Can Noise Create the Size and Value Effects?*

Robert Arnott
Research Affiliates, LLC

Jason Hsu
Research Affiliates, LLC and University of California, Irvine

Jun Liu
CKGSB, SWUFE, and UCSD

Harry Markowitz
University of California, San Diego

First Draft: March 2005
Current Draft: March 29, 2008

---

*Jun Liu is on leave from University of California, San Diego and is visiting Cheung-Kong Graduate School of Business (CKGSB) and Southwestern University of Finance and Economics (SWUFE). We are grateful for comments from Andrew Ang, Jonathan Berk, Phil Dybvig, Harrison Hong, Ming Huang, John Hughes, Feifei Li, Jing Liu, Jun Pan, Allan Timmermann, and Rossen Valkanov, and seminar participants at 2007 Chinese International Conference in Finance, 2007 Financial Management Association Meeting, 2008 American Finance Association Meeting, the Central University of Finance and Economics of China, Cheung-Kong Graduate School of Business, the Chinese University of Hong Kong, Georgia Tech, Southwestern University of Finance and Economics, Shandong University, Peking University, Shanghai University of Finance and Economics, TsingHua University, and UCSD. We would like to thank Vitali Kalesnik for helpful technical assistance.

Abstract

Can Noise Create the Size and Value Effects?

Blume and Stambaugh (1983) argue that that the bid-ask spread leads to a random noise in stock price and show that, with noise in price, a stock with higher noise volatility has a higher **unconditional expected return**. Black (1986) and Summers (1986) suggest that market inefficiency also produce noise in price at quarterly and annual frequencies, thus beyond the extent of market microstructure.

In this paper, we show that price noise can produce rich patterns in **conditional expected returns**. In particular, cross-sectional variations in expected returns conditional on price and price-ratios can be generated by variations in realizations of the price noise and thus exist even in the stock have an identical unconditional distribution. We show that stocks with lower prices or lower price ratios have higher expected returns. These higher expected returns are not compensation for risk but are generated because a stock with a low price or a price-ratio is more likely to have a negative price noise which leads to a higher expected return.

One advantage of our approach is parsimony. All stocks are identical ex-ante in our model. The cross-sectional variation in expected return is determined from the parameters that specify the time-series property of one representative stock. Another advantage is a small stock in our approach is a stock that has a low market capitalization, whereas Blume and Stambaugh have to make an extra assumption that a smaller stock has higher noise volatility.

Fama and French (1992) use the matrix of expected returns conditional on size-value deciles as a demonstration of size and value effects. This matrix can be computed in closed form using our model and, for plausible parameters, is similar to its empirical counterpart (Table V of Fama and French). Our study suggests that noise can create the size and value effect.

1 Introduction

Blume and Stambaugh (1983) argue that the bid-ask spread leads to a random noise in stock price and show that, in the presence of noise in price, the **unconditional expected return** of a stock with a higher noise volatility is higher. Thus, a smaller firm has a higher unconditional expected return assuming that it is more likely for smaller to have a higher noise volatility. Black (1986), and Summers (1986) suggest that market inefficiency can produce a random noise in price beyond the scope of market microstructure and propose that autocorrelation in stock returns as an evidence.

In this paper, we show that noise in price can produce rich patterns of **conditional expected returns**. We show that a stock with a low price[1] and/or price-dividend ratio has a lower expected return. With a plausible set of parameters, where the conditional volatility of return due to noise is about 6%, the matrix of expected returns conditional on size and value deciles computed using our model is similar to the matrix of Fama and French (1992).

The novelty of our approach is that stocks have identical unconditional distributions. The cross-sectional variation in expected returns is generated by the variation of realizations of price noise: a stock with a negative (positive) realization of noise have a higher (lower) expected return. This has two advantages. The first one is in the parsimony of the model, we only need to specify the time-series property of one representative stock. For example, we only need to specify 7 parameters to produce the 10×10 matrix of expected return conditional on size/value deciles in Fama-French (1992). The second advantage is that small (value[2]) stocks in our model are stocks with low market caps (low price-dividend ratios), whereas Blume and Stambaugh (1983) have to make additional assumptions that small stocks have high noise volatility.

In our model, the size and value effects have the same source–*noise*. The intuition is the following. A stock with a positive noise should have a lower expected return. Although noise is unobservable, they can be inferred from prices: noise for a stock is more likely positive if its price is high. The same intuition applies for price-book as well as a variety of other price-fundamental ratios. In Blume and Stambaugh (1983) the higher unconditional expected return for a stock with higher volatility because of a stronger Jensen's inequality effect.

We should point out that our model is different from Berk (1995, 1997) in subtle ways. Berk also studies unconditional expected returns, thus their variation has to be generated from difference in unconditional distributions. In our calibration exercise, stocks have identical unconditional distribution and thus have no variation in unconditional expected returns.

Our model predicts that small and value stocks are on average riskier, in the sense

---

[1] Throughout this paper, we assume that firms have only one stock share outstanding. Therefore, we can use ”price”, market capitalization, and market equity (as in Fama and French (1992)) interchangeably.

[2] In this paper, we use “value” to mean the fundamental or intrinsic value of a stock and use it also in “value effects”. Which usage of the term should be clear from context.

that both systematic and idiosyncratic risks are higher. The average beta[3] and the average idiosyncratic volatility with noise are a few percents higher than the averages without noise, given the parameters calibrated to US market data. However, the higher expected returns in small and value stocks cannot be accounted for by slightly higher (systematic) risks. They are driven mostly by pricing noise in the stock market. Our result suggest that value stocks are, indeed, more likely to be undervalued.

We should point out that it is possible that higher expected returns of small-cap and value stocks may not persist over time. On the other hand, they may persist over time due to the limits of arbitrage, associated with either risks of small and value stocks or transactional costs.

In our paper, both the noise and the value process are exogenously given. The value process, which is a Gaussian random walk in our paper, is used in many academic studies and can be generated in an equilibrium model. This specification is useful for closed-form solution for the size and value spread. In general, the value process from asset pricing theories may not have the exact form we assumed, however the intuition still applies. The noise process, which describes deviations from equilibrium, is exogenously specified as a mean reverting process. Our specification of the noise is quite intuitive and plausible and is used extensively in literature (Summers (1986), Poterba and Summers (1988), and Fama and French (1988), to name a few). To endogenize the noise process, a model of off-equilibrium is needed, which is beyond the scope of this paper.

However, it is not easy to detect the presence of these temporary deviations, as pointed out by Summers (1986), Fama and French (1988), and Poterba and Summers (1988). At the same time, the cross section of expected returns predicted by economic theories does not match that observed in the data. In particular, stocks with a low price(market capitalization) and/or price-to-fundamental ratio have higher expected returns, as summarized by the matrix (Fama and French (1996) deciles.

Our paper is organized as follows. In Section 2, we review the related literature briefly. In Section 3, we formally introduce the model of noise and specify the parameters of the model. We explore the implication on unconditional expected stock returns in the presence of pricing noise. We show that stocks with greater noise earn higher returns, on average. In Section 4, we compute the expected returns conditional on price and price ratios. In section 5, we compute the matrix of expected return conditional on size and value simultaneously. We then compute the matrix of expected returns, beta, and alpha conditional on size and value deciles. We make concluding remarks in Section 6.

---

[3] Lakonishok, Shleifer, and Vishny (1994) found that the beta of the value stocks is about 0.1 higher than the beta of the growth stocks.

2 Literature Review

Our paper is closely related to two brands of literature, namely, that on noise in price and on value and size effects.

Blume and Stambaugh (1983) argue that observed price is either the bid or the ask, not the fundamental value, thus price is different from the fundamental value by a random noise term.[4] They show that the unconditional expected returns increases with the volatility of the noise. A smaller stock has a higher unconditional expected return if a smaller stock has a higher noise volatility. In our paper, a smaller stock is a stock with a smaller price. Furthermore, the variation in expected returns in Blume and Stambaugh has to be generated from variation in parameters such as difference in noise volatility, whereas the variation is generated by difference in realization of noise. Finally, the bid-ask bounce is useful for explaining effects in daily returns but is less likely the cause for effects that occur at quarterly or annual horizons and the size effect is observed in these horizons.

Black (1986) proposes that financial markets are noisy (that prices are different from the fundamental values) due to trading by investors without information. He believes that "noise causes the market to be somewhat inefficient but yet prevent people from taking advantage of inefficiencies."

Summers (1986) also argues that prices are noisy and build a formal model to show that the power of the standard econometric tests are simply too weak to either detect noise or reject the Efficient Market Hypothesis.

Fama and French (1988) and Poterba and Summers (1988) study mean-reversion in prices and point out that one of the possible explanation for mean reversion is the deviation of price from the efficient market value. They infer the existence and properties of noise from the autocorrelation of returns.

In behavioral finance literature, pricing error can arise from investor overreaction or underreaction, as suggested by Shiller (1981), DeBondt and Thaler (1985, 1987), Lakonishok, Vishny, and Shleifer (1994), among others. Price can be different from value if investors under- or over-react. With random realization of positive or negative news, over- or under-reaction presumably should generate noise–random deviation from value. Note that over- or under-reaction is different from optimism or pessimism, which presumably lead to biased deviations from the value.

In Campbell and Kyle (1993) value is determined endogenously, but the price is different from value by a mean-reverting noise that is exogenously specified. They show that this model can explain the volatility and predictability of the US stock returns.

In term structure models, where the number of shocks is usually smaller than the number of independent securities, it is assumed that the market prices for bonds are different from

---

[4] There are subsequently many studies in market microstructure literature on noise in prices. See for example, Daniel, Hirshleifer, and Subrahmanyam (2001) and Chordia, Roll, and Subrahmanyam (2005). However, noise considered in this paper is less likely due to market microstructure.

the model-derived fair values by a noise.

The size and value effects have spurred spirited debates since Banz (1981) and Reinganum (1981) documented that smaller capitalization stocks tend to outperform on a risk-adjusted basis, and Stattman (1980) and Rosenberg, Reid and Lanstein (1985) documented that high book-market stocks also outperform. Similarly, other ratios such as earnings-price, documented by Basu (1977) and dividend yield, documented by Razeff (1984), Shiller (1984), Blume (1980) and Keim (1985), also predict future performance.

There are many explanations for the observed size and value effects. Fama and French (1992) show that size and value, along with market beta, capture well the cross-sectional variation in stock returns and subsume the explanatory powers of other financial variables. They propose that the size and value premia are compensation for risk. Lakonishok, Shleifer and Vishny (1994) argue that the size and value premia are due to investor overreaction rather than to risk. Gomes, Kogan, and Zhang (2003) Zhang (2006) argues that the value effect can be explained in a production economy. Yogo (2006) proposes that the size and value effects can be explained by investor preferences that are non-separable in nondurable and durable consumption.

Berk (1995, 1997) suggests noise as a source of size and value effects. Although Berk shows (under mild conditions) that any source of cross-sectional variations in expected returns will produce a size effect, he focuses on risk. We focus on another cross-sectional variation that he does not focus on–noise in asset prices. Berk considers unconditional expected returns, thus the difference in expected returns is generated by difference in parameters, whereas we study conditional returns where the difference in expected returns is generated by difference in realization of noise. Berk's model applied with our assumption of identical unconditional distribution would produce no cross-sectional variation in expected returns.

Arnott, Hsu, and Moore (2005) and Arnott (2005a, b) also propose that 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. Arnott and Hsu (2008) show that mean-reverting mispricing can lead to small cap and value stock outperformance; they predict that size and value might be two manifestations of one effect, pricing noise.

Brennan and Wang (2006) also study, empirically as well as theoretically, the effect of mispricing on unconditional expected returns for a larger class of models, where mispricings can be due to slowness in adjustment of price and systematic mispricing in addition to random noise. They did not study conditional expected returns which are our focus.

Finally, our paper is also related to noisy rational expectation equilibrium (NREE) models. In these models, the price is noisy to uninformed investors in a sense that these investors do not observe the signal about the stock but infer it with noise from the price. In our paper, the price is a noisy signal about the fundament value of the stock in the same sense. However, the price is derived endogenously and is equal to the value in NREE models. As such, expected return in excess over the riskfree return is the compensation for risk and there is no alpha. In our paper, the price is assumed exogenously and equal to the value only up to a noise term. As such, part of expected return in excess over the riskfree return is not the compensation for risk and alpha is not zero.

3 Noise

In this section, we discuss key assumptions and technical assumptions 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 fundamental 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*.*

This is the key assumption of the paper. Blume and Stambaugh (1983) first made the assumption that the market price is different from the fundamental value based on the random bounce between the bid and ask prices. Black (1986) argues that there should be noise in price because of noise trading. Summers (1986) 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." Formally, Summers (1986) proposes a similar model of noise in price as an alternative to the efficient market hypothesis and argues that it is not easy to reject this alternative using traditional standard tests.

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. We will assume that the value V_t has all the "nice" properties of standard asset pricing models, 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 only by beta if the asset pricing model is APT. For our purpose, it is not necessary to specify 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.

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 $\mathrm{E}[\Delta_t] = 0$.

The value V_t is the price if there were no noise. However, we will assume that there is noise Δ_t and V_t is not observed, only P_t is observed. Assumption 1 implies that

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

That is, for any given value of V_t, the price P_t for a stock is a noisy proxy for the value V_t, which we assume is unobservable, and the price is, on average, right. In noisy rational expectation (NREE) models, given an information s, the equilibrium price P is a noisy signal of s if P satisfies

$$\mathrm{E}[P|s] = s.$$

In exactly the same sense, in our model, P_t is a noisy signal of V_t. Just as in NREE models where information s can be inferred with noise from P, the fundamental value V_t can be inferred from P_t with noise in our model. As we will show later, if P_t is high, $\mathrm{E}[V_t|P_t]$ will be high and vice versa.

Furthermore, the above assumption implies that P_t and Δ_t are positively correlated. Thus,

$$\mathrm{E}[V_t|P_t] = P_t \mathrm{E}[e^{-\Delta_t}|P_t]\mathrm{E}[e^{\Delta_t}].$$

In general,

$$\mathrm{E}[e^{-\Delta_t}|P_t]\mathrm{E}[e^{\Delta_t}] \neq 1.$$

Thus, the expectation of the fundamental value V_t given P_t does not equal to P_t.

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 gradually 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 force will move price P_t towards 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 $\rho = 1$.

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 (2008), 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).[5]

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.

---

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

Assumptions 2 and 3 are made 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 now study the implications of noise on unconditional expected returns. We show that noise can generate cross-sectional variations in unconditional expected stock returns.

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

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*

$$\mathrm{E}\left[\frac{P_{t+1}+D_{t+1}}{P_t}\right] = \mathrm{E}\left[\frac{V_{t+1}}{V_t}\right] e^{\frac{\sigma^2_{\epsilon_\Delta}}{1+\rho}} + \mathrm{E}\left[\frac{D_{t+1}}{V_t}\right] e^{\frac{\sigma^2_{\epsilon_\Delta}}{1-\rho^2}}, \tag{8}$$

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*

$$\mathrm{E}\left[\frac{P_{t+1}+D_{t+1}}{P_t}\right] = e^{\mu+\frac{1}{2}\sigma_r^2}\left(e^{\frac{\sigma^2_{\epsilon_\Delta}}{1+\rho}} + e^{-\bar{x}_v+\frac{\sigma^2_{\epsilon_x}}{2(1-\rho_x^2)}+\frac{\sigma^2_{\epsilon_\Delta}}{1-\rho^2}}\right). \tag{9}$$

We note that 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.

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 noise. 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. In addition to the implications we discuss below, the unconditional expected return given by equation (9) can be to compare the conditional expected return we study in later sections.

Blume and Stambaugh (1983) suggest that bid-ask spreads produce noise in price. They compute unconditional expected return approximately for the case that is similar to $\rho = 0$ and $D = 0$ case of the above Proposition (we can the closed form unconditional expected return because of the exponential form of the noise in our model). They show that the noise increases the unconditional expected return which is increasing with noise variance and use this fact to explain the size effects in daily data.

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 (9) 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 different 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%.

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 much 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).

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 selection of these parameters are described in Section 3.

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,

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

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 $\sigma^2_{\epsilon_\Delta}$ 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 in stock returns 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 $\sigma^2_{\epsilon_\Delta}$.

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 did. 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,[6] 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 $\sigma_{\epsilon_x} = 10\%$.

---

[6] 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.

4 The Time-Series Variation of Expected Returns

In this section, we show that the expected return conditional on the price or the price-dividend ratio increases with the price or price-dividend ratio. The expected return of a stock conditional on the market price or the price-dividend ratio or both is computed in closed form.

Whereas the result on unconditional expected return is due to Jensen's inequality, the intuition for conditional expected returns is different but straightforward. First, suppose that the noise is observed. If Δ_t is negative, then the market price of the stock is lower than its fundamental values and the expected return conditional on a negative Δ_t should be high.

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. 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.

4.1 The "Time-Series Size Effect"

We now 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,

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

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.

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_{vt}^2. Then the expected return conditional on P_t is*

$$\begin{aligned} & \mathrm{E}\left[\frac{P_{t+1}+D_{t+1}}{P_t}|p_t\right] \\ = & \; e^{\mu+\frac{1}{2}\sigma_r^2}\left(e^{\frac{\sigma_{\epsilon_\Delta}^2}{1+\rho}}\frac{P_t^{-(1-\rho)\gamma_1}}{\bar{P}_{t,-(1-\rho)\gamma_1}}+e^{-\bar{x}_v+\frac{\sigma_{\epsilon x}^2}{2(1-\rho_x^2)}+\frac{\sigma_{\epsilon_\Delta}^2}{1-\rho^2}}\frac{P_t^{-\gamma_1}}{\bar{P}_{t,-\gamma_1}}\right), \end{aligned} \tag{11}$$

where $\gamma_1 = \frac{\sigma_{\epsilon_\Delta}^2}{(1-\rho^2)\sigma_{vt}^2+\sigma_{\epsilon_\Delta}^2}$ and $\bar{P}_{t,\gamma}$ denotes the expected value of P_t^γ given v_t is normal with mean $\bar{v}_t$ and variance σ_{vt}^2.

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,

$$\frac{P_{t+1}}{P_t} = \frac{V_{t+1}}{V_t}e^{\Delta_{t+1}-\Delta_t}. \tag{12}$$

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,

$$p_t = v_t + \Delta_t - \ln(\mathrm{E}[e^{\Delta_t}]). \tag{13}$$

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.

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}.

4.2 The "Time-Series 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. We now 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.

Proposition 4 *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 intuition for the x_t dependence is the same as the intuition for the p_t dependence explored in in Section 4.1. A high price-dividend ratio implies a high noise Δ_t, on average, thus a low expected return.

Proposition 4 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.

4.3 The "Time-Series Size-Value Effect"

So far, we have studied the expected return conditional on either the price or the price-dividend ratio separately. 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 5 *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_{vt}^2*. Then the expected return conditional on* p_t *and* x_t *is,*

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

where $\gamma_3 = \frac{\frac{1}{\sigma_{vt}^2}}{\frac{1}{\sigma_{vt}^2}+\frac{1-\rho_x^2}{\sigma_{\epsilon_x}^2}+\frac{1-\rho^2}{\sigma_{\epsilon_\Delta}^2}}$ *and* $\gamma_4 = \frac{\frac{(1-\rho_x^2)}{\sigma_{\epsilon_x}^2}}{\frac{1}{\sigma_{vt}^2}+\frac{1-\rho_x^2}{\sigma_{\epsilon_x}^2}+\frac{1-\rho^2}{\sigma_{\epsilon_\Delta}^2}}$. $\overline{PX}_{t,\gamma_1,\gamma_2}$ *is the expected value of* $P_t^{\gamma_1}X_t^{\gamma_2}$ *given* v_t *is a normal with mean* $\bar{v}_t$ *and variance* σ_{vt}^2*.*

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.

5 The Cross-Sectional Variation of Expected Returns

So far, we have studied the expected return of one stock conditional on its market cap or price-dividend ratio. In this section, we show that noise can also produce significant cross-sectional variations in expected returns. Fama and French (1992) demonstrate the size and value effects using the matrix of expected return conditional on size and value deciles. These cross-sectional conditional expected returns can be computed using our model.

We first compute the expected return conditional on deciles. We then make assumption to translated into cross-sectional implications.

Assumption 4 *There are* N *stocks. Stock* k *has a value* V_{kt}*, which is determined by economic theory. The price* P_{kt} *of a stock deviates from its fundamental value* V_{kt} *by a noise* Δ_{kt}*. Specifically,*

$$P_{kt} = V_{kt}\frac{e^{\Delta_{kt}}}{\mathrm{E}[e^{\Delta_{kt}}]},$$

where Δ_{kt} *is independent of* V_{ks} *for all* t *and* s *and* $\mathrm{E}[e^{\Delta_{kt}}]$ *is the unconditional expectation of* $e^{\Delta_{kt}}$*. The dividend* D_{kt} *of the stock is also independent of* Δ_{ks}*, for all* t *and* s*. Let* $v_{kt} = \ln V_{kt}$ *and* $p_{kt} = \ln P_{kt}$*. We also assume*

$$
\begin{aligned}
\Delta_{kt+1} &= \rho\Delta_{kt} + \sigma_{\epsilon_\Delta}\epsilon_{kt+1}^{\Delta}, \\
v_{kt+1} &= \mu_v + v_{kt} + \beta F_{t+1} + \sigma\epsilon_{kt+1}^{v}, \\
v_{kt+1} - d_{kt+1} &= (1-\rho_x)\bar{x}_v + \rho_x(v_{kt} - d_{kt}) + \sigma_{\epsilon_x}\epsilon_{xt+1}^{v},
\end{aligned}
$$

where ρ*,* σ_{ϵ_Δ}*,* μ_v*,* β*,* σ*,* ρ_x*,* $\bar{x}_v$*, and* σ_{ϵ_x} *are all constant. All shocks are independent across stocks and time.*

Assumption 4 states that the Assumptions 1–3 hold for each and every stock. All stocks have the identical distributions but are not independent. The correlation among stocks is introduced through a systematic factor F_t and the beta coefficient is the same for all stocks. All the stock have the same idiosyncratic volatility although the shocks are independent.

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 $\sqrt{\sigma_{vt}^2 + \frac{\sigma_e^2}{1-\rho^2}}$ 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 $\mathrm{E}\Big[f(z)|_{\underline{z}}^{\overline{z}}\Big]$ denote the expectation of $f(z)$ for z between $\underline{z}$ and $\overline{z}$ for a standard normal random variable z.

Proposition 6 (Size-Value Effect) *Suppose that Assumption 4 holds. Furthermore, the number N of stock tends to infinity. Then the expected return conditional on (i, j) decile of (p_t, x_t) space is,*

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

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

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 systematic factor F_t is introduced to describe the correlation among the stock. However, as far as inference of Δ_t is concerned, we can ignore this factor because it can be inferred completely from the prices when the number of stocks is infinite.

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 3–6 hold when the correlations are through factors, provided we replace the variance parameters by their idiosyncratic components.

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 6 with the parameters given by Table 1. The beta in the absence of noise is assumed to be 1.

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 is the same as population average, and so can be computed using Propositions 3-6. So, these propositions 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, will lead to additional cross-sectional variations in expected returns. Next we will show that, with plausible parameters, these variations are consistent with those observed in the US data.

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-cap and value stocks' higher expected returns are attributable to higher systematic risks. In Table 3, we present the beta matrix for size-

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.

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."

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 the higher betas.

As a model for the cross section of expected returns, our paper is different from Berk (1995, 1997), where the heterogeneity in expected return is specified in terms of the heterogeneity of parameters. In fact, in our calibration exercise, stock returns are identically distributed, thus there is no heterogeneity in parameters. Under Berk (1995, 1997), there is no cross-sectional variation in expected returns. The heterogeneity in expected return in our calibration exercise is completely driven by the random realization of the noise in our paper.

There are empirical studies documenting that the size and value spreads are different

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.

between booms and recessions. In our paper, the expected returns conditional on deciles in are state independent. 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 that financial market prices deviate widely and frequently from rational valuations." Our study suggest that the size and value effects are actually evidence for the view that the market prices of stocks deviate from their intrinsic values.

6 Conclusion

This paper builds on the work of Blume and Stambaugh (1983) and Black (1986) and Summers (1986). We show that random realizations of noise generate cross-sectional variations in expected return conditional on price and price ratios. In particular, with plausible parameters, such as a noise volatility of 6% per annum, the matrix of expected returns conditional on size and value deciles, is similar to that of Fama and French (1992). Since the difference in beta for different size and value deciles is small in our model, small and value stocks have higher expected return because they are under-valued due to price noise, not because of higher systematic risk. Thus our results suggest that a modest amount of noise can create size and value effects.

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.

Black (1986) argues that noise should always be present because investors are risk averse and are not sure whether information is just pure noise. According to Black, "noise creates the opportunities to trade profitably, but at the same time makes it difficult to trade profitably." If Black is right, size and value effects are likely to continue to persist.

Our model assumes that noise is independent of the value and the dividend. One could examine the implications of relaxing that assumption. Indeed, for certain forms of dependence, we would expect that the value and size effects should disappear. Empirical evidence clearly does not support this form of the model. But, if enough investors trade on size and value, arbitrage could force this outcome.

Our model assumes static parameters describing the noise function. By empirically observing the time-varying nature of these parameters, we may well find that the growth-value cycle is nothing more than a manifestation of expansion and contraction of the noise variance. We may find a linkage between economic expansion and contraction, or bull and bear markets, and the parameters of our model, notably noise variance. This may allow us to better understand the link between the economic cycle and the growth-value cycle.

If noise varies cross-sectionally, as it presumably will, one can model and empirically test the impact of a world in which some stocks may have more noise than others, and some stocks may mean-revert more quickly than others. Because the difference between average return and average value return is proportional to the variance of the noise and to the rate of mean reversion, this would suggest starkly different behavior and mean returns for assets with little uncertainty about value (e.g., short-term bonds), relative to assets with large uncertainty (e.g., venture capital and private equity).

In short, this simple change in the classic Efficient Market Hypothesis, acknowledging the possibility of mean-reverting noise, perhaps too small to statistically discern, not only better conforms with past empirical findings, but also opens wide opportunities for further research.

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 1

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

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

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

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

The unconditional expected return is,

$$\mathrm{E}\left[\frac{P_{t+1} + D_{t+1}}{P_t}\right] = \mathrm{E}\left[\frac{V_{t+1}}{V_t}\right]\mathrm{E}\left[e^{\Delta_{t+1}-\Delta_t}\right] + \frac{D_{t+1}}{V_t}\mathrm{E}\left[e^{\Delta_t}\right]e^{-\Delta_t}. \tag{16}$$

By stationarity,

$$\mathrm{E}\left[\Delta_{t+1}\right] = \mathrm{E}\left[\Delta_t\right], \tag{17}$$

therefore,

$$\mathrm{E}\left[\Delta_{t+1} - \Delta_t\right] = 0. \tag{18}$$

By Jensen's inequality,

$$\mathrm{E}\left[e^{\Delta_{t+1}-\Delta_t}\right] \geq e^{\mathrm{E}[\Delta_{t+1}-\Delta_t]} = 1. \tag{19}$$

Equation (16) then gives,

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

Furthermore,

$$\mathrm{E}\left[\frac{D_{t+1}}{P_t}\right] = \mathrm{E}\left[\frac{D_{t+1}\mathrm{E}[e^{\Delta_t}]}{V_t e^{\Delta_t}}\right] = \mathrm{E}\left[\frac{D_{t+1}}{V_t}\right]\mathrm{E}[e^{\Delta_t}]\mathrm{E}\left[\frac{1}{e^{\Delta_t}}\right] \geq \mathrm{E}\left[\frac{D_{t+1}}{V_t}\right]. \tag{21}$$

Combining inequalities in (20) and (21), we conclude that

$$\mathrm{E}\left[\frac{P_{t+1}+D_{t+1}}{P_t}\right] \geq \mathrm{E}\left[\frac{V_{t+1}+D_{t+1}}{V_t}\right]. \tag{22}$$

Proof of Proposition 2

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 (9) is proved by noting that

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

Proof of Proposition 3

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^2_{vt}, the distribution of Δ_t is its unconditional distribution of mean 0 and variance $\frac{\sigma^2_{\epsilon_\Delta}}{1-\rho^2}$. 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

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

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 the expectations in the proposition and prove the equivalence between the above equation implies the equation given in the proposition.

Proof of Proposition 4

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 $\frac{\sigma^2_{\epsilon_x}}{1-\rho_x^2}$ and is independent of Δ_t and Δ_t is normal with a mean of 0 and a variance of $\frac{\sigma^2_{\epsilon_\Delta}}{1-\rho_e^2}$.

Therefore, conditional on x_t, the mean of Δ_t is

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

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

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

Thus, we get

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

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

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

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

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

we get

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

Finally,

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

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 5

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

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

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

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 6

In vector notation, we can write

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

We can write

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

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), $B = (\sigma I, \beta I_1)$, where I_1 is a $N \times 1$ vector of 1's. Therefore,

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

and

$$A_1 \Sigma_\theta A_1' + BB' = (\sigma^2 + \sigma_{\epsilon_\Delta}^2) I + \beta^2 I_1 I_1'.$$

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

$$(A_1 \Sigma_\theta A_1' + BB')^{-1} = (\sigma^2 + \sigma_{\epsilon_\Delta}^2)^{-1} I - (\sigma^2 + \sigma_{\epsilon_\Delta}^2)^{-2} \beta^2 (1 + N\beta^2 (\sigma^2 + \sigma_{\epsilon_\Delta}^2)^{-1}) I_1 I_1'.$$

An application of Lemma 1 implies that

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

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

When $N \to \infty$, $(1 + N\beta^2(\sigma^2 + \sigma_{\epsilon_\Delta}^2)^{-1})^{-1} \to 0$, thus the second term goes to zero, the above formula reduces to the formula for the case[7] of $\beta = 0$,

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

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. We only need to consider the case of single stock, as long as only the idiosyncratic volatility σ are used.

Now 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 $R_1 = \{\sigma_{pt}\delta_i \le p_t \le \sigma_{pt}\delta_{i+1}, \sigma_{xt}\delta_{i,j} \le x_t \le \sigma_{xt}\delta_{i,j+1}\}$, 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,

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

---

[7] 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.

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 $e^{\frac{1}{2}(\phi_1^2\sigma_{pt}^2+\rho\phi_1\phi_2\sigma_{pt}\sigma_{xt}+\phi_2^2\sigma_{xt}^2)}\mathrm{E}[= \mathrm{E}[e^{-(\phi_1 p_t+\phi_2 x_t)}]$, we get

$$\mathrm{E}[e^{-(\phi_1 p_t+\phi_2 x_t)}|R_1] = \mathrm{E}[e^{-(\phi_1 p_t+\phi_2 x_t)}]\mathrm{E}[(N(x_3)-N(x_4))|R_4].$$

The proposition is proved by noting that the expected value of $e^{-(\phi_1 p_t+\phi_2 x_t)}$ conditional R_1 is $\mathrm{E}[e^{-(\phi_1 p_t+\phi_2 x_t)}|R_1]$ divided by the probability of R_1, which is 0.01.

References

[1] Aboody, David, John Hughes, and Jing Liu, 2002, "Measuring Value Relevance in a (possibly) Inefficient Market," *Journal of Accounting Research*, v40, n4, 965-986.

[2] Ang, Andrew and Jun Liu, 2006, "Risk, Return, and Dividends," forthcoming, *Journal of Financial Economics.*

[3] Arnott, Robert D., 2005a, "What Cost Noise?" *Financial Analysts Journal*, v61, n2, 10-14.

[4] Arnott, Robert D., 2005b, "Disentangling Size and Value," *Financial Analysts Journal*, v61, n5, 12-15.

[5] Arnott, Robert and Jason Hsu, 2008, "Noise, CAPM and the Size and Value Effects", working paper.

[6] Arnott, Robert D., Jason C. Hsu and Philip Moore, 2005, "Fundamental Indexation," *Financial Analysts Journal*, v61, n2, 83-99.

[7] Banz, Rolf W., 1981, "The Relationship Between Market Value and Return of Common Stocks," *Journal of Financial Economics,* v9, n1, 3-18.

[8] Basu, S., 1977, " Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis", *The Journal of Finance*, v32, n3, 663-682.

[9] Berk, Jonathan B., 1995, "A Critique of Size Related Anomalies," *Review of Financial Studies*, n8, 275-286.

[10] Berk, Jonathan B., 1997, "Does Size Really Matter?" *Financial Analysts Journal*, v52, n5, 12-18.

[11] Black, Fischer, 1986, "Noise,", *The Journal of Finance*, v41, n3, 529-543.

[12] Black, Fischer and Robert Litterman, 1992, "Global Portfolio Optimization", *Financial Analysts Journal*, September/October, v48, n5, 28-43.

[13] Blume, Marshall E., 1980, "Stock Returns and Dividend Yields: Some More Evidence", *The Review of Economics and Statistics*, v62, n4, 567-577.

[14] Blume, Marshall E., and Robert Stambaugh, 1983, "Biases in Computed Returns", *Journal of Financial Economics*, v12, n3, 387-404.

[15] Brennan, Michael and Ashley Wang, 2006, "Asset Pricing and Mispricing", working paper.

[16] Campbell, John Y. and Albert S. Kyle, 1993, "Smart Money, Noise Trading and Stock Price Behaviour", *The Review of Economic Studies*, v60, n1, 1-34.

[17] Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam, 2005, "Evidence on the Speed of Convergence to Market Efficiency", *Journal of Financial Economics*, v76, n2, 271-292.

[18] Daniel, Kent D., David Hirshleifer, and Avanidhar Subrahmanyam, 2001, "Overconfidence, Arbitrage, and Equilibrium Asset Pricing", *The Journal of Finance*, v56, n3, 921-965.

[19] De Bondt, Werner F. M. and Richard Thaler, 1985, "Does the Stock Market Overreact?" *The Journal of Finance*, v40, n3, 793-805.

[20] De Bondt, Werner F. M. and Richard Thaler, 1987, "Further Evidence on Investor Overreaction and Stock Market Seasonality," *The Journal of Finance*, v42, n3, 557-581.

[21] Fama, Eugene F. and Kenneth R. French, 1988, "Permanent and Temporary Components of Stock Prices", *Journal of Political Economy*, v96, n2, 246-73.

[22] Fama, Eugene F. and Kenneth R. French, 1992, "The Cross-Section of Expected Stock Returns", *Journal of Finance*, v47, n2, 427-465.

[23] French, Kenneth and Richard Roll, 1986, "Stock Return Variances: The Arrival of Information and the Reaction of Traders", *Journal of Financial Economics*, v17, n1, 5-26.

[24] Gomes, Joao, Leonid Kogan, and Lu Zhang, 2003, "Equilibrium Cross Section of Returns", *Journal of Political Economy*, v111, n4, 693-732.

[25] Grossman, Sanford J. and Joseph E. Stiglitz, 1980, "On the Impossibility of Informationally Efficient Markets", *American Economic Review*, v70, n3, 393-408.

[26] Hsu, Jason C., 2006, "Cap-Weighted Portfolios Are Sub-optimal Portfolios," *Journal of Investment Management*, 3rd Quarter.

[27] Hughes, John, Jing Liu, and Jun Liu, 2006, "Information Asymmetry, Diversification, and Cost of Capital", UCLA and UCSD working paper.

[28] Jegadeesh, Narasimhan, 1990, "Evidence of Predictable Behavior of Security Returns," *The Journal of Finance*, v45, n3, 881-898.

[29] Jegadeesh, Narasimhan and Sheridan Titman, 1993, "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency," *The Journal of Finance*, v48, n1, 65-91.

[30] Kahneman, D. and A. Tversky, “Intuitive Prediction: Biases and Corrective Procedures,” in D. Kahneman, P. Slovic an A. Tversky (eds.), Judgment Under Uncertainty: Heuristics and Biases, London: Cambridge University Press, 1982.

[31] Keim, D., 1985 “Dividend Yields and Stock Returns: Implications of Abnormal January Returns”, *Journal of Financial Economics*, v14, 473-489.

[32] Lakonishok, Josef, Andrei Shleifer and Robert Vishny, 1994, “Contrarian Investment, Extrapolation, and Risk ,” *The Journal of Finance*, v49, n5, 1541-1578.

[33] Liptser, R. S. and Shiryayev, A. N., 1977, *Staticsitics of Random Processes II*, Berlin: Springer-Verlag.

[34] Poterba, James M. and Lawrence H. Summers, 1986, “The Persistence of Volatility and Stock Market Fluctuations,” *The American Economic Review*, v76, n5, 1142-1151.

[35] Poterba, James M. and Lawrence H. Summers, 1988, “Mean reversion in stock prices: Evidence and implications,” *Journal of Financial Economics,* v22, 27-59.

[36] Reinganum, M. R., 1981, “Misspecification of Capital Asset Pricing: Empirical Anomalies Based on Earnings’ Yields and Market Values,” *Journal of Financial Economics*, v9, n1, 19-46.

[37] Rosenberg, B., K. Reid, and R. Lanstein, 1985, “Persuasive Evidence of Market Inefficiency,” *Journal of Portfolio Management*, v11, 9-17.

[38] Roll, Richard, 1983, “On Computing Mean Return and Small Firm Premium”, *Journal of Financial Economics*, v12, 371-386.

[39] Rozeff, M., 1984, ”Dividend Yields and Equity Risk Premiums,” *Journal of Portfolio Management*, v11, 68-75.

[40] Shiller, Robert, 1981, “Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?” *American Economic Review,* v71, 421-36.

[41] Shiller, Robert, 1984, “Stock Prices and Social Dynamics,” *Brookings Papers on Economics Activity*, v2, 458-510.

[42] Shiller, Robert, 1989, “Comovements in Stock Prices and Comovements in Dividends”, *The Journal of Finance*, v44, n3, 719-729.

[43] Stattman, D., 1980, “Book values and stock returns,” *The Chicago MBA: A Journal of Selected Papers*, v4, 25-45.

[44] Summers, Lawrence, 1986, “Does the Stock Market Rationally Reflect Fundamental Values?” *The Journal of Finance*, v41, n3, 591-601.

[45] Yogo, Motohiro, 2006, “A Consumption-Based Explanation of Expected Stock Returns,” *Journal of Finance*, v61, n2, 539-580.

[46] Zhang, Lu, 2005, “The value premium,” *Journal of Finance*, v60 n1, 67-103.