(x_t-\bar{x})}{\frac{1}{\sigma_{vt}^2}+\frac{1-\rho_x^2}{\sigma_{\epsilon_x}^2}+\frac{1-\rho^2}{\sigma_{\epsilon_\Delta}^2}}}e^{\frac{(1-\rho)^2}{2\left(\frac{1}{\sigma_{vt}^2}+\frac{1-\rho_x^2}{\sigma_{\epsilon_x}^2}+\frac{1-\rho^2}{\sigma_{\epsilon_\Delta}^2}\right)}} \\
&= e^{\frac{(1-\rho)\left(\frac{1}{\sigma_{vt}^2}\bar{p}_t+\frac{1-\rho_x^2}{\sigma_{\epsilon_x}^2}\bar{x}\right)+\frac{(1-\rho)^2}{2}}{\frac{1}{\sigma_{vt}^2}+\frac{1-\rho_x^2}{\sigma_{\epsilon_x}^2}+\frac{1-\rho^2}{\sigma_{\epsilon_\Delta}^2}}}e^{\frac{\frac{-(1-\rho)}{\sigma_{vt}^2}p_t}{\frac{1}{\sigma_{vt}^2}+\frac{1-\rho_x^2}{\sigma_{\epsilon_x}^2}+\frac{1-\rho^2}{\sigma_{\epsilon_\Delta}^2}}}e^{\frac{-(1-\rho)\frac{1-\rho_x^2}{\sigma_{\epsilon_x}^2}x_t}{\frac{1}{\sigma_{vt}^2}+\frac{1-\rho_x^2}{\sigma_{\epsilon_x}^2}+\frac{1-\rho^2}{\sigma_{\epsilon_\Delta}^2}}}.
\end{aligned}
$$

The first equality of the equation in the proposition obtains by noting that

$$
\mathrm{E}\left[\frac{P_{t+1}}{P_t}|x_t, p_t\right] = e^{\mu+\frac{1}{2}(\sigma_r^2+\sigma_{\epsilon_\Delta}^2)}\mathrm{E}\left[e^{-(1-\rho)\Delta_t}|x_t, p_t\right].
$$

The second equality follows from the definitions $P_t = e^{p_t}$ and $X_t = e^{x_t}$. Note that

$$v_{t+1} - d_{t+1} = (1-\rho_x)\bar{x}_v + \rho_x(v_t - d_t) + \sigma_{\epsilon_x}\epsilon^x_{t+1},$$

$$\begin{aligned}
\mathrm{E}\left[\frac{D_{t+1}}{P_t}|x_t\right] &= \mathrm{E}\left[\frac{D_{t+1}}{V_t e^{\Delta_t - \ln(\mathrm{E}[e^{\Delta_t}])}}|x_t\right] = \mathrm{E}\left[\frac{V_{t+1}}{V_t}\frac{D_{t+1}}{V_{t+1}}e^{-\Delta_t + \ln(\mathrm{E}[e^{\Delta_t}])}|p_t, x_t\right] \\
&= e^{\mu + \frac{1}{2}(\sigma_r^2 + \sigma_{\epsilon_x}^2) - (1-\rho_x)\bar{x} - \rho_x x_t}\mathrm{E}\left[e^{-(1-\rho_x)\Delta_t}|p_t, x_t\right].
\end{aligned}$$

Finally,

$$\begin{aligned}
&\mathrm{E}\left[\frac{P_{t+1}+D_{t+1}}{P_t}|x_t\right] = \mathrm{E}\left[\frac{P_{t+1}}{P_t}|x_t\right] + \mathrm{E}\left[\frac{D_{t+1}}{V_t e^{\Delta_t}}|x_t\right] \\
= \;& e^{\mu+\frac{1}{2}(\sigma_r^2+\sigma_{\epsilon_\Delta}^2)}\mathrm{E}\left[e^{-(1-\rho)\Delta_t}|x_t, p_t\right] + e^{\mu+\frac{1}{2}(\sigma_r^2+\sigma_{\epsilon_x}^2)-(1-\rho_x)\bar{x}-\rho_x x_t}\mathrm{E}\left[e^{-(1-\rho_x)\Delta_t}|p_t, x_t\right] \\
= \;& e^{\mu+\frac{1}{2}(\sigma_r^2+\sigma_{\epsilon_\Delta}^2)} e^{\frac{\frac{(1-\rho)^2}{2}}{\frac{1}{\sigma_{vt}^2}+\frac{1-\rho_x^2}{\sigma_{\epsilon_x}^2}+\frac{1-\rho^2}{\sigma_{\epsilon_\Delta}^2}}} e^{\frac{\frac{-(1-\rho)}{\sigma_{vt}^2}(p_t-\bar{p}_t)}{\frac{1}{\sigma_{vt}^2}+\frac{1-\rho_x^2}{\sigma_{\epsilon_x}^2}+\frac{1-\rho^2}{\sigma_{\epsilon_\Delta}^2}}} e^{\frac{\frac{-(1-\rho)(1-\rho_x^2)}{\sigma_{\epsilon_x}^2}(x_t-\bar{x})}{\frac{1}{\sigma_{vt}^2}+\frac{1-\rho_x^2}{\sigma_{\epsilon_x}^2}+\frac{1-\rho^2}{\sigma_{\epsilon_\Delta}^2}}} \\
&+ e^{\mu+\frac{1}{2}(\sigma_r^2+\sigma_{\epsilon_x}^2)-(1-\rho_x)\bar{x}-\rho_x x_t} e^{\frac{\frac{(1-\rho_x)^2}{2}}{\frac{1}{\sigma_{vt}^2}+\frac{1-\rho_x^2}{\sigma_{\epsilon_x}^2}+\frac{1-\rho^2}{\sigma_{\epsilon_\Delta}^2}}} e^{\frac{\frac{-(1-\rho_x)}{\sigma_{vt}^2}(p_t-\bar{p}_t)}{\frac{1}{\sigma_{vt}^2}+\frac{1-\rho_x^2}{\sigma_{\epsilon_x}^2}+\frac{1-\rho^2}{\sigma_{\epsilon_\Delta}^2}}} e^{\frac{\frac{-(1-\rho_x)(1-\rho_x^2)}{\sigma_{\epsilon_x}^2}(x_t-\bar{x})}{\frac{1}{\sigma_{vt}^2}+\frac{1-\rho_x^2}{\sigma_{\epsilon_x}^2}+\frac{1-\rho^2}{\sigma_{\epsilon_\Delta}^2}}} \\
= \;& e^{\mu+\frac{1}{2}(\sigma_r^2+\sigma_{\epsilon_\Delta}^2)} e^{\frac{(1-\rho)\left(\frac{1}{\sigma_{vt}^2}\bar{p}_t+\frac{1-\rho_x^2}{\sigma_{\epsilon_x}^2}\bar{x}\right)+\frac{(1-\rho)^2}{2}}{\frac{1}{\sigma_{vt}^2}+\frac{1-\rho_x^2}{\sigma_{\epsilon_x}^2}+\frac{1-\rho^2}{\sigma_{\epsilon_\Delta}^2}}} P_t^{\frac{\frac{-(1-\rho)}{\sigma_{vt}^2}}{\frac{1}{\sigma_{vt}^2}+\frac{1-\rho_x^2}{\sigma_{\epsilon_x}^2}+\frac{1-\rho^2}{\sigma_{\epsilon_\Delta}^2}}} X_t^{\frac{\frac{-(1-\rho)(1-\rho_x^2)}{\sigma_{\epsilon_x}^2}}{\frac{1}{\sigma_{vt}^2}+\frac{1-\rho_x^2}{\sigma_{\epsilon_x}^2}+\frac{1-\rho^2}{\sigma_{\epsilon_\Delta}^2}}} \\
&+ e^{\mu+\frac{1}{2}(\sigma_r^2+\sigma_{\epsilon_x}^2)-(1-\rho_x)\bar{x}} e^{\frac{(1-\rho_x)\left(\frac{1}{\sigma_{vt}^2}\bar{p}_t+\frac{1-\rho_x^2}{\sigma_{\epsilon_x}^2}\bar{x}\right)+\frac{(1-\rho_x)^2}{2}}{\frac{1}{\sigma_{vt}^2}+\frac{1-\rho_x^2}{\sigma_{\epsilon_x}^2}+\frac{1-\rho^2}{\sigma_{\epsilon_\Delta}^2}}} P_t^{\frac{\frac{-(1-\rho_x)}{\sigma_{vt}^2}}{\frac{1}{\sigma_{vt}^2}+\frac{1-\rho_x^2}{\sigma_{\epsilon_x}^2}+\frac{1-\rho^2}{\sigma_{\epsilon_\Delta}^2}}} X_t^{\frac{\frac{-(1-\rho_x)(1-\rho_x^2)}{\sigma_{\epsilon_x}^2}}{\frac{1}{\sigma_{vt}^2}+\frac{1-\rho_x^2}{\sigma_{\epsilon_x}^2}+\frac{1-\rho^2}{\sigma_{\epsilon_\Delta}^2}}-\rho_x}.
\end{aligned}$$

It is straightforward to evaluate the expectations in the proposition and prove the equivalence between the above equation implies the equation given in the proposition.

Proof of Proposition 8

Let $\sigma_{xt} = \sqrt{\frac{1-\rho_x^2}{\sigma_{\epsilon_x}^2} + \frac{1-\rho^2}{\sigma_{\epsilon_\Delta}^2}}$. Without loss of generality, we can assume that the means of p_t and x_t are zero. We need to compute

$$\mathrm{E}[e^{-(\phi_1 p_t + \phi_2 x_t)}|R_1]$$

where $R_1 = \{\sigma_{pt}\delta_i \leq p_t \leq \sigma_{pt}\delta_{i+1}, \sigma_{xt}\delta_{i,j} \leq x_t \leq \sigma_{xt}\delta_{i,j+1}\}$, for various ϕ_1 and ϕ_2. Define q and z by the following equations.

$$\begin{aligned}
p_t &= \sqrt{1-\hat{\rho}^2}\sigma_{pt}q + \hat{\rho}\sigma_{pt}z, \\
x_t &= \sigma_{xt}z.
\end{aligned}$$

Using the fact that p_t and x_t have variances of σ_{pt}^2 and σ_{xt}^2 and covariance of $\hat{\rho}\sigma_{pt}\sigma_{xt}$, we can show that q and z are independent standard normals. By changing the variable from (p_t, x_t) to (q, z), we get,

$$\begin{aligned}
&\mathrm{E}[e^{-(\phi_1 p_t + \phi_2 x_t)}|R_1] = \mathrm{E}[e^{-(\phi_1(\sqrt{1-\rho^2}\sigma_{pt}q + \rho\sigma_{pt}z) + \phi_2\sigma_{xt}z)}|R_2] \\
= \;& \mathrm{E}[e^{-\phi_1\sqrt{1-\rho^2}\sigma_{pt}q - (\phi_1\rho\sigma_{pt} + \phi_2\sigma_{xt})z}|R_2],
\end{aligned}$$

where $R_2 = \{\delta_i \leq \sqrt{1-\rho^2}q + \rho z \leq \delta_{i+1}, \delta_{i,j} \leq z \leq \delta_{i,j+1}\}$. Integrating out q, we get,

$$\mathrm{E}[e^{-\phi_1\sqrt{1-\rho^2}\sigma_{pt}q - (\phi_1\rho\sigma_{pt}+\phi_2\sigma_{xt})z}|R_2] = e^{\frac{1}{2}\phi_1^2\sigma_{pt}^2(1-\rho^2)}\mathrm{E}[e^{-(\phi_1\rho\sigma_{pt}+\phi_2\sigma_{xt})z}(N(x_1) - N(x_2))|R_3],$$

where $x_1 = \frac{\delta_{i+1} - \rho z}{\sqrt{1-\rho^2}} + \phi_1\sqrt{1-\rho^2}\sigma_{pt}$, $x_2 = \frac{\delta_i - \rho z}{\sqrt{1-\rho^2}} + \phi_1\sqrt{1-\rho^2}\sigma_{pt}$, $R_3 = (\delta_{i,j}, \delta_{i,j+1})$. One can show that

$$\begin{aligned}
&e^{\frac{1}{2}\phi_1^2\sigma_{pt}^2(1-\rho^2)}\mathrm{E}[e^{-(\phi_1\rho\sigma_{pt}+\phi_2\sigma_{xt})z}(N(x_1) - N(x_2))|R_3] \\
&= e^{\frac{1}{2}(\phi_1^2\sigma_{pt}^2 + \rho\phi_1\phi_2\sigma_{pt}\sigma_{xt} + \phi_2^2\sigma_{xt}^2)}\mathrm{E}[(N(x_3) - N(x_3))|R_4] \\
&= \mathrm{E}[e^{-(\phi_1 p_t + \phi_2 x_t)}(N(x_3) - N(x_4))|R_4],
\end{aligned}$$

$x_3 = \frac{\delta_{i+1}-\rho z+(\phi_2\rho\sigma_{xt}+\phi_1\sigma_{pt})}{\sqrt{1-\rho^2}}$, $x_4 = \frac{\delta_i-\rho z+(\phi_2\rho\sigma_{xt}+\phi_1\sigma_{pt})}{\sqrt{1-\rho^2}}$, $R_4 = (\delta_j+\phi_1\rho\sigma_{pt}+\phi_2\sigma_{xt}, \delta_{j+1}+\phi_1\rho\sigma_{pt}+\phi_2\sigma_{xt})$. Noting that $e^{\frac{1}{2}(\phi_1^2\sigma_{pt}^2+\rho\phi_1\phi_2\sigma_{pt}\sigma_{xt}+\phi_2^2\sigma_{xt}^2)}\mathrm{E}[= \mathrm{E}[e^{-(\phi_1p_t+\phi_2x_t)}]$, we get

$$\mathrm{E}[e^{-(\phi_1p_t+\phi_2x_t)}|R_1] = \mathrm{E}[e^{-(\phi_1p_t+\phi_2x_t)}]\mathrm{E}[(N(x_3)-N(x_4))|R_4].$$

The proposition is proved by noting that the expected value of $e^{-(\phi_1p_t+\phi_2x_t)}$ conditional R_1 is $\mathrm{E}[e^{-(\phi_1p_t+\phi_2x_t)}|R_1]$ divided by the probability of R_1, which is 0.01.

Proof of Proposition 9

We will first consider the expected return conditional on return. That is, we are interested in the mean of

$$\frac{P_{t+1}+D_{t+1}}{P_t}$$

conditional on the return of previous period

$$\frac{P_t}{P_{t-1}} = e^{r_t}.$$

From the assumption that $\frac{V_{t+1}}{V_t} = e^{\mu+\sigma_r\epsilon_{rt+1}}$, we get

$$\mu+\sigma_r\epsilon_{rt}-(1-\rho)\Delta_{t-1}+\sigma_{\epsilon_\Delta}\epsilon_{et} = r_t.$$

Thus,

$$\sigma_r\epsilon_{rt}-(1-\rho)\Delta_{t-1}+\sigma_{\epsilon_\Delta}\epsilon_{et} = r_t-\mu.$$

Therefore,

$$\begin{aligned}
\mathrm{E}[-(1-\rho)\Delta_{t-1}|r_t] &= \frac{(1-\rho)^2\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}(r_t-\mu)}{\sigma_{\epsilon_\Delta}^2+\sigma_r^2+(1-\rho)^2\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}; \\
\mathrm{E}[\sigma_{\epsilon_\Delta}\epsilon_{et}|r_t] &= \frac{\sigma_{\epsilon_\Delta}^2(r_t-\mu)}{\sigma_{\epsilon_\Delta}^2+\sigma_r^2+(1-\rho)^2\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}},
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{E}[\big(-(1-\rho)\Delta_{t-1}\big)^2|r_t] &= \frac{(\sigma_{\epsilon_\Delta}^2+\sigma_r^2)(1-\rho)^2\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}{\sigma_{\epsilon_\Delta}^2+\sigma_r^2+(1-\rho)^2\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}; \\
\mathrm{E}[(\sigma_{\epsilon_\Delta}\epsilon_{et})^2|r_t] &= \frac{\sigma_{\epsilon_\Delta}^2\left(\sigma_r^2+(1-\rho)^2\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}\right)}{\sigma_{\epsilon_\Delta}^2+\sigma_r^2+(1-\rho)^2\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}, \\
\mathrm{E}[\Big(-(1-\rho)\Delta_{t-1}\Big)\sigma_{\epsilon_\Delta}\epsilon_{et}|r_t] &= \frac{-\sigma_{\epsilon_\Delta}^2(1-\rho)^2\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}{\sigma_{\epsilon_\Delta}^2+\sigma_r^2+(1-\rho)^2\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathrm{E}[\Delta_t|r_t] &= \rho\mathrm{E}[\Delta_{t-1}|r_t]+\mathrm{E}[\sigma_{\epsilon_\Delta}\epsilon_t|r_t] = \frac{-\rho(1-\rho)\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}+\sigma_{\epsilon_\Delta}^2}{\sigma_{\epsilon_\Delta}^2+\sigma_r^2+(1-\rho)^2\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}(r_t-\mu) \\
&= \frac{\frac{\sigma_{\epsilon_\Delta}^2}{1+\rho}}{\sigma_{\epsilon_\Delta}^2+\sigma_r^2+(1-\rho)^2\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}(r_t-\mu).
\end{aligned}$$

The variance conditional on r_t is

$$\begin{aligned}
\mathrm{E}[\big(\rho\Delta_{t-1}+\sigma_{\epsilon_\Delta}\epsilon_t\big)^2|r_t] &= \mathrm{E}[\rho^2\Delta_{t-1}^2+2\rho\Delta_{t-1}\sigma_{\epsilon_\Delta}\epsilon_t+\big(\sigma_{\epsilon_\Delta}\epsilon_t\big)^2|r_t] \\
&= \frac{\rho^2(\sigma_{\epsilon_\Delta}^2+\sigma_r^2)\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}+2\rho(1-\rho)\sigma_{\epsilon_\Delta}^2\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}+\left(\sigma_r^2+(1-\rho)^2\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}\right)\sigma_{\epsilon_\Delta}^2}{\sigma_{\epsilon_\Delta}^2+\sigma_r^2+(1-\rho)^2\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}} \\
&= \frac{\sigma_{\epsilon_\Delta}^2\frac{\sigma_r^2+\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}{\sigma_{\epsilon_\Delta}^2+\sigma_r^2+(1-\rho)^2\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}} = \frac{\sigma_{\epsilon_\Delta}^2(\sigma_r^2+\sigma_{\epsilon_\Delta}^2)}{(1-\rho^2)(\sigma_{\epsilon_\Delta}^2+\sigma_r^2)+(1-\rho)^2\sigma_{\epsilon_\Delta}^2}.
\end{aligned}$$

Thus,

$$
\begin{aligned}
\mathrm{E}\left[\frac{P_{t+1}}{P_t}|r_t\right] &= e^{\mu+\frac{1}{2}(\sigma_r^2+\sigma_{\epsilon_\Delta}^2)}\mathrm{E}[e^{-(1-\rho)\Delta_t}|r_t] \\
&= e^{\mu+\frac{1}{2}(\sigma_r^2+\sigma_{\epsilon_\Delta}^2)}e^{\frac{-\frac{1-\rho}{1+\rho}\sigma_{\epsilon_\Delta}^2(r_t-\mu)}{\sigma_{\epsilon_\Delta}^2+\sigma_r^2+(1-\rho)^2\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}+\frac{1}{2}\frac{(1-\rho)^2\sigma_{\epsilon_\Delta}^2(\sigma_r^2+\sigma_{\epsilon_\Delta}^2)}{(1-\rho^2)(\sigma_{\epsilon_\Delta}^2+\sigma_r^2)+(1-\rho^2)\sigma_{\epsilon_\Delta}^2}} \\
&= e^{\mu+\frac{1}{2}(\sigma_r^2+\sigma_{\epsilon_\Delta}^2)}e^{\frac{\frac{1-\rho}{1+\rho}\left(\mu+\frac{1}{2}(\sigma_r^2+\sigma_{\epsilon_\Delta}^2)\right)\sigma_{\epsilon_\Delta}^2}{\sigma_{\epsilon_\Delta}^2+\sigma_r^2+(1-\rho)^2\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}}e^{\frac{-\frac{1-\rho}{1+\rho}\sigma_{\epsilon_\Delta}^2 r_t}{\sigma_{\epsilon_\Delta}^2+\sigma_r^2+(1-\rho)^2\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}}.
\end{aligned}
$$

Furthermore,

$$
\begin{aligned}
&\mathrm{E}\left[\frac{D_{t+1}}{P_t}|r_t\right] = \mathrm{E}\left[\frac{V_{t+1}}{V_t}\right]\mathrm{E}\left[\frac{D_{t+1}}{V_{t+1}}\right]\mathrm{E}\left[e^{-\Delta_t}|r_t\right] \\
&= e^{\mu+\frac{1}{2}\sigma_r^2}e^{-\bar{x}_v+\frac{(1-\rho_x)\sigma_{\epsilon_x}^2}{2(1+\rho_x)}}e^{\frac{\left(\frac{1}{1+\rho}\mu+\frac{1}{2(1-\rho^2)}(\sigma_r^2+\sigma_{\epsilon_\Delta}^2)\right)\sigma_{\epsilon_\Delta}^2}{\sigma_{\epsilon_\Delta}^2+\sigma_r^2+(1-\rho)^2\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}}e^{\frac{-\frac{1}{1+\rho}\sigma_{\epsilon_\Delta}^2 r_t}{\sigma_{\epsilon_\Delta}^2+\sigma_r^2+(1-\rho)^2\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}}.
\end{aligned}
$$

The proposition is proved by combining the above two equations.

Proof of Proposition 10

Let $\bar{\Delta}_{t,t_0}$ denote the mean of Δ_t conditional on prices from time t_0 to t, $\bar{\Delta}_{t,t_0} = \mathrm{E}[\Delta_t|\{p_s\}_{t_0}^t]$, and $\bar{\sigma}_{t,t_0}^2$ the variance of Δ_t conditional on prices from time t_0 to t, $\bar{\sigma}_{t,t_0}^2 = \mathrm{E}[(\Delta_t - \bar{\Delta}_{t,t_0})^2|\{p_s\}_{t_0}^t]$. If $t_0 = t$, it is the case considered in Proposition 3. Now consider $t_0 = t-1$. Using the fact that Δ_t is stationary, we can write

$$p_t - p_{t-1} = v_t - v_{t-1} + \Delta_t - \Delta_{t-1} = \mu + \sigma_r\epsilon_{rt} - (1-\rho)\Delta_{t-1} + \sigma_{\epsilon_\Delta}\epsilon_{et}.$$

We can re-write the above equation as

$$p_t - p_{t-1} - \mu + (1-\rho)\bar{\Delta}_{t-1} = \sigma_r\epsilon_{rt} - (1-\rho)(\Delta_{t-1} - \bar{\Delta}_{t-1}) + \sigma_{\epsilon_\Delta}\epsilon_{et},$$

where $\bar{\Delta}_{t-1} \equiv \bar{\Delta}_{t-1,t-1}$. Therefore,

$$
\begin{aligned}
\mathrm{E}[-(1-\rho)(\Delta_t - \bar{\Delta}_{t-1})|\{p_s\}_{t-1}^t] &= \frac{(1-\rho)^2\bar{\sigma}_{t-1}^2}{\sigma_r^2+\sigma_{\epsilon_\Delta}^2+(1-\rho)^2\bar{\sigma}_{t-1}^2}(p_t - p_{t-1} - \mu + (1-\rho)\bar{\Delta}_{t-1}); \\
\mathrm{E}[\sigma_{\epsilon_\Delta}\epsilon_{et}|\{p_s\}_{t-1}^t] &= \frac{\sigma_{\epsilon_\Delta}^2}{\sigma_r^2+\sigma_{\epsilon_\Delta}^2+(1-\rho)^2\bar{\sigma}_{t-1}^2}(p_t - p_{t-1} - \mu + (1-\rho)\bar{\Delta}_{t-1}),
\end{aligned}
$$

where $\bar{\sigma}_{t-1}^2 \equiv \bar{\sigma}_{t-1,t-1}^2$, and

$$
\begin{aligned}
\mathrm{E}[\big(-(1-\rho)(\Delta_{t-1} - \bar{\Delta}_{t-1})\big)^2|\{p_s\}_{t-1}^t] &= \frac{(1-\rho)^2\bar{\sigma}_{t-1}^2(\sigma_r^2+\sigma_{\epsilon_\Delta}^2)}{\sigma_r^2+\sigma_{\epsilon_\Delta}^2+(1-\rho)^2\bar{\sigma}_{t-1}^2}; \\
\mathrm{E}[(\sigma_{\epsilon_\Delta}\epsilon_{et})^2|\{p_s\}_{t-1}^t] &= \frac{\sigma_{\epsilon_\Delta}^2(\sigma_r^2+(1-\rho)^2\bar{\sigma}_{t-1}^2)}{\sigma_r^2+\sigma_{\epsilon_\Delta}^2+(1-\rho)^2\bar{\sigma}_{t-1}^2}; \\
\mathrm{E}[-(1-\rho)(\Delta_{t-1} - \bar{\Delta}_{t-1}), \sigma_{\epsilon_\Delta}\epsilon_{rt}|\{p_s\}_{t-1}^t] &= \frac{-(1-\rho)^2\bar{\sigma}_{t-1}^2\sigma_{\epsilon_\Delta}^2}{\sigma_r^2+\sigma_{\epsilon_\Delta}^2+(1-\rho)^2\bar{\sigma}_{t-1}^2}.
\end{aligned}
$$

Furthermore,

$$\bar{\Delta}_{t,t-1} = \bar{\Delta}_{t-1} + \mathrm{E}[\Delta_{t-1} - \bar{\Delta}_{t-1}|\{p_s\}_{t-1}^t].$$

This implies that

$$
\begin{aligned}
\bar{\Delta}_{t,t-1} &= \mathrm{E}[\rho\Delta_{t-1} + \sigma_{\epsilon_\Delta}\epsilon_{et}|\{p_s\}_{t-1}^t] = \rho\bar{\Delta}_{t-1} + \mathrm{E}[\rho(\Delta_{t-1} - \bar{\Delta}_{t-1}) + \sigma_{\epsilon_\Delta}\epsilon_{et}|\{p_s\}_{t-1}^t] \\
&= \rho\bar{\Delta}_{t-1} + \frac{-\rho(1-\rho)\bar{\sigma}_{t-1}^2+\sigma_{\epsilon_\Delta}^2}{\sigma_r^2+\sigma_{\epsilon_\Delta}^2+(1-\rho)^2\bar{\sigma}_{t-1}^2}(p_t - p_{t-1} - \mu + (1-\rho)\bar{\Delta}_{t-1}) \\
&= \frac{\sigma_{\epsilon_\Delta}^2 - \rho(1-\rho)\bar{\sigma}_{t-1}^2}{\sigma_r^2+\sigma_{\epsilon_\Delta}^2+(1-\rho)^2\bar{\sigma}_{t-1}^2}(p_t - p_{t-1} - \mu) + \frac{\rho\sigma_r^2+\sigma_{\epsilon_\Delta}^2}{\sigma_r^2+\sigma_{\epsilon_\Delta}^2+(1-\rho)^2\bar{\sigma}_{t-1}^2}\bar{\Delta}_{t-1} \\
&= \frac{\sigma_{\epsilon_\Delta}^2 - \rho(1-\rho)\bar{\sigma}_{t-1}^2}{\sigma_r^2+\sigma_{\epsilon_\Delta}^2+(1-\rho)^2\bar{\sigma}_{t-1}^2}(p_t - p_{t-1} - \mu) + \frac{\rho\sigma_r^2+\sigma_{\epsilon_\Delta}^2}{\sigma_r^2+\sigma_{\epsilon_\Delta}^2+(1-\rho)^2\bar{\sigma}_{t-1}^2}\bar{\Delta}_{t-1}.
\end{aligned}
$$

The variance of Δ_t conditional on $\{p_s\}_{t-1}^t$ is

$$\bar{\sigma}_{t,t-1}^2 = \frac{\rho^2\bar{\sigma}_{t-1}^2(\sigma_r^2+\sigma_{\epsilon_\Delta}^2)+\sigma_{\epsilon_\Delta}^2(\sigma_r^2+(1-\rho)^2\bar{\sigma}_{t-1}^2)+2\rho(1-\rho)\bar{\sigma}_{t-1}^2\sigma_{\epsilon_\Delta}^2}{\sigma_r^2+\sigma_{\epsilon_\Delta}^2+(1-\rho)^2\bar{\sigma}_{t-1}^2} = \frac{(\rho^2\bar{\sigma}_{t-1}^2+\sigma_{\epsilon_\Delta}^2)\sigma_r^2+\sigma_{\epsilon_\Delta}^2\bar{\sigma}_{t-1}^2}{\sigma_{\epsilon_\Delta}^2+\sigma_r^2+(1-\rho)^2\bar{\sigma}_{t-1}^2}.$$

Iterating this relation, we get

$$\begin{aligned}\bar{\Delta}_{t,t_0} &= \sum_{s=t_0+1}^{t}\left(\prod_{u=s+1}^{t}\frac{\rho\sigma_r^2+\sigma_{\epsilon_\Delta}^2}{\sigma_r^2+\sigma_{\epsilon_\Delta}^2+(1-\rho)^2\bar{\sigma}_{u-1}^2}\right)\frac{\sigma_{\epsilon_\Delta}^2-\rho(1-\rho)\bar{\sigma}_{s-1}^2}{\sigma_r^2+\sigma_{\epsilon_\Delta}^2+(1-\rho)^2\bar{\sigma}_{s-1}^2}(p_s-p_{s-1}-\mu)\\ &\quad+\left(\prod_{u=t_0+1}^{t}\frac{\rho\sigma_r^2+\sigma_{\epsilon_\Delta}^2}{\sigma_r^2+\sigma_{\epsilon_\Delta}^2+(1-\rho)^2\bar{\sigma}_{u-1}^2}\right)\bar{\Delta}_{t_0}.\end{aligned}$$

We use the convention that the summation is zero if the lower index of the summation operator Σ is greater than the upper index and the product is 1 if if the lower index of the product operator Σ is greater than the upper index. If $\bar{\sigma}_0^2=\bar{\sigma}^2$, then $\bar{\sigma}_t^2=\bar{\sigma}^2$ for all t, with

$$\bar{\sigma}^2 = \sqrt{\sigma_{\epsilon_\Delta}^2+\frac{(1+\rho)^2}{4}\sigma_r^2}\frac{\sigma_r}{1-\rho}-\frac{1+\rho}{2(1-\rho)}\sigma_r^2.$$

The above expression simplifies to

$$\begin{aligned}\bar{\Delta}_{t,t_0} &= \sum_{s=t_0+1}^{t}\left(\frac{\rho\sigma_r^2+\sigma_{\epsilon_\Delta}^2}{\sigma_r^2+\sigma_{\epsilon_\Delta}^2+(1-\rho)^2\bar{\sigma}^2}\right)^{t-s}\frac{\sigma_{\epsilon_\Delta}^2-\rho(1-\rho)\bar{\sigma}^2}{\sigma_r^2+\sigma_{\epsilon_\Delta}^2+(1-\rho)^2\bar{\sigma}^2}(p_s-p_{s-1}-\mu)\\ &\quad+\left(\frac{\rho\sigma_r^2+\sigma_{\epsilon_\Delta}^2}{\sigma_r^2+\sigma_{\epsilon_\Delta}^2+(1-\rho)^2\bar{\sigma}^2}\right)^{t-t_0}\bar{\Delta}_{t_0}.\end{aligned}$$

At time t_0, the expected noise $\bar{\Delta}_{t_0}$ conditional on p_{t_0} is the same as in Proposition 3,

$$\bar{\Delta}_{t_0} = \frac{\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}(p_{t_0}-\bar{p}_{t_0})}{\sigma_{vt_0}^2+\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}.$$

By assumption,

$$\frac{1}{\bar{\sigma}^2} = \frac{1}{\sigma_{vt_0}^2}+\frac{1-\rho^2}{\sigma_{\epsilon_\Delta}^2},$$

we can write

$$\bar{\Delta}_{t_0} = \frac{\bar{\sigma}^2}{\sigma_{vt_0}^2}(p_{t_0}-\bar{p}_{t_0}) = \left(1-\frac{1-\rho^2}{\sigma_{\epsilon_\Delta}^2}\bar{\sigma}^2\right)(p_{t_0}-\bar{p}_{t_0}).$$

The dependence on p_0 is given by

$$\begin{aligned}h_{p_0} &= (\sigma_{\epsilon_\Delta}^2-\rho(1-\rho)\bar{\sigma}^2)-(\rho\sigma_r^2+\sigma_{\epsilon_\Delta}^2)\left(1-\frac{1-\rho^2}{\sigma_{\epsilon_\Delta}^2}\bar{\sigma}^2\right)\\ &= -\sigma_{\epsilon_\Delta}^2+\rho(1-\rho)\bar{\sigma}^2+\rho\sigma_r^2+\sigma_{\epsilon_\Delta}^2-\rho\sigma_r^2\frac{1-\rho^2}{\sigma_{\epsilon_\Delta}^2}\bar{\sigma}^2-(1-\rho^2)\bar{\sigma}^2\\ &= -(1-\rho)\bar{\sigma}^2+\rho\sigma_r^2\left(1-\frac{1-\rho^2}{\sigma_{\epsilon_\Delta}^2}\bar{\sigma}^2\right).\end{aligned}$$

We can prove that the above expression is negative using the definition of $\bar{\sigma}^2$. Using the notation

$$\begin{aligned}h_p &= \frac{\sigma_{\epsilon_\Delta}^2-\rho(1-\rho)\bar{\sigma}^2}{\sigma_r^2+\sigma_{\epsilon_\Delta}^2+(1-\rho)^2\bar{\sigma}^2};\\ h_e &= \frac{\rho\sigma_r^2+\sigma_{\epsilon_\Delta}^2}{\sigma_r^2+\sigma_{\epsilon_\Delta}^2+(1-\rho)^2\bar{\sigma}^2},\end{aligned}$$

we can express $\bar{\Delta}_{t,t_0}$ as

$$\begin{aligned}\bar{\Delta}_{t,t_0} &= \sum_{s=t_0+1}^{t}h_e^{t-s}h_p(p_s-p_{s-1}-\mu)+h_e^{t-t_0}\bar{\Delta}_{t_0}\\ &= -\frac{1-h_e^{t-t_0}}{1-h_e}h_p\mu-h_e^{t-t_0}\left(1-\frac{1-\rho^2}{\sigma_{\epsilon_\Delta}^2}\bar{\sigma}^2\right)\bar{p}_{t_0}\\ &\quad+h_pp_t-\sum_{s=t_0+1}^{t-1}h_e^{t-s-1}(1-h_e)h_pp_s-h_e^{t-t_0-1}h_{p_0}p_{t_0}.\end{aligned}$$

Multiple Assets with Factor Structure

We assume that there are N assets where the values are given by

$$v_{it} = \mu_{vi} + \beta_i F_t + \sigma_i \epsilon_{it}^v, \quad i = 1, ..., N.$$

At time t, the price satisfies

$$p_{it} = \mu_{vi} + \beta_i F_t + \sigma_i \epsilon_{it}^v + \beta_{ei} F_t + \sigma_{ei} \epsilon_{it}^e - \ln(\mathrm{E}[e^{\Delta_{it}}]).$$

Therefore, there are systematic risks as well as idiosyncratic risks in both the value and the noise. In vector notation, we can write

$$p_t = \bar{p}_t + \beta F_t + \sigma \epsilon_t^v + \beta_e F_t + \sigma_{\epsilon_\Delta} \epsilon_t^\Delta,$$

where σ and σ_{ϵ_Δ} are $N \times N$ diagonal matrices with diagonal elements being σ_i and σ_{ei} respectively, and $\bar{p}_t = \mu_v + \ln \mathrm{E}[e^{\Delta_t}])$. We can write

$$p_t - \bar{p}_t = \sigma_{\epsilon_\Delta} \epsilon_t^v + (\beta + \beta_e) F_t + \sigma \epsilon_t^v.$$

In terms of the notation of Lemma 1, $\theta = \sigma_{\epsilon_\Delta} \epsilon_t^e$, $\bar{\theta} = 0$, $A_0 = 0$, $A_1 = I$ (where I is the N-dimensional identity matrix), $B = (\sigma, \beta)$. Therefore,

$$\Sigma_\theta = \sigma^2$$

and

$$A_1 \Sigma_\theta A_1' + BB' = \sigma^2 + \sigma_{\epsilon_\Delta}^2 + (\beta + \beta_e)(\beta + \beta_e)'.$$

Let $D = \sigma^2 + \sigma_{\epsilon_\Delta}^2$ and $\beta_0 = \beta + \beta_e$, we get

$$(A_1 \Sigma_\theta A_1' + BB')^{-1} = D^{-1} - D^{-1} \beta_0 (1 + \beta_0' D^{-1} \beta_0)^{-1} \beta_0' D^{-1}.$$

An application of Lemma 1 implies that

$$\begin{aligned}
\bar{\Delta}_t &= \Sigma_\theta A_1' (A_1 \Sigma_{\theta_0} A_1' + BB')^{-1} \xi \\
&= \sigma_{\epsilon_\Delta}^2 \left(D^{-1} - D^{-1} \beta_0 (1 + \beta_0' D^{-1} \beta_0)^{-1} \beta_0' D^{-1} \right) (p_t - \bar{p}_t) \\
&= \sigma_{\epsilon_\Delta}^2 \left(D^{-1} p_t - D^{-1} \beta_0 (1 + \beta_0' D^{-1} \beta_0)^{-1} \beta_0' D^{-1} (p_t - \bar{p}_t) \right) \\
&= \sigma_{\epsilon_\Delta}^2 D^{-1} (p_t - \bar{p}_t) - \sigma_{\epsilon_\Delta}^2 D^{-1} \beta_0 (1 + \beta_0' D^{-1} \beta_0)^{-1} \beta_0' D^{-1} (p_t - \bar{p}_t).
\end{aligned}$$

The first term is corresponds to the case of $\beta = 0$.

When $N \to \infty$, $(1 + \beta_0' D^{-1} \beta_0)^{-1} \to 0$, thus the second term goes to zero, the above formula reduces to the formula for the case[5] of $\beta = 0$,

$$\bar{\Delta}_t = \sigma_{\epsilon_\Delta}^2 D^{-1} (p_t - \bar{p}_t) = \sigma_{\epsilon_\Delta}^2 (\sigma_{\epsilon_\Delta}^2 + \sigma^2)^{-1} (p_t - \bar{p}_t).$$

Intuitively, each stock price is a signal on F_t. When there are infinitely many of stock thus infinitely many of the signals, the factor uncertainty is eliminated and thus can be ignored for the inference about the noise Δ_t and thus the computation of the expected return conditional on prices and price ratios.

The above formula is also important for calibration exercises. It implies that only the idiosyncratic volatility σ should be used for computing the expected returns conditional prices and price ratios.

[5] Note that in Proposition 3, the variance of Δ_t is $\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}$ and variance of the value is σ_{vt}^2.