NBER WORKING PAPER SERIES

STOCK VALUATION AND LEARNING ABOUT PROFITABILITY

Lubos Pastor
Pietro Veronesi

Working Paper 8991
http://www.nber.org/papers/w8991

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2002

Helpful comments were gratefully received from Dan Bens, Tony Bernardo, Michael Brennan, John Campbell, George Constantinides, Gene Fama, Ken French, Chris Geczy, Rick Green, John Heaton, Boyan Jovanovic, Rob Stambaugh, Per Strömberg, Tuomo Vuolteenaho, Larry Wall, Franco Wong, Yihong Xia, an anonymous referee, and seminar participants at the Federal Reserve Bank of Atlanta, the Federal Reserve Bank of Chicago, the University of Chicago, and the University of Pennsylvania. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

Stock Valuation and Learning about Profitability
Lubos Pastor and Pietro Veronesi
NBER Working Paper No. 8991
June 2002
JEL No. G12

ABSTRACT

We develop a simple approach to valuing stocks in the presence of learning about average profitability. The market-to-book ratio (M/B) increases with uncertainty about average profitability, especially for firms that pay no dividends. M/B is predicted to decline over a firm's lifetime due to learning, with steeper decline when the firm is young. These predictions are confirmed empirically. Data also support the predictions that younger stocks and stocks that pay no dividends have more volatile returns. Firm profitability has become more volatile recently, helping explain the puzzling increase in average idiosyncratic return volatility observed over the past few decades.

Lubos Pastor
Graduate School of Business
University of Chicago
1101 E. 58th St.
Chicago, IL 60637
and NBER
lubos.pastor@gsb.uchicago.edu

Pietro Veronesi
Graduate School of Business
University of Chicago
1101 E. 58th St.
Chicago, IL 60637
and NBER
pietro.veronesi@gsb.uchicago.edu

1. Introduction

The past two decades have witnessed an unprecedented surge in the number of newly listed firms on the major U.S. stock exchanges. According to Fama and French (2001b), over 550 new firms per year appeared between years 1980 and 2000, on average, compared to less than 150 firms in the previous two decades. Some of these new firms command valuations that might seem too high to be justified by reasonable assumptions about expected future profitability. For example, more than one in ten of all firms listed between years 1962 and 2000 are worth more than seven times their book value at the end of their year of listing, and almost one in 50 firms are worth more than 20 times their book value. Naturally, investors attempting to value the newly listed firms are confronted with substantial uncertainty about their future profitability. We argue that this uncertainty contributes to the high valuations of young firms, and that the resolution of this uncertainty over time tends to be accompanied by a decline in the valuation ratios.

The basic idea is simple. Let B denote a firm's book equity today (at time 0) and g its constant growth rate, so the value of book equity at time T is $B \exp(gT)$. Assuming that competition eliminates the firm's expected abnormal earnings by T, the firm's market value at T equals its book value, and the market value today is the expected book value at T discounted at some rate r. If g is unknown and assumed to be normally distributed with mean $\bar{g}$ and variance σ^2, the market-to-book ratio (M/B) today is

$$\frac{M}{B} = \mathrm{E}\left\{\exp\left[(g-r)T\right]\right\} = \exp\left[(\bar{g} + \sigma^2/2 - r)T\right]. \quad (1)$$

The M/B ratio is increasing in the uncertainty about book equity growth, σ^2, thanks to the convex relation between the growth rate and terminal value.[1] We argue that uncertainty declines over a firm's lifetime due to learning. As a result, younger firms have higher σ^2 and hence also higher M/B ratios, holding other things (such as $\bar{g}$ and r) constant.

The valuation model developed in the paper is more realistic than the baby model outlined above, and thus it has a richer set of implications. Firm profitability, measured as the accounting rate of return on book equity, is assumed to revert to an unknown mean whose

[1]To further illustrate the convexity, suppose that firms A and B differ only in the future growth rates of their book equity, whose current value is \$1m. While A is going to grow by 10% per year with certainty, B will grow either by 5% or by 15% with equal probabilities. Due to the convex nature of compounding, the difference between compounding the 15% and 10% growth rates over time is substantially higher than the difference between compounding the 10% and 5% growth rates. Over 10 years, for example, \$1m cumulates to \$4.05m at the 15% rate, to \$2.59m at the 10% rate, and to \$1.63m at the 5% rate. Hence, firm B has a higher expected future book value of equity than firm A (because $(4.05 + 1.63)/2 = 2.84 > 2.59$).

value investors learn over time. Using techniques of Bayesian updating, we obtain a closed-form solution for the firm's M/B ratio. M/B is shown to increase with expected profitability and decrease with expected stock return, consistent with the existing literature. The previous paragraph shows that M/B increases with uncertainty about the average growth rate of book equity. Since the growth rate of book is profitability minus the dividend yield, M/B also increases with uncertainty about average profitability. This uncertainty is shown to have no effect on expected stock returns. Even if firm profitability is correlated with the exogenously specified stochastic discount factor, the updates in the estimates of mean profitability are uncorrelated with it, so any additional return volatility is idiosyncratic. Uncertainty about mean profitability therefore leads to higher valuations because it increases expected future payoffs without affecting the discount rate.

In the presence of external financing constraints, which are common for young firms and assumed in the paper, dividend payouts reduce the growth rate of book equity, thereby weakening the convexity between the growth rate and the future value of book equity. As a result, the model implies that the relation between M/B and the uncertainty about average profitability should be stronger for no-dividend-paying stocks. Calibrating the model to a typical firm in our sample, we find that the effect of the uncertainty on M/B can be quantitatively large, especially for stocks that pay no dividends.

Our empirical analysis confirms the model's predictions on a large panel of data covering the years 1963 through 2000. Since uncertainty declines over time due to learning, the model implies that a younger firm should have a higher M/B ratio than an otherwise identical older firm. Indeed, we find a significantly negative cross-sectional relation between firm age and M/B, even after controlling for other well-known determinants of M/B. We also find that this relation is stronger for firms that pay no dividends, confirming another prediction of the model. The results are both statistically and economically significant. For example, we find a difference of over 5% between the valuations of a typical two-year-old firm and an otherwise identical one-year-old firm. Among firms that pay no dividends, this valuation difference is over 12%, and the difference between the valuations of a typical five-year-old dividend non-payer and an otherwise identical one-year-old non-payer is almost 27%. In addition, age is shown to predict future changes in M/B. The decline of M/B over time is significantly steeper for younger firms, consistent with the model.

One alternative explanation for the observed high valuations of young firms is that these firms are expected to be highly profitable. However, M/B is related to age even after controlling for current and future profitability. Another possibility is that (irrational) investors are on average too optimistic about the future profitability of young firms. Average stock returns for young firms should then be low or even negative, as investors gradually revise their high expectations downward. In contrast, we provide a rational model of learning in which young firms can have high valuations without having low returns. The debate on whether the returns on young firms are too low is still raging.[2] Since we find a relation between M/B and age even after controlling for future returns, the learning effect seems distinct from any potential over-optimism effect. Of course, we do not argue that our model fully explains the observed valuations. We only argue that valuations that appear excessively high at first sight do not necessarily imply that investors are irrational.[3]

Apart from implications for valuation ratios, our model has consequences for return volatility as well. Idiosyncratic return volatility increases with uncertainty about the firm's average profitability as well as with idiosyncratic volatility of profitability. In the data, volatility indeed tends to be higher for younger firms and for firms with more volatile profitability. Volatility is also higher for firms that pay no dividends, as the model predicts.

While the bulk of asset pricing research related to stock valuation focuses on the discount rate, our emphasis is on cash flow. We model cash flow using accounting information such as earnings and book equity, similar to several other recent studies. Ang and Liu (2001), building on Feltham and Ohlson (1999), specify affine processes for selected accounting variables and derive a nonlinear relation between M/B and stochastic interest rates, profitability, and growth in book value. Bakshi and Chen (2001) develop a stock valuation model in which the expected earnings growth rate follows a mean-reverting process, and obtain a number of implications for the price-earnings ratio. Schwartz and Moon (2000) propose a valuation approach in which the expected sales growth rate follows a mean-reverting process with a time-varying drift. They argue that high firm valuations can be justified if the mean and volatility of the sales growth rate are sufficiently high. In addition, Berk, Green, and Naik (1999) and Gomes, Kogan, and Zhang (2001) derive the value of an optimally-investing firm as a sum of assets in place and growth options. In their models, M/B appears in the dynamics of conditional expected returns, helping explain various properties of stock returns. Although all of these papers address various interesting issues related to dynamic stock valuation, none of them incorporates uncertainty and learning about company fundamentals into the valuation framework.

In this paper, learning is done by investors valuing a firm. Related literature explores the

[2] See Schultz (2002) for a literature review and a recent contribution.

[3] This argument is broadly consistent with Lewellen and Shanken (2002) who "emphasize that many tests of market efficiency cannot distinguish between a market with learning and an irrational market."

implications of learning by firms about their own efficiency and resources (e.g. Jovanovich, 1982, Bernardo and Chowdry, 2002). The finance literature on learning also includes Timmerman (1993), David (1997), Veronesi (1999, 2001), Brennan and Xia (2001), Lewellen and Shanken (2002), and others. Among other things, these articles show that learning about a hidden state variable such as the dividend growth rate has equilibrium implications that are in line with certain empirical regularities, such as excess volatility and predictability of stock returns. These articles do not investigate the implications of learning for stock valuation ratios such as M/B, which is the subject of the present article.

The paper is organized as follows. Section 2 develops our valuation approach. The first subsection derives the expressions for M/B and stock returns in the benchmark case when average firm profitability is assumed to be known. The second subsection extends the analysis to the case of unknown average profitability. Section 3 empirically tests the main implications of the model, and Section 4 concludes.

2. The Valuation Framework

This section develops a continuous-time framework for valuing stocks of firms whose profitability is mean-reverting. Profitability is defined as the firm's instantaneous accounting return on equity:

$$\rho_t = \frac{Y_t}{B_t},$$

where Y_t denotes the firm's earnings at time t, and B_t denotes the book value of the firm's equity at time t. Profitability is assumed to follow a simple mean-reverting process:

$$d\rho_t = \phi(\overline{\rho} - \rho_t)dt + \sigma_\rho dW_t, \tag{2}$$

where $\overline{\rho}$ is mean profitability, sometimes also referred to as average profitability, ϕ is the speed of mean reversion, $\sigma_\rho = (\sigma_{\rho,1}, \sigma_{\rho,2})$ is a 1×2 vector of constants, and dW_t is a 2×1 vector of independent Brownian motions. To avoid unnecessary complexity, we do not model the firm's investment policy. Instead, we simply assume that the resulting profitability is mean-reverting, consistent with the existing empirical literature.[4] Note that ρ_t is not required to be positive, which is useful in light of our interest in the valuation of young firms, whose earnings are sometimes negative.

[4]See, for example, Beaver (1970), Lookabill (1976), Freeman, Ohlson, and Penman (1982), Penman (1991), Fama and French (2000), and Pakoš (2001).

The firm is assumed to pay out a constant fraction of its book equity in dividends:

$$D_t = c\ B_t, \qquad c \geq 0. \tag{3}$$

This assumption of constant dividend yield with respect to book value reflects the common corporate policy of smoothing dividends over time.[5] Since young firms often pay no dividends and since the fraction of dividend-paying firms has been decreasing steadily (see Fama and French, 2001a), the special case of $c = 0$ receives particular attention in our analysis.

We assume that the firm is financed only by equity and that there are no new equity issues. External financing is costly for many reasons, including transaction costs and asymmetric information problems (see Myers, 1984, and Myers and Majluf, 1984). Both reasons are likely to be relevant for young firms and firms with uncertain profitability, which are at the focus of this study.

The process for book equity is governed by the clean surplus relation. This accounting identity states that, in the absence of external capital contributions or withdrawals, book equity increases by current earnings and decreases by current dividends:[6]

$$dB_t = (Y_t - D_t)\, dt = (\rho_t - c)\, B_t\ dt. \tag{4}$$

In words, the growth rate of book equity equals profitability minus the dividend yield.

To build a bridge between accounting values and market values, we assume that the market value of equity will equal the book value at some future time T. To motivate this assumption, consider the abnormal earnings model of Ohlson (1990, 1995) and Feltham and Ohlson (1995), which is essentially just another accounting identity. This model equates market equity with book equity plus the discounted sum of abnormal earnings, defined as earnings in excess of those earned at the rate equal to the cost of capital. Since firms are often created to pursue profitable ventures based on novel ideas and technologies, which are often patent-protected, the profitability of a new firm can be abnormal for an extended period of time. However, it seems reasonable to assume that abnormal earnings are gradually reduced by competitive market forces. We assume that, at some known future time T, the firm reaches "maturity" in the sense that the present value of expected future abnormal earnings is zero. At that point, $M_T = B_T$.

[5]Some simple alternative assumptions about dividend policy seem less appealing. For example, assuming a constant ratio of dividends to earnings gives rise to negative dividends when earnings are negative. Assuming a constant ratio of dividends to market equity leads to dividend fluctuations that seem unrealistically high. The possibility of making c a decreasing function of ρ_t is discussed in Section 2.3.

[6]Some violations of clean surplus accounting exist. They include accounting for unrealized gains and losses on financial items, foreign currency translation gains and losses, and tax benefits on preferred dividends. Brennan and Schwartz (1982) rely on equation (4) to study the effects of regulatory policy on firm value.

Some readers may find it unnatural to relate market value to book value, as many accounting features seem irrelevant for valuation. Such readers may find it useful to replace earnings by net cash flows and to view book value as the cumulative net cash flow that is reinvested in the firm. The assumption that $M_T = B_T$ then means that average net future cash flow is zero at maturity, and all of our theoretical results can be interpreted from the cash flow perspective. Our choice of accounting-based exposition is motivated mainly by practical concerns – the concepts of earnings, book value, and profitability are well-known and commonly reported, making the model easier to test and to use in practice.

The market value of the firm's equity at any point in time is given by the sum of the discounted value of all future dividends and the terminal value $M_T = B_T$:

$$M_t = E_t \left[\int_t^T \frac{\pi_s}{\pi_t} D_s ds \right] + E_t \left[\frac{\pi_T}{\pi_t} B_T \right]. \tag{5}$$

The stochastic discount factor π_t is assumed to follow the log-normal process

$$\frac{d\pi_t}{\pi_t} = -r dt - \sigma_\pi dW_t, \tag{6}$$

where r denotes a constant risk-free interest rate and $\sigma_\pi = (\sigma_{\pi,1}, 0)$ is a 1×2 vector. We assume that changes in π_t are driven only by the first element of dW_t (i.e. $\sigma_{\pi,2} = 0$), so that we can interpret $dW_{1,t}$ as a systematic shock and $dW_{2,t}$ as an idiosyncratic shock.

The valuation framework presented above can be generalized in numerous ways. Many assumptions can be relaxed or modified without affecting any conclusions, as discussed in Section 2.3. Our aim is to present a parsimonious yet reasonably realistic model that highlights the importance of learning about mean profitability in stock valuation.

2.1. No Uncertainty about Mean Profitability

This subsection assumes perfect knowledge of all model parameters, including mean profitability $\overline{\rho}$. The results developed here are useful for comparison with the next subsection, in which $\overline{\rho}$ is treated as unknown. The following function is used repeatedly:

$$Z(\overline{\rho}, \rho_t, s) = \exp\left\{ -(r + c - \overline{\rho}) s + \frac{1}{\phi} \left(1 - e^{-\phi s}\right) (\rho_t - \overline{\rho}) + Q(s) \right\}, \tag{7}$$

where

$$Q(s) = \frac{\sigma_\rho \sigma_\rho'}{2\phi^3} \left[\frac{1 - e^{-2\phi s}}{2} + \phi s - 2\left(1 - e^{-\phi s}\right) \right] + \frac{\sigma_\pi \sigma_\rho'}{\phi^2} \left(1 - e^{-\phi s} - \phi s\right). \tag{8}$$

As shown in the appendix, the function Z has a simple economic interpretation as expected discounted growth in book equity:

$$Z\left(\overline{\rho}, \rho_t, s\right) = E_t\left[\frac{\pi_{t+s} B_{t+s}}{\pi_t B_t}\right]. \tag{9}$$

The following proposition, proved in the appendix, obtains closed-form solutions for the firm's M/B ratio and for the first two moments of stock returns. Throughout, $\tau = T - t$.

Proposition 1. Assume $\overline{\rho}$ is known with certainty. Then

(a) The firm's ratio of market value of equity to book value of equity is given by

$$\frac{M_t}{B_t} = G\left(\overline{\rho}; \rho_t, \tau\right) \equiv c \int_0^\tau Z\left(\overline{\rho}, \rho_t, s\right) ds + Z\left(\overline{\rho}, \rho_t, \tau\right). \tag{10}$$

(b) The process for excess stock returns, $dR_t = \left(dM_t + D_t dt\right)/M_t - r dt$, is given by

$$dR_t = \mu_{R,t} dt + \sigma_{R,t} dW_t,$$

where

$$\mu_{R,t} = F(\rho_t, \tau, c)\, \frac{\sigma_\rho \sigma_\pi'}{\phi}, \tag{11}$$

$$\sigma_{R,t} = F(\rho_t, \tau, c)\, \frac{1}{\phi} \sigma_\rho, \tag{12}$$

and $F(\rho_t, \tau, c)$ is given in equation (33) in the appendix. In the special case when the firm pays no dividends ($c = 0$), we have

$$\mu_{R,t} = \left(1 - e^{-\phi\tau}\right) \frac{\sigma_\rho \sigma_\pi'}{\phi}, \tag{13}$$

$$\sigma_{R,t} = \left(1 - e^{-\phi\tau}\right) \frac{1}{\phi} \sigma_\rho. \tag{14}$$

Note that the expressions for M/B and the moments of stock returns are fully explicit for $c = 0$. In the presence of dividends, $c > 0$, the integral in equation (10) as well as its counterpart in the expression for $F(\rho_t, \tau, c)$ in the appendix can be computed in a split-second using standard numerical techniques. Moreover, the solution is sufficiently explicit to allow full characterization of the properties of the M/B ratio. These properties are summarized in the following corollaries, obtained immediately from equations (7) and (10) with the help of the fact that $s\phi > 1 - e^{-\phi s} > 0$ for all $s > 0$.

Corollary 1. The M/B ratio increases if

(i) average profitability $\overline{\rho}$ increases;

(ii) current profitability ρ_t increases;

(iii) $\sigma_\rho \sigma_\pi'$, the key determinant of expected excess stock return $\mu_{R,t}$, decreases;

(iv) the interest rate r decreases;

(v) $\sigma_\rho \sigma_\rho'$, the instantaneous variance of the profitability process, increases.

Corollary 2. The M/B ratio is convex in $\overline{\rho}$.

The results in Corollary 1 are intuitive. Increases in $\overline{\rho}$ or ρ_t push M/B up because they raise expectations of future profitability.[7] Decreases in $\mu_{R,t}$ or r lift M/B as they reduce expected stock returns. In other words, M/B increases with expected accounting returns (profitability) and decreases with expected stock returns, as discussed for example in Vuolteenaho (2000). More volatile profitability increases M/B because it increases expected future payoffs due to the convex relation between future payoffs and their growth rate. This relation is discussed later in the paper in more detail.

The M/B ratio also depends on the speed of mean reversion ϕ and the dividend yield c, but the dependence is slightly more complex. The effect of ϕ on M/B is ambiguous as a result of two effects working in the opposite directions. Increasing ϕ ensures that ρ_t varies less around its mean $\overline{\rho}$. Since future payoffs become more certain, both systematic and nonsystematic components of return volatility decline, together with the required rate of return, and M/B increases. On the other hand, the lower dispersion of ρ_t decreases expected future payoffs, again due to the convex relation between future payoffs and their growth rate, and M/B decreases. The effect of c on M/B is ambiguous as well. A higher dividend yield increases M/B since the payoffs are received earlier on average, but it also decreases M/B since it reduces the growth rate of book equity. One would expect that when profitability is sufficiently high, paying dividends reduces M/B, and vice versa. This intuition is confirmed in Figure 1, which plots M/B against $\overline{\rho}$ for three values of c: 0, 0.04, and 0.10. For sufficiently large values of $\overline{\rho}$, increasing c reduces M/B, and vice versa.

Figure 1 is constructed using parameter values that correspond to typical values observed in the data (described in the next section). The value of ρ_t is set equal to 0.11, which is the grand median, across stocks and years, of all valid annual returns on equity (ROE) in our sample. To estimate ϕ and σ_ρ, we estimate an AR(1) model for each stock's ROE, using the longest continuous series of the stock's valid annual ROEs. That series is required to be at least 10 years long (using 15 or 20 years leads to almost identical results). The slope

[7] It is easy to show from equation (2) that $\mathrm{E}_t(\rho_{t+s}) = (1 - e^{-\phi s})\overline{\rho} + e^{-\phi s}\rho_t$.

coefficients are adjusted for small-sample bias following the approach of Marriott and Pope (1954) and Kendall (1954). The value of $\phi = 0.3968$ used in the calibration is the median of the estimated coefficients across all 4,571 stocks.[8] In addition, we choose $\tau = 15$ years, $r = 0.03$, $\sigma_{\pi,1} = 0.6$, $\sigma_{\rho_1} = 0.0584$, and $\sigma_{\rho_2} = 0.0596$. The latter three values were chosen to imply $\mu_{R,t} = 0.0883$, a reasonable value for expected annual excess return in light of the data, as well as $\sqrt{\sigma_\rho \sigma_\rho'} = 0.0834$, which is equal to the median of the residual volatility of profitability across all 4,571 stocks. Finally, the grand median of the ratios of common stock dividends to last year's book equity (c) is 0.0105. (Interestingly, the cross-sectional median is zero in each year since 1985. Most firms do not pay dividends nowadays.) The grand median of c calculated only across firms that pay dividends is 0.0434. These numbers have guided our choices for c in Figure 1 as well as in the subsequent figures.

Figure 1 shows that the relation between M/B and $\overline{\rho}$ is convex. This relation is stated formally in Corollary 2, and the proof of the convexity of $G(\overline{\rho}; \rho_t, \tau)$ in equation (10) is provided in the appendix. The intuition is explained in the introduction. Also note that the convexity is higher at higher growth rates. Since dividends reduce the growth rate of book equity, the convexity of M/B in $\overline{\rho}$ should be less pronounced for higher values of c. This observation is also confirmed in Figure 1.

Finally, the stock return process displayed in Proposition 1 exhibits a mechanical property worth commenting on. As $\tau \to 0$, $F(\rho_t, \tau, c) \to 0$ and hence also $\sigma_{R,t} \to 0$ and $\mu_{R,t} \to 0$. In words, as maturity approaches, stock returns become less risky, and at maturity they become riskless. This property is an artifact of the finite horizon assumption, since at time T, the stockholder receives a locally riskless payoff B_T. We are not overly concerned about the return behavior close to maturity, as we are interested in valuing young firms whose horizons are relatively long. If the horizon τ is long enough, expected returns and return volatility are essentially constant over time. For example, using the parameter values from Figure 1, we have $\mu_{R,t} = 7.1\%$ for $\tau = 5$ years, $\mu_{R,t} = 8.0\%$ for $\tau = 10$ years, and $\mu_{R,t} = 8.1\%$ for $\tau = 15$ years. Moreover, these finite horizon effects disappear in frameworks with random or infinite horizon (see Section 2.3.), which have very similar qualitative implications.

2.2. Learning about Mean Profitability

This subsection investigates the effects of uncertainty and learning about mean profitability on valuation and stock returns. Throughout, mean profitability $\overline{\rho}$ is treated as unknown.

[8] Fama and French (2000) estimate the average rate of mean reversion in the return on assets as 0.38.

At any time t, investors' beliefs about $\overline{\rho}$ are summarized by a probability density function $p_t(\overline{\rho})$, and investors rationally update their beliefs about $\overline{\rho}$ over time.

Equations (5), (10), and the law of iterated expectations immediately imply

$$\frac{M_t}{B_t} = \int_R G\left(\overline{\rho};\rho_t,\tau\right)\ p_t\left(\overline{\rho}\right)\ d\overline{\rho}. \tag{15}$$

Loosely speaking, the firm's M/B ratio under uncertainty is the average of the M/B ratios for all possible values of $\overline{\rho}$, weighted by the current probabilities assigned to each $\overline{\rho}$.

One of our key results is that higher uncertainty about $\overline{\rho}$ increases M/B. Recall from Corollary 2 that $G\left(\rho_t,\overline{\rho};\tau\right)$ is a convex function of $\overline{\rho}$. Hence, greater dispersion in $p_t\left(\overline{\rho}\right)$ increases the expected value of $G\left(\rho_t,\overline{\rho};\tau\right)$. (Greater dispersion is interpreted as a mean-preserving-spread, defined in Rothschild and Stiglitz, 1970.) Higher uncertainty about $\overline{\rho}$ increases the probability that future growth rate of book equity will be persistently high or persistently low. Due to the convex nature of compounding, a persistently high growth rate has a bigger impact on the future book value than a persistently low growth rate and the expected future book value increases, together with M/B.

To obtain closed-form solutions for prices and returns, we assume that the prior distribution $p_0\left(\overline{\rho}\right)$ at time $t=0$ is normally distributed, and that investors update their beliefs using the Bayes rule. The resulting posterior distribution $p_t\left(\overline{\rho}\right)$ is also normal, as shown by the following lemma, proved in the appendix.

Lemma 1. Suppose the prior distribution of $\overline{\rho}$ at time $t=0$ is normal, $\overline{\rho}\sim\mathcal{N}(\widehat{\rho}_0,\widehat{\sigma}_0^2)$. Then the posterior distribution of $\overline{\rho}$ at time $t>0$ conditional on $\mathcal{F}_t=\{(\pi_\tau,\rho_\tau):0\leq\tau\leq t\}$ is also normal, $\overline{\rho}|_{\mathcal{F}_t}\sim\mathcal{N}(\widehat{\rho}_t,\widehat{\sigma}_t^2)$, where

(a) The conditional mean $\widehat{\rho}_t=E[\overline{\rho}|\mathcal{F}_t]$ evolves according to the process

$$d\widehat{\rho}_t=\widehat{\sigma}_t^2\frac{\phi}{\sigma_{\rho,2}}d\widetilde{W}_{2,t}, \tag{16}$$

where $\widetilde{W}_{2,t}$ is the idiosyncratic component of the Wiener process capturing investors' perceived expectation errors (see equation 35 in the appendix).

(b) The mean squared error $\widehat{\sigma}_t^2=E[(\overline{\rho}-\widehat{\rho}_t)^2|\mathcal{F}_t]$ is non-stochastic and given by

$$\widehat{\sigma}_t^2=\frac{1}{\frac{1}{\widehat{\sigma}_0^2}+\frac{\phi^2}{\sigma_{\rho,2}^2}t}. \tag{17}$$

Note that the variance $\widehat{\sigma}_t^2$ of the posterior distribution declines over time due to learning, and the speed of this decline increases with the ratio $\phi/\sigma_{\rho,2}$. A higher value of ϕ means that ρ_t will be close to the mean $\overline{\rho}$ more often, making it easier to learn the value of the latter. A smaller value of $\sigma_{\rho,2}$ also leads to quicker learning because it implies a less noisy profitability process. Figure 2 plots the evolution of $\widehat{\sigma}_t$ using the same parameter values as in Figure 1, except that ϕ assumes three different values, 0.2, 0.4, and 0.6. Note that learning is rather slow. With the prior standard deviation of $\widehat{\sigma}_0 = 0.10$, or 10% per year, the posterior standard deviation $\widehat{\sigma}_t$ after 10 years is still large, over 4%. This number corresponds to $\phi = 0.4$, obtained for the median firm in our sample. Even when mean reversion is faster, $\phi = 0.6$, the uncertainty after 10 years remains substantial, about 3%.

It is useful to define the following function:

$$Z^U\left(\widehat{\rho}_t, \widehat{\sigma}_t^2, \rho_t, s\right) = Z\left(\widehat{\rho}_t, \rho_t, s\right) \exp\left\{\frac{1}{2\phi^2}\widehat{\sigma}_t^2\left(\phi s - 1 + e^{-\phi s}\right)^2\right\}, \tag{18}$$

where $Z\left(\widehat{\rho}_t, \rho_t, s\right)$ is defined in equation (7). Proposition 2, proved in the appendix, is a counterpart of Proposition 1 in the case of unknown mean profitability. As before, $\tau = T - t$.

Proposition 2. Assume that investors rationally learn about the unknown value of $\overline{\rho}$, and that the assumptions stated in Lemma 1 are satisfied. Then

(a) The firm's ratio of market value of equity to book value of equity is given by

$$\frac{M_t}{B_t} = c\int_0^\tau Z^U\left(\widehat{\rho}_t, \widehat{\sigma}_t^2, \rho_t, s\right) ds + Z^U\left(\widehat{\rho}_t, \widehat{\sigma}_t^2, \rho_t, \tau\right). \tag{19}$$

(b) The process for excess stock returns, $dR_t = \left(dM_t + D_t dt\right)/M_t - r dt$, is given by

$$dR_t = \mu_{R,t} dt + \sigma_{R,t} d\widetilde{W}_t,$$

where

$$\mu_{R,t} = F\left(\rho_t; \widehat{\rho}_t, \widehat{\sigma}_t^2, \tau, c\right) \frac{1}{\phi}\sigma_\rho \sigma_\pi' \tag{20}$$

$$\sigma_{R,t} = F\left(\rho_t; \widehat{\rho}_t, \widehat{\sigma}_t^2, \tau, c\right) \frac{1}{\phi}\sigma_\rho + F_1\left(\rho_t; \widehat{\rho}_t, \widehat{\sigma}_t^2, \tau, c\right) \frac{1}{\phi}\widehat{\sigma}_{\rho,t} \tag{21}$$

$$\widehat{\sigma}_{\rho,t} = \left(0,\ \widehat{\sigma}_t^2 \frac{\phi}{\sigma_{\rho,2}}\right). \tag{22}$$

The functions $F\left(\rho_t; \widehat{\rho}_t, \widehat{\sigma}_t^2, \tau, c\right)$ and $F_1\left(\rho_t; \widehat{\rho}_t, \widehat{\sigma}_t^2, \tau, c\right)$ are strictly positive and given in equations (41) and (42) in the appendix. In the special case when the firm pays no dividends ($c = 0$), we have

$$\mu_{R,t} = \left(1 - e^{-\phi\tau}\right) \frac{\sigma_\rho \sigma_\pi'}{\phi} \tag{23}$$

$$\sigma_{R,t} = \left(1 - e^{-\phi\tau}\right) \frac{1}{\phi}\sigma_\rho + \left(\phi\tau - 1 + e^{-\phi\tau}\right) \frac{1}{\phi}\widehat{\sigma}_{\rho,t}. \tag{24}$$

2.2.1. Implications for market-to-book

The expression for M/B in Proposition 2 (where $\overline{\rho}$ is unknown) has the same form as in Proposition 1 (where $\overline{\rho}$ is known), except that the function Z in equation (10) is replaced by Z^U in equation (19). There are two differences. First, the known value of $\overline{\rho}$ is replaced by its current estimate $\widehat{\rho}_t$, which varies over time. Second, Z^U is equal to Z multiplied by a term related to the uncertainty about mean profitability (see equation 18). Since this additional term involves only $\widehat{\sigma}_t^2$, ϕ, and s, Corollary 1 applies here as well, with $\overline{\rho}$ replaced by $\widehat{\rho}_t$. The additional term increases with $\widehat{\sigma}_t^2$, which implies that higher $\widehat{\sigma}_t^2$ increases M/B as well. This fact is highlighted in the following corollary, which follows immediately from equation (18).

Corollary 3. The M/B ratio increases with uncertainty about mean profitability, $\widehat{\sigma}_t^2$.

The intuition behind this effect is provided earlier. As one would expect, the effect strengthens with horizon, since the term multiplying $\widehat{\sigma}_t^2$ in equation (18) increases with τ. Based on Figure 1, one might also expect this effect to be stronger for firms that pay no dividends. This insight is formalized in the following corollary, proved in the appendix.

Corollary 4. The effect of $\widehat{\sigma}_t^2$ on log(M/B) is stronger for firms that pay no dividends.

Corollaries 3 and 4 are investigated in the empirical analysis. Note that it is also possible to prove a related claim that there exists a time horizon T^* such that for every $T > T^*$, the effect of $\widehat{\sigma}_t^2$ on M/B becomes stronger as the dividend yield c decreases (i.e., $\partial^2 \left(M_t/B_t\right) / \partial\widehat{\sigma}_t^2 \partial c < 0$). The numerical analysis of that case reveals that $T^* = 0$ for most reasonable parameter values, including those used in Figures 1–4.

To assess the strength of the effect of $\widehat{\sigma}_t$ on M/B for a typical firm, the top panel of Figure 3 plots M/B against $\widehat{\sigma}_t$ for various levels of the dividend yield c. All other parameters are as in Figure 1. The effect of $\widehat{\sigma}_t$ on M/B is substantial. For a firm that pays no dividends, M/B equals 1.41 under perfect certainty ($\widehat{\sigma}_t = 0$). This M/B rises to 1.72 when the standard deviation around the current estimate of mean profitability is 5% per year ($\widehat{\sigma}_t = 0.05$), and to 3.08 when $\widehat{\sigma}_t = 0.10$. The effect is the strongest for $c = 0$, supporting Corollary 4.

As time passes, $\widehat{\sigma}_t$ declines as investors learn about $\overline{\rho}$. This resolution of uncertainty leads to a gradual decline in M/B (Corollary 3). The top panels of Figure 4 plot the evolution of M/B over time, using the parameter values from Figure 1. The figure assumes $\rho_t = \widehat{\rho}_t$ for all t, to highlight the effect of the resolution of uncertainty and to remove the effect of learning that $\overline{\rho}$ is higher or lower than expected. Prior uncertainty about $\overline{\rho}$ is chosen based on the cross-sectional dispersion of average annual ROE. When at least 10 (20) valid ROEs are required to compute the average, the dispersion is 0.1368 (0.0728). Our choice of $\hat{\sigma}_0 = 0.10$ per year is a round-number compromise between these two values.

The top-left panel of Figure 4 shows a rapid decline in M/B over the first few years of the firm's life. If the firm pays no dividends, its M/B declines almost by half, from just over 3.0 to about 1.6, over the first five years. With a 4% dividend yield, the M/B ratio falls from about 2.4 at birth to about 1.4 after five years. The top-right panel shows that M/B declines slightly over time even when $\overline{\rho}$ is known. This is an artifact of our finite horizon assumption, which implies that M/B must converge to 1 as maturity approaches. However, if the horizon is sufficiently long, M/B under certainty is virtually constant over time. Comparing the top two panels of Figure 4 reveals strong effects of uncertainty and learning on the time path of M/B of a young firm. M/B of a new firm is high due to uncertainty about its future profitability, but it declines quickly as investors learn about $\overline{\rho}$.

To conclude the description of the various properties of M/B, consider a firm that pays no dividends ($c = 0$). It follows from equations (18) and (19) that

$$\log\left(\frac{M_t}{B_t}\right) = \alpha_0(\tau) + \alpha_1(\tau)\,\widehat{\rho}_t + \alpha_2(\tau)\,\rho_t + \alpha_3(\tau)\,\mu_{R,t} + \alpha_4(\tau)\,\sigma_\rho\sigma_\rho' + \alpha_5(\tau)\,\widehat{\sigma}_t^2, \qquad (25)$$

where the coefficients $\alpha_1(\tau)$ through $\alpha_5(\tau)$ depend only on horizon. As shown in the appendix, $\alpha_1(\tau) > 0$, $\alpha_2(\tau) > 0$, $\alpha_3(\tau) < 0$, $\alpha_4(\tau) > 0$, and $\alpha_5(\tau) > 0$. In words, log M/B increases with the estimate of average profitability as well as with current profitability, decreases with expected stock return, increases with the volatility of profitability, and, last but not least, increases with the uncertainty about average profitability. While deriving an analogous expression for dividend-paying stocks appears infeasible, we view equation (25) as motivation for the log-linear functional form specified in the empirical analysis in Section 3.

2.2.2. Implications for stock returns

The first obvious implication for stock returns is that return volatility increases with the volatility of profitability, which is apparent from equation (21). This relation is present even when $\overline{\rho}$ is known, as can be seen from equation (12). Another intuitive implication is that uncertainty about $\overline{\rho}$ amplifies the volatility of stock returns:

Corollary 5. Idiosyncratic return volatility increases with $\widehat{\sigma}_t^2$.

This result follows immediately from equation (21). Similar results for aggregate market volatility have been obtained in the existing literature on learning, as cited in the introduction. The intuition here is simple. Market value of equity is quite sensitive to $\overline{\rho}$, as illustrated in Figure 1. When the true value of $\overline{\rho}$ is unknown, the perception of $\overline{\rho}$ changes over time due to learning, and the market value fluctuates by more than it would if $\overline{\rho}$ were known.

Corollary 6. Firms that pay no dividends have higher return volatility than dividend payers.

The intuition behind this corollary, proved in the appendix, is as follows. If the firm pays dividends, market value of equity depends not only on the terminal value B_T but also on near-term dividends, which are less sensitive to $\overline{\rho}$. (At the extreme, the dividend received in the next instant is riskless.) As a result, the sensitivity of the market value to $\overline{\rho}$ is higher for dividend non-payers. This higher sensitivity to $\overline{\rho}$ then translates into higher return volatility, using the intuition behind Corollary 5.

The bottom panel of Figure 3 plots return volatility against $\widehat{\sigma}_t$. Following Corollaries 5 and 6, volatility increases with $\widehat{\sigma}_t$ and is higher when no dividends are paid. For example, an increase of $\widehat{\sigma}_t$ from zero to 5% raises the volatility from 20% to 35% for a stock with 4% dividend payout, and to about 40% for a non-dividend-paying stock. The bottom panels of Figure 4 plot the evolution of volatility over time with and without learning, assuming as before that $\rho_t = \widehat{\rho}_t$ for every t. Volatility declines substantially in the presence of learning, but it remains essentially constant when there is no learning. (It declines to zero close to maturity only as an artifact of the finite-horizon model.) Also note that with no uncertainty/learning, there is almost no difference between the volatilities of dividend payers and non-payers. Volatility is slightly lower for the payers simply because dividends are instantaneously riskless, but the intuition behind Corollary 6 is missing and the effect of dividends on volatility is negligible compared to the uncertainty case.

Finally, consider the effect of learning on expected stock returns. Comparing Propositions 1 and 2, expected returns in equations (13) and (23) are identical in the absence of dividends. In other words, uncertainty and its resolution have no effect on expected returns. Although return volatility is higher in the presence of uncertainty about $\overline{\rho}$, all additional volatility is idiosyncratic, as it has no loading on the systematic shock $d\widetilde{W}_1$ (see equation 22). There is no reason why learning about a firm-specific parameter $\overline{\rho}$ should be related to the exogenously specified stochastic discount factor. As is apparent from equation (16), innovations in the posterior mean of $\overline{\rho}$ are uncorrelated with $d\widetilde{W}_1$ and hence also with the stochastic discount factor. In the absence of dividends, this zero correlation implies that the covariance between stock returns and the stochastic discount factor is unaffected by uncertainty about $\overline{\rho}$, so that expected returns under certainty and uncertainty are equal. The same intuition applies in the presence of dividends, but their perfect local predictability induces a slight dependence of expected returns on $\widehat{\sigma}_t$. (This dependence is very weak numerically.) With this minor caveat, we conclude that the learning process in our framework is idiosyncratic and does not affect expected returns. It does affect expected cash flows, as discussed earlier, and therefore plays an important role in stock valuation.

2.3. Robustness and Extensions

The results presented earlier are remarkably robust to various modifications and generalizations of our simple framework. To start with, while we assume that the firm is financed entirely by equity, there is a weaker assumption that delivers identical theoretical results. Debt financing can be allowed under the condition that its dynamics do not affect the profitability process in equation (2). That is likely to be the case when the firm maintains an approximately constant ratio of debt to equity, for example. On a related point, our results do not hinge on the assumption of no new equity issues (which can be viewed as negative dividends). It can be shown by continuity that, for any given domain of ρ_t, there exists a $c^* < 0$ such that all results go through for each $c > c^*$. Hence, all we need is some upper bound on external financing. Intuitively, our results follow as long as reinvested earnings are the primary source of growth in the firm's book equity.

Our basic framework uses constant c for simplicity and to be broadly consistent with empirical evidence. However, the firm's optimal dividend policy is likely to be time-varying. When expected profitability is low relative to the cost of capital, the firm raises dividends and perhaps even shuts down (i.e. pays out a dividend equal to the firm's liquidation value), thereby increasing its market value compared to the constant-dividend-policy scenario. When expected profitability is high, the firm reduces dividends and perhaps even raises more capital, again increasing the market value. Only when expected profitability is typical, there is little adjustment to dividend policy and to market value. While incorporating such dividend timing in our model appears too difficult, it would strengthen our results, because it amplifies the convexity between the market value and average profitability.

Assuming a stochastic process for the interest rate in equation (6) would present no complications. With the interest rate dynamics à la Vasicek (1977), our specification of the stochastic discount factor is the same as in Berk, Green, and Naik (1999) and all results go through. Allowing σ_ρ to vary over time would not affect the qualitative results either. Also, the results do not hinge on the assumption that time to maturity is known. We have examined a framework in which T is unknown and maturity can take place any time with certain probability. Although that framework is less tractable, the basic convexity relation remains and the qualitative implications are very similar to those presented here.

We do not model the firm's development after time T, but we have also explored a more complex framework that does and found the same results. In that alternative framework, the assumption that $M_T = B_T$ is replaced with the assumption that $\overline{\rho}$ jumps to the cost of capital at time T and remains there forever. The market value in the absence of dividends is calculated as the limit of the discounted book value at infinity. The M/B ratio at T is no longer certain to be equal to one, because ρ_t continues to evolve randomly after T. As a result, the return volatility no longer goes to zero at T, but the results are otherwise very similar to those reported here, with both M/B and return volatility positively related to uncertainty about $\overline{\rho}$. We chose to present the model with $M_T = B_T$ mainly because it is easier to explain and the effect of uncertainty is more transparent.

The practical use of our valuation model must consider the convention of conservative accounting (that profits are recognized when earned but losses when anticipated), which implies that M/B could exceed one even after profits are competed away. In addition, the absence of intangible assets from the firm's books could lead to differences in the long-run M/B ratios across industries, as the steady-state reliance on intangibles varies across industries. These issues present no problem for us – all results go through when $M_T = (1 + \eta)B_T$, where η is a constant. Even the specifications with random η modeled as a function of profitability are tractable and yield equivalent implications.

It seems worth remarking that our framework with uncertainty about a constant $\overline{\rho}$ has the same pricing implications as an alternative framework in which known $\overline{\rho}$ varies over time following the law of motion from Lemma 1. These two frameworks are observationally equivalent, but the learning framework strikes us as more plausible. While the convergence of $\hat{\sigma}_t$ to zero in equation (17) appears easy to justify through learning, a similar simple justification seems elusive in the alternative framework with certainty. On a related note, the framework in which known $\overline{\rho}$ follows a more general stochastic mean-reverting process is also tractable and delivers similar pricing implications.

To summarize, the simple framework presented here can be extended in various ways, adding complexity but not affecting the main conclusions. The parsimony of the basic specification reflects our desire to present the simplest possible model that illustrates the effects of uncertainty about mean profitability on stock valuation and returns.

3. Empirical Analysis

This section empirically investigates some implications of the learning model presented in the previous section. One important implication is that firms with high uncertainty about future profitability should have high M/B ratios (Corollary 3). In the model, uncertainty about mean profitability $\bar{\rho}$ declines over time due to learning. Hence, other things equal, the M/B ratio of a typical young firm should be high and it should fall as investors learn about the firm's prospects. Of course, M/B for any given firm can rise if investors learn that profitability is higher than expected. However, assuming that investors' expectations about the profitability of young firms are right on average, we expect to find a negative cross-sectional relation between M/B and firm age, after controlling for other known determinants of M/B. This section provides corroborating evidence on this point. Our model also implies that the effect of age on M/B should be stronger for firms that pay no dividends (Corollary 4). Besides, M/B is predicted to increase with both the level and the volatility of profitability, and to decrease with expected future returns (Corollary 1). All of these implications are also confirmed in the empirical analysis.

In addition to the predictions about M/B, the model has implications for stock return volatility. Idiosyncratic volatility should be high for firms with higher uncertainty about average profitability (Corollary 5), firms that pay no dividends (Corollary 6), and firms with more volatile profitability. Ample support for these predictions is found in the data.

We use annual data for the years 1962 through 2000 extracted from the CRSP/Compustat database. (The appendix describes the data in detail.) A crucial variable in our empirical investigation is firm age. Similar to Fama and French (2001b), we consider each firm as "born" in the year of its first appearance in the CRSP database.[9] Specifically, we look for the first occurence of a valid stock price on CRSP, as well as the first occurrence of the valid market value in the CRSP/Compustat database, and take the earlier of the two. The firm's age is assigned the value of one in the year in which the firm is born and increases by one

[9]Fama and French note that a vast majority of their new lists are IPOs, especially after 1972. About 80% of the firms in our sample are born after 1972.

in each subsequent year. The maximum age in our sample is 76 years (the age in 2000 of a firm born in 1925), and the grand median (across stocks and years) is 11 years.

Table 1 reports some summary statistics of various characteristics computed across firms of the same age (regardless of the calendar year in which that age was reached). Median M/B declines monotonically with age, from 2.25 for the newly-listed firms to 1.25 for firms that are 10 years old. The grand median of M/B is 1.41. Return volatility, expressed in monthly percentage terms, also declines with age. Profitability (ROE) is the highest in year 1, 12.93% per year, but flat around 11% thereafter. Stock returns, which are not risk-adjusted in any way, are negative in the first three years. Finally, as firms grow older, they tend to grow in size and become more levered and more likely to pay dividends. The cross-sectional standard deviation of ROE is rather flat across different ages.

3.1. Results for Market-to-Book

Figure 5 shows that ageing in the life of a firm tends to be accompanied by a decrease in the M/B ratio. For each age, the solid line in the top panel plots the median M/B across all firms of that age. At 2.25, the median is the highest for the newly-listed firms, and it declines for about 10 years after listing. The decline in M/B during a typical firm's childhood, as well as the convex pattern of this decline, are consistent with the learning model presented in the previous section. As investors learn about a young firm's mean profitability, uncertainty is resolved and M/B falls.[10] The top panel also plots the median M/B separately for dividend payers and non-payers. Each year, a firm is classified as a payer or non-payer depending on whether it paid any dividends in that year. The dashed line reveals a pronounced age pattern in the M/B ratios of the non-payers. The median M/B, equal to 2.73 for the newly-listed non-payers, drops almost monotonically for at least 20 years after listing. A typical 10-year-old non-payer has M/B of 1.37, only half of M/B of a newborn non-payer. The dotted line shows a much weaker relation between age and M/B for the payers. The median M/B ratio declines shortly after listing, from 1.55 in year one to 1.20 in year three, but the pattern is flat beyond year three. The observation that the age pattern in M/B is stronger for non-payers than for payers is again consistent with the learning model.

The top panel of Figure 5, obtained from the data, can be compared to the top-left panel of Figure 4, obtained by calibrating the model with reasonable parameter values. The similarity between the shapes of the M/B paths in the model-based and data-based figures

[10] Related results on M/B are reported by Fama and French (2001b) and Jain and Kini (1994).

is striking. This comparison provides some encouragement for the idea that learning effects might be relevant for valuation, but it provides no substitute for the more careful analysis provided in the rest of this section.

Figure 6 plots the evolution of M/B in calendar time. Each year, firms are categorized as young or old depending on whether their age exceeds the midpoint between the minimum and maximum age in the cross-section. Firms are also separately sorted into dividend payers and non-payers. Four groups of stocks are formed each year by intersecting the two independent sorts, and their median M/B ratios are plotted in the top panel. Consistent with Figure 5, young firms and dividend non-payers have higher M/B than old firms and dividend payers in almost every year. The plot also shows that while there is a substantial difference between the median M/B of the young and old dividend non-payers, that difference is much smaller for the payers. All of these observations also emerge later in this section in the regression analysis of the relation between M/B, age, and dividend payouts.

Of course, the relation between M/B and age observed in Figures 5 and 6 could be to some extent due to effects distinct from learning. For example, younger firms should have higher M/B if they have higher profitability or lower returns than older firms (see Table 1). Cohen, Polk, and Vuolteenaho (2001) report that about 80% of the cross-sectional dispersion in M/B is due to the cross-sectional dispersion in expected future profitability, and the remaining 20% is due to the dispersion in expected future returns. The key question is whether the negative relation between M/B and age remains after controlling for the known determinants of M/B ratios. This issue is investigated next in a regression framework.

Each year between 1963 and 2000, we regress M/B cross-sectionally on a function of firm age and other potential determinants of M/B:

$$\begin{aligned}\log(M/B)_i \;&=\; a + b\; AGE_i + c\; DD_i + d\; LEV_i + e\; SIZE_i + f\; VOLP_i + g_0\; ROE(0)_i \;\;\ldots \\ &\quad + g_1\; ROE(1)_i + \ldots + g_q\; ROE(q)_i \;\;\ldots \\ &\quad + h_1\; RET(1)_i + \ldots + h_q\; RET(q)_i, \qquad\qquad i = 1, \ldots, N, \end{aligned} \tag{26}$$

where N denotes the number of firms with valid data in the current year, and the right-hand-side variables are described below. Taking the model seriously, we use the natural logarithm of M/B, motivated by the theoretical prediction in equation (25).

We also turn to the model to define the AGE variable as minus the reciprocal of one plus the firm age. This choice is based on equation (17), which prescribes a particular functional form for the relation between uncertainty and age. Uncertainty has a linear function of age in the denominator, with different coefficients across firms. To keep the same definition of

AGE for all firms, we plug in the median values of ϕ and $\sigma_{\rho,2}$ and $\hat{\sigma}_0 = 0.15$, which makes uncertainty proportional to $1/(1+\text{age})$. In an earlier version of the paper, the AGE variable was defined as the natural logarithm of age, which also captures the idea that one year of age should matter more for young firms than for old firms. The results for log(age) are very similar to those reported here. Note that the results are slightly weaker but still significant even when AGE is defined as plain age.

The choice of the other right-hand-side variables is based on equation (25) as well as on the existing literature on the determinants of M/B. Recall that the model implies that M/B increases with the current as well as estimated average future profitability, decreases with expected stock return, increases with the volatility of profitability, and increases with the uncertainty about average profitability. Moreover, these relations are all linear, at least for firms that pay no dividends. Future stock returns (RET) and profitability (ROE) are included on the right-hand side as proxies for expected returns and expected profitability, up to 25 years into the future ($q = 0, \ldots, 25$). When expectations are rational, they should be captured reasonably well by the ex post realized values.[11] These future values are included in the same linear fashion as in Vuolteenaho (2000), who derives an approximate linear identity that equates the log M/B ratio with an infinite discounted sum of future ROE and RET. Conditioning on the presence of future accounting and investment returns obviously leads to a sample populated only by firms that survived throughout the period. While such conditioning leads to survivorship biases in some applications, it is not clear why any bias at all should be present in the slope coefficients that are at the focus of our analysis.

In addition to AGE and future ROE and RET, we also include other regressors that could be related to M/B. Equation (25) asks for current profitability, ROE(0), and for the volatility of profitability, VOLP, which is estimated for a given stock as the residual variance from an AR(1) model, as described in Section 2.1. Motivated by Figure 1, we also add the dividend dummy, DD, which is equal to one if the firm paid any dividends in the current year and zero otherwise. The natural logarithm of total assets (SIZE) and the ratio of long-term debt to total assets (LEV) are included as measures of firm size and leverage.

Table 2 reports the estimated coefficients from regression (26), together with their t-statistics. The inference is conducted in the style of Fama and MacBeth (1973). The reported slope coefficients are time-series averages of the estimated cross-sectional slope coefficients.

[11] We also considered regression specifications that include only the averages of the future values of ROE and RET in place of those values per se. All M/B results (Tables 2 and 3) are very similar in terms of magnitude as well as significance to the results reported here, only the R^2s are lower. Including the future values enables us to assess how many future values need to be considered, as explained below.

The corresponding standard errors are also computed from the time series, and therefore include estimation error due to the residual correlation in the cross-sectional regressions. In those specifications in which there is any significant serial correlation in the time series, the standard errors are adjusted accordingly.[12] The results are reported for the specifications with $q = 0, 1, 5, 10, 15, 20,$ and 25 future values (i.e. "leads") of ROE and RET, as well as for a simple regression of M/B on AGE only.

Table 2 confirms the prediction of Corollary 3 that younger firms have higher M/B ratios, other things equal. The coefficient on AGE is significantly negative in all specifications. When AGE is the only variable included, the AGE coefficient is -1.15 ($t = -9.36$), confirming the pattern observed in Figure 5. The magnitude of the coefficient decreases to -0.94 ($t = -6.19$) when the current values of DD, LEV, SIZE, VOLP, and ROE are added, and it continues to decline as the leads of ROE and RET are included. However, the magnitude of the coefficient never falls below -0.32 ($t = -4.71$) observed for $q = 20$. Taking this conservative value, $\log(M/B)$ of a firm whose $1/(1 + \text{age})$ is lower by one is likely to be higher by 0.32. To interpret this magnitude, consider a typical two-year-old firm, whose M/B equal to 1.80 (Table 1). The regression estimate predicts that an otherwise identical one-year-old firm will have a M/B ratio of 1.90. A one-year difference in age thus translates into a valuation difference of more than 5.5%. Although the effect is weaker for older firms, it is safe to conclude that the effect of AGE on M/B is not only statistically but also economically significant, even after accounting for other known determinants of M/B.

Firms whose profitability is more volatile tend to have higher M/B, confirming another prediction of the model. The coefficient on VOLP is significantly positive throughout, ranging from 2.01 ($t = 6.01$) for $q = 0$ to 1.11 ($t = 4.82$) for $q = 25$.[13] Also, M/B ratios tend to be higher for smaller firms, firms that are less levered, and firms that pay no dividends.

All coefficients on ROE, current and future, are positive, consistent with the model. The coefficients generally decline as we go further into the future. In the most comprehensive specification ($q = 25$), all coefficients are positive, but none of the coefficients more than 19 years ahead is significant. Similarly, all RET coefficients are negative, consistent with the

[12] Different ways of adjusting the standard errors for potential serial correlation in the time series lead to exactly the same conclusions. For example, not adjusting at all or adjusting for any (even insignificant) serial correlation produce very similar results and do not affect any inferences.

[13] This result is somewhat sensitive to the exclusion of outliers. We exclude the values of VOLP above the 99th percentile of the cross-sectional distribution, thereby excluding stocks whose residual standard deviations of profitability exceed 100% per year. When the outliers in VOLP are included, two of the seven coefficients on VOLP lose their significance due to the extra noise. However, including these outliers has only a negligible effect on the coefficients on all other variables and their t-statistics.

model, and the magnitude of the coefficients also drops as we go further into the future. The most distant significantly negative coefficient is again 19 years ahead, and the values turn insignificantly positive at lead 24. This analysis reveals that the accounting and investment returns more than 20 years ahead have little if any effect on the current M/B, helping justify our choice of specifications with up to 25 leads of ROE and RET in Tables 2 and 3.

It is important to distinguish the age effects from the effects due to the variation in M/B over time (time effects) and the effects due to the differences between firms born in different years (cohort effects). The time effects are controlled for, since we run cross-sectional regressions year by year. Whether the M/B ratios are generally high or low across stocks in a given year, the cross-sectional results are unaffected. As for the cohort effects, we can show that our results are not driven by firms born late or early in the sample. The AGE coefficients tend to be negative for all years, with only occasional exceptions (typically around 1980, if any). Since the coefficients are negative throughout the 1960s, firms born after the 1960s are not crucial for our results. Moreover, since the AGE coefficients are negative even when firms born before 1963 are excluded from the regressions (numbers not reported to save space), those firms are not crucial for the results either. The data appear to reveal true age effects on M/B.

Table 3 tests another implication of the model, namely that the effect of AGE on M/B should be stronger for firms that pay no dividends (Corollary 4). This test is conducted by including an additional term, the cross-product of AGE and the dividend dummy DD, on the right-hand side of regression (26). The table reports the AGE coefficients for dividend payers and non-payers, together with their differences and t-statistics. The AGE coefficient is significantly more negative for non-payers than for payers, as the model predicts. The effect of AGE on M/B for payers is negative and significant with $q > 15$, but this effect is much stronger for non-payers – the AGE coefficients for non-payers are all highly significant, ranging from -2.14 ($t = -6.92$) for $q = 0$ to -0.71 ($t = -3.46$) for $q = 25$. To interpret the magnitudes, note that a typical non-payer of age two has M/B of 2.23. Taking the conservative value of -0.71 for the AGE coefficient, an otherwise identical firm that is just one year younger is estimated to have M/B of 2.51. A one-year difference in age thus leads to a valuation difference of over 12.5%. An analogous exercise for a typical five-year-old firm shows that the valuation of an otherwise identical one-year-old firm is higher by almost 27%. The relation between firm valuation and uncertainty about future profitability, proxied by age, is remarkably strong.

Share repurchases have become an increasingly important alternative to dividends over the past few decades. As a robustness check, we redid the analysis defining dividend payers more broadly as firms that paid any dividends or repurchased any common shares in the given year. The results are very similar to those reported in Table 3, despite the shorter sample period (repurchase data can be constructed only back to 1971, as described in the appendix). The AGE coefficients for the non-payers as well as the differences between the coefficients for the payers and non-payers remain significant in all specifications.

We also analyze the changes in M/B over time for a given stock. Calculating the grand mean of such changes across firms and years, we find that M/B of a typical firm declines substantially, by 0.095 per year on average. In addition, our model predicts that the decline in M/B should be steeper for younger firms. This prediction is strongly endorsed by the regression of changes in log(M/B) on the same regressors as in (26): AGE, DD, LEV, SIZE, VOLP, ROE, and future values of ROE and RET. (Those future values that are mechanically correlated with future changes in M/B are excluded.) Changes over one and three years are both considered. The AGE coefficients in those regressions are significantly positive, which means that changes in M/B are more negative for younger firms, as the model predicts. A second-order prediction is that the learning effects are stronger for dividend non-payers with less volatile profitability. While the coefficients on DD and VOLP are insignificant, both generally go in the right direction. Overall, this additional analysis, whose results are not reported to save space, provides further support for the model.

3.2. Results for Return Volatility

For each stock, idiosyncratic return volatility in any given year is estimated from the market model regression of (at least ten) monthly stock returns on the returns of the value-weighted portfolio of all NYSE, AMEX, and NASDAQ-traded stocks. All results in this subsection seem quite robust to the definition of idiosyncratic volatility as they hold even when this volatility is replaced by total volatility of stock returns.

The bottom panel of Figure 5 plots the median idiosyncratic return volatility for firms of different ages. Younger firms tend to have higher volatility: the median declines almost monotonically from over 11% per month for the newly-listed firms to about 8% for firms that are 20 years old. The bottom panel of Figure 6 plots the median volatility in every year between 1963 and 2000. Without exception, young firms again have higher median volatility than old firms. In addition, dividend non-payers have substantially higher volatility than payers in both figures. Both facts are consistent with the model.

The same observations come to light in the regression analysis that controls for other potential determinants of volatility. Each year between 1963 and 2000, idiosyncratic variance of stock returns is regressed cross-sectionally on various stock characteristics: AGE, the log of M/B, DD, LEV, SIZE, VOLP, and ROE. The inference is conducted as in Table 2, and the results are summarized in Table 4. Regardless of which subset of variables is included in the regression, the AGE coefficient is significantly negative, suggesting that younger firms have higher return volatility. While AGE is our favorite proxy for uncertainty about profitability, M/B could act as a proxy as well, since it increases with that uncertainty. Therefore, the result that firms with higher M/B ratios tend to have higher return volatility nicely complements the evidence based on AGE. We also find that firms that pay no dividends have more volatile returns, with t-statistics of around -13. (This relation remains highly significant when DD is redefined to reflect also share repurchases.) More volatile profitability translates into more volatile returns, as the coefficient on VOLP is significantly positive ($t = 5.91$ or higher). All of these results provide a strong endorsement to our learning model.

Naturally, some firms depart from our dataset. A negative relation between return volatility and AGE can be expected if firms with high volatility depart more often. To investigate whether firms with certain characteristics are more likely to depart than others, we implement a simple linear probability model and run a pooled time-series cross-sectional regression of a dummy variable for a firm's departure from the dataset on a number of characteristics. Departure is defined here as the beginning of the last sequence of the firm's missing volatility values, if any. The characteristics include M/B, DD, LEV, SIZE, VOLP, and stock return volatility. The results (not reported to save space) reveal that firms with higher return volatility are significantly more likely to disappear. Of course, this result does not invalidate our model. The disappearing firms are likely to have high volatility precisely because their future profitability is highly uncertain, whether due to their relatively young age or due to reasons specific to their disappearance. Firms with high M/B ratios are in fact less likely to disappear, and the other characteristics enter insignificantly.

Similar to M/B, we also analyze the changes in volatility for a given stock. The grand mean of such changes across firms and years is positive (at 0.0180 per year), not negative as the model predicts. This result can be traced to a surprising fact, discussed in the conclusion, that firm profitability has become substantially more volatile over the last few decades. As profitability grows more volatile, return volatility grows despite the learning effect in the opposite direction. We also regress changes in volatility three years ahead on the same characteristics as in Table 4. The model predicts that volatility changes should be lower (more negative) when there is more uncertainty about average profitability or when learning is faster. Volatility changes indeed tend to be significantly lower for firms with higher M/B, which are likely to face bigger uncertainty. The changes are also significantly lower for firms with lower volatility of profitability and for dividend non-payers, for which learning should be faster, as discussed in Section 2. At least to this extent, the results on volatility changes, not reported to save space, support the model.

4. Conclusions

This paper develops a framework for valuing stocks whose average future profitability is unknown. We show that uncertainty about a firm's average profitability increases the firm's M/B ratio as well as its idiosyncratic return volatility. This uncertainty is especially large for the newly listed firms, but it declines over time due to learning. Our model therefore predicts that both M/B and the return volatility of a typical young firm should decline as the firm ages. We find that younger firms indeed tend to have higher M/B ratios than older firms, after controlling for the known determinants of M/B such as future profitability and returns. Moreover, this effect is stronger for firms that pay no dividends, confirming another prediction of the model. The model is also endorsed by the observation that M/B declines faster for younger firms. Finally, we show that return volatility tends to be higher for younger firms, for firms with more volatile profitability, and for firms that pay no dividends, consistent with the model.

Due to the idiosyncratic nature of learning in our model, expected stock returns are unaffected by uncertainty about average profitability. As a result, the model cannot explain any cross-sectional relation between M/B and expected stock returns. For the same reason, we make no contribution to the literature on the potential long-run underperformance of initial public offerings.

This paper does not explore the equilibrium implications of learning about firm profitability. An exogenous specification of the stochastic discount factor seems reasonable given our focus on the valuation of young firms, which typically do not account for a sizable fraction of the total market capitalization. Nonetheless, endogenizing the stochastic discount factor in a general equilibrium framework with learning would be a useful and ambitious direction for future research.

Future work can also model the firm creation process. Since uncertainty increases valuations, entrepreneurs have an incentive to start a new firm with uncertain profitability even when expected profitability $\overline{\rho}$ is low. In addition, private owners are tempted to take their firms public when current profitability is unusually high. Investors should recognize this and adjust their priors accordingly. One might therefore expect some new firms with high prior uncertainty to have low prior means of $\overline{\rho}$, and the overall effect on valuation is unclear. What is clear is that uncertainty about $\overline{\rho}$ increases valuations holding $\overline{\rho}$ constant, as well as that investors learn about $\overline{\rho}$ over time regardless of the prior. The model's main predictions therefore easily survive the above arguments. Moreover, the empirical challenge is to explain why the valuations of young firms tend to be high, not low. If $\overline{\rho}$'s of young firms are indeed low as argued above, then the high valuations of young firms are even more puzzling and the uncertainty effect must be even stronger to justify them.

The empirical analysis uses age as a proxy for uncertainty about average profitability. In our model, uncertainty declines over time due to learning, making age a natural choice. One alternative proxy could be the dispersion of security analysts' profitability forecasts, constructed from the earnings forecasts contained in the IBES database of Thomson Financial. However, we are concerned that using the IBES sample would likely introduce an important selection bias, as that sample is heavily tilted toward big and well-established stocks.[14] Since stocks for which future profitability is likely to be the most uncertain are largely absent from the IBES database, we would not necessarily expect to find strong learning effects there. Nevertheless, some alternative proxies should be investigated in future work.

Campbell, Malkiel, Lettau, and Xu (2001) show that average idiosyncratic volatility of individual stock returns has increased since the 1960s. Indeed, the top panel of Figure 7 shows that average idiosyncratic volatility in our sample rises from 6.7% per month in 1963 to 17.5% in 2000. In our model, idiosyncratic volatility of returns has two parts, idiosyncratic volatility of profitability and uncertainty about profitability (equation 24). Both parts seem to have contributed to the increase in average idiosyncratic return volatility. Average uncertainty about profitability has risen due to the recent explosion in the number of newly listed firms, depicted in the third panel of Figure 7 (Fama and French, 2001b). Interestingly, average volatility of firm profitability has risen as well, from 10% per year in 1963 to over 40% in 2000, as shown in the second panel.[15] This striking increase might perhaps be due to falling

[14] Diether, Malloy, and Scherbina (2002) show that over 95% of stocks in the first size decile (using NYSE breakpoints) are not covered by the IBES database in January 1983, while this fraction is below 8% for stocks in deciles 9 and 10. Similarly, Hong, Lim, and Stein (2000) show that 82% of stocks in the bottom size quintile (using NYSE/AMEX breakpoints) are not covered by the IBES database in 1988, while this fraction is below 6% for stocks in the top quintile.

[15] The figure plots the cross-sectional standard deviation of ROE. Using variance decomposition, the cross-sectional variance of ROE is the sum of the cross-sectional average of the variances of ROE and the cross-sectional variance of expected ROE. Since the latter component is relatively small, we take the cross-sectional

barriers to entry throughout the economy – when new firms can enter the marketplace easily, the profits of the incumbent firms are likely to become more volatile. Finally, the fraction of firms that pay no dividends has increased dramatically (Fama and French, 2001a). The bottom panel of Figure 7 shows that while only 27% of all firms paid no dividends in 1963, this fraction rises to 68% in 2000. Similarly, the fraction of firms that neither pay dividends nor repurchase shares grows from 31% in 1971 (when repurchase data becomes available) to 52% in 2000. According to Corollary 6, this rise should also raise return volatility. The common increase in the four series plotted in Figure 7 hardly seems coincidental, given the link between them within our model. This evidence has the potential to explain the baffling result of Campbell et al. (2001). We hope that future research will provide more conclusive evidence by carefully decomposing the rise in average return volatility into the effects described here.

variance of ROE as an estimate of the average variance of ROE. Also note that we discard outliers, defined as ROE smaller than -200% and bigger than 1000% per year. The rise in volatility observed in Figure 7 is robust to different cutoffs that identify the outliers, and a similar (albeit more jagged) figure obtains even when the outliers are included.

Table 1
Summary Statistics

The table summarizes various statistics for groups of firms of the same age, where age measures the number of years since the firm's listing. Panel A reports the medians across firms of the annual characteristics listed in the row label. Residual return volatility is calculated with respect to the market model. Return volatility, return on equity, stock return, and leverage are all expressed in percentage terms. Assets are in millions of dollars. Panel B shows the number of firms with valid M/B ratios, the fraction of firms that pay dividends, and the standard deviation of ROE across firms. The latter two values are expressed in percent.

Age	1	2	3	4	5	6	7	8	9	10
Panel A. Medians across firms.										
M/B	2.25	1.80	1.57	1.49	1.39	1.38	1.35	1.33	1.27	1.25
Total return volatility	13.96	13.35	13.09	13.05	12.72	12.38	12.23	11.80	11.77	11.19
Resid return volatility	11.11	11.16	11.00	10.94	10.64	10.31	10.02	9.61	9.69	9.26
ROE	12.93	11.06	10.37	10.31	10.73	10.95	11.66	11.53	11.51	11.56
Stock return	-2.70	-7.19	-1.09	4.86	2.82	7.04	7.91	6.56	4.33	9.22
Assets	29.12	37.21	43.16	48.33	53.85	58.62	65.11	71.46	77.07	83.00
Leverage	7.14	9.43	11.62	12.46	13.17	13.96	13.95	13.96	13.63	13.98
Panel B. Other statistics.										
Number of firms	14522	14543	12806	11319	9942	8699	7746	6985	6218	5606
Dividend payers	27.87	31.16	34.16	36.59	39.69	41.82	44.14	46.60	48.47	50.69
Std deviation of ROE	29.38	41.72	33.56	38.83	30.99	82.68	32.82	35.74	27.20	32.74

Table 2
Determinants of Market-to-Book Ratios

Each year between 1963 and 2000, the log of the market-to-book ratio (M/B) is regressed cross-sectionally on minus the reciprocal of one plus firm age (AGE), dividend dummy (DD), leverage (LEV), the log of total assets (SIZE), the volatility of profitability (VOLP), current return on equity (ROE), and future values of ROE and stock returns (RET), up to the number of leads listed in the column headings. The reported slope coefficients and their standard errors are computed from the time-series of the estimated cross-sectional slope coefficients. The t-statistics, adjusted for any significant serial correlation in the time series, are in parentheses. The numbers of years across which the averages of the coefficients are computed are given in the last row. Also given are averages across these years of the R^2's and of the numbers of firms from the cross-sectional regressions. The values in the first column are obtained from the regression of log M/B on AGE only.

		Number of future ROE and RET included						
Regressor		0	1	5	10	15	20	25
Intercept	0.21	-0.04	-0.04	0.05	0.13	0.17	0.12	0.19
	(2.35)	(-0.45)	(-0.44)	(0.48)	(1.21)	(1.68)	(1.37)	(1.36)
AGE	-1.15	-0.94	-0.93	-0.63	-0.44	-0.37	-0.32	-0.33
	(-9.36)	(-6.19)	(-6.04)	(-5.63)	(-6.10)	(-4.07)	(-4.71)	(-4.71)
DD		-0.01	-0.04	-0.08	-0.13	-0.16	-0.17	-0.18
		(-0.29)	(-1.14)	(-3.47)	(-5.30)	(-7.43)	(-4.97)	(-4.14)
LEV		-0.38	-0.36	-0.34	-0.32	-0.33	-0.28	-0.23
		(-5.14)	(-6.24)	(-6.23)	(-4.78)	(-3.17)	(-2.65)	(-2.06)
SIZE		0.02	0.01	-0.01	-0.03	-0.03	-0.02	-0.02
		(1.51)	(1.03)	(-1.53)	(-3.78)	(-4.76)	(-3.21)	(-1.60)
VOLP		2.01	1.96	1.72	1.42	1.53	1.74	1.11
		(6.01)	(5.78)	(5.61)	(5.51)	(4.29)	(5.93)	(4.82)
ROE		1.59	0.99	0.82	0.81	0.76	0.98	1.28
		(5.80)	(6.24)	(6.79)	(7.23)	(7.99)	(6.66)	(8.38)
ROE(1)			1.32	0.97	0.97	0.98	1.05	1.35
			(9.10)	(9.14)	(8.79)	(7.81)	(8.24)	(5.88)
ROE(2)				0.68	0.60	0.60	0.56	0.62
				(7.85)	(5.69)	(4.99)	(5.58)	(4.14)
ROE(3)				0.48	0.47	0.57	0.50	0.48
				(6.92)	(5.84)	(5.44)	(3.63)	(2.42)
RET(1)			-0.22	-0.35	-0.39	-0.48	-0.50	-0.55
			(-7.48)	(-10.79)	(-11.49)	(-10.20)	(-9.03)	(-7.36)
RET(2)				-0.27	-0.32	-0.41	-0.43	-0.43
				(-9.80)	(-10.91)	(-10.54)	(-8.89)	(-7.04)
RET(3)				-0.21	-0.29	-0.38	-0.40	-0.42
				(-7.29)	(-8.61)	(-8.99)	(-7.83)	(-6.37)
Average R^2	0.05	0.21	0.30	0.42	0.51	0.61	0.69	0.78
Average N	4234	2318	2145	1840	1393	956	662	437
Years	38	38	37	33	28	23	18	13

Table 3
The AGE coefficients for Dividend Payers vs. Non-Payers

Each year between 1963 and 2000, the log of the market-to-book ratio (M/B) is regressed cross-sectionally on minus the reciprocal of one plus firm age (AGE), dividend dummy (DD), the interaction term AGE*DD, leverage (LEV), the log of total assets (SIZE), the volatility of profitability (VOLP), current return on equity (ROE), and future values of ROE and stock returns (RET), up to the number of leads listed in the column headings. The reported AGE coefficients and their standard errors are computed from the time-series of the estimated cross-sectional slope coefficients. The t-statistics, adjusted for any significant serial correlation in the time series, are in parentheses. The numbers of years across which the averages of the coefficients are computed are given in the last row. Also given are the averages across these years of the R^2's from the cross-sectional regressions, as well as of the numbers of dividend payers and non-payers in each year. The values in the first column are obtained from the regression of log M/B on AGE and AGE*DD only. To obtain the t-statistics on the coefficients for dividend payers, the regression is rerun with DD redefined as its own complement.

		Number of future ROE and RET included						
Regressor		0	1	5	10	15	20	25
Non-payers	-1.39	-2.14	-1.95	-1.50	-1.13	-0.99	-0.89	-0.71
	(-8.88)	(-6.92)	(-7.58)	(-8.02)	(-10.76)	(-5.54)	(-4.77)	(-3.46)
Payers	-0.58	-0.19	-0.19	-0.05	-0.04	-0.11	-0.17	-0.26
	(-6.42)	(-1.51)	(-1.87)	(-0.54)	(-0.51)	(-1.66)	(-2.42)	(-2.95)
Difference	0.81	1.95	1.76	1.46	1.09	0.88	0.71	0.45
	(5.32)	(6.32)	(6.58)	(7.49)	(9.94)	(4.93)	(3.82)	(1.90)
Average R^2	0.06	0.22	0.30	0.43	0.51	0.61	0.69	0.78
Avg N: Non-payers	1967	750	702	562	369	190	112	72
Avg N: Payers	2181	1567	1444	1278	1024	767	550	365
Years	38	38	37	33	28	23	18	13

Table 4
Determinants of Return Volatility

Each year between 1963 and 2000, residual return variance from the market model is regressed cross-sectionally on various subsets of the following set of variables: minus the reciprocal of one plus firm age (AGE), the log of M/B (M/B), dividend dummy (DD), leverage (LEV), the log of total assets (SIZE), the volatility of profitability (VOLP), and current return on equity (ROE). The reported slope coefficients and their standard errors are computed from the time-series of the estimated cross-sectional slope coefficients. The t-statistics, adjusted for any significant serial correlation in the time series, are in parentheses. Also given are averages across these years of the R^2's and of the numbers of firms from the cross-sectional regressions. All reported coefficients are multiplied by 100 for convenience.

Intercept	1.29	1.43	0.94	0.61	0.50	1.68	2.65	2.75	2.61
	(5.77)	(8.17)	(7.43)	(9.33)	(9.77)	(15.37)	(12.47)	(11.57)	(11.62)
AGE	-3.65		-3.56	-5.35	-6.43	-3.30	-1.56		-1.23
	(-5.87)		(-6.73)	(-2.24)	(-2.65)	(-2.49)	(-2.19)		(-2.64)
M/B		0.45	0.34		-0.00	0.05		0.18	0.17
		(3.59)	(2.94)		(-0.10)	(1.22)		(2.64)	(2.55)
DD						-1.22	-0.85	-0.83	-0.82
						(-12.56)	(-13.03)	(-13.25)	(-12.90)
LEV						-0.43	-0.17	-0.07	-0.05
						(-2.69)	(-0.95)	(-0.77)	(-0.61)
SIZE							-0.19	-0.20	-0.20
							(-13.23)	(-10.08)	(-9.59)
VOLP				15.19	12.54	7.47	6.23	5.96	5.96
				(5.91)	(7.22)	(6.51)	(6.35)	(6.66)	(6.66)
ROE							-0.75	-0.91	-0.93
							(-3.55)	(-4.79)	(-4.92)
Average R^2	0.02	0.01	0.04	0.06	0.08	0.15	0.18	0.18	0.18
Average N	4842	4073	4073	2533	2423	2392	2341	2311	2311

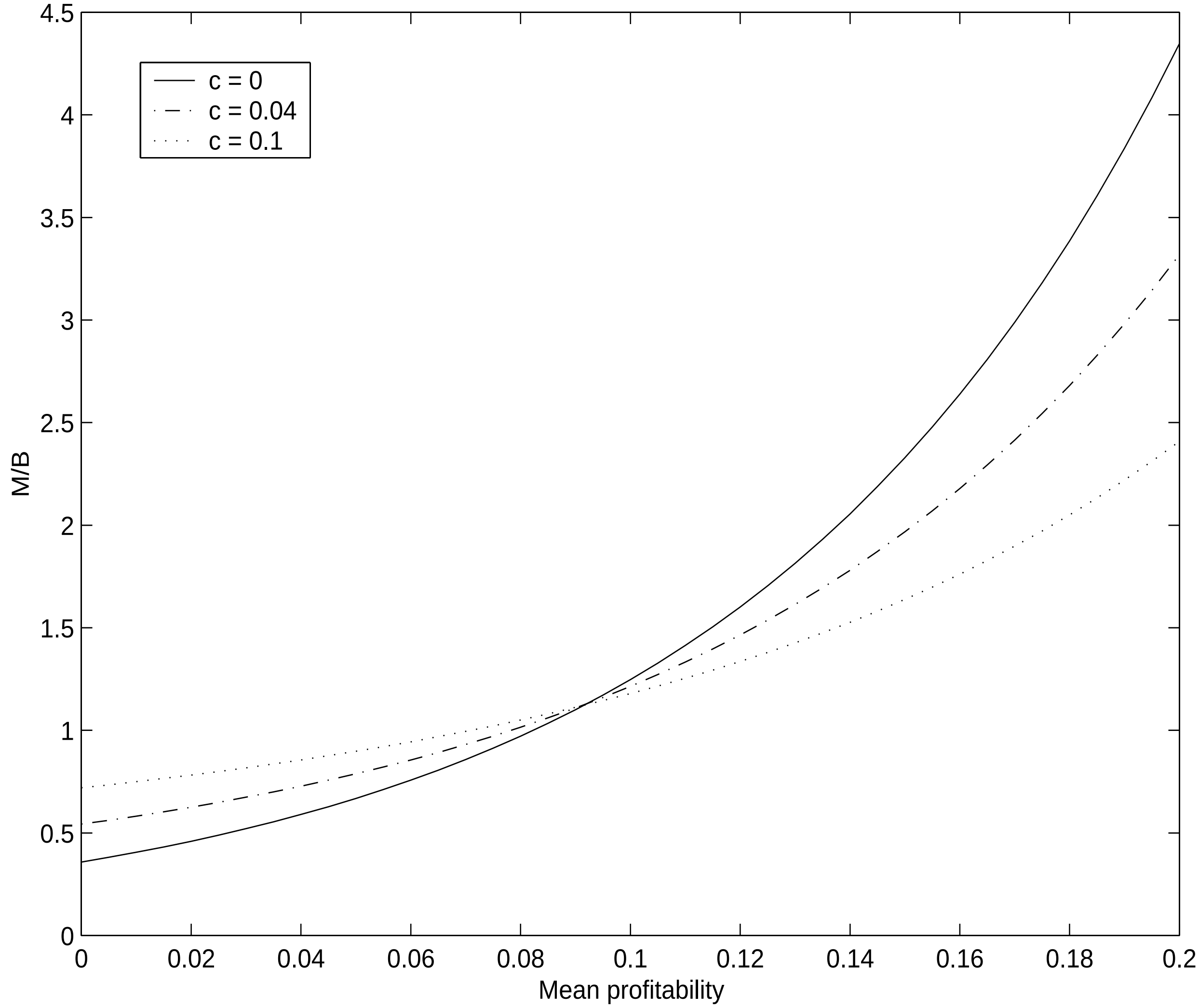

Figure 1. M/B versus mean profitability.

The figure plots the model-implied M/B ratio against the known value of mean profitability $\overline{\rho}$ for various levels of the dividend yield c. The model parameters are specified as follows: $\phi = 0.3963$, $\rho_t = 0.11$, $\sigma_{\rho,1} = 0.0584$, $\sigma_{\rho,2} = 0.0596$, $\sigma_{\pi,1} = 0.6$, $r = 0.03$, and $\tau = 15$.

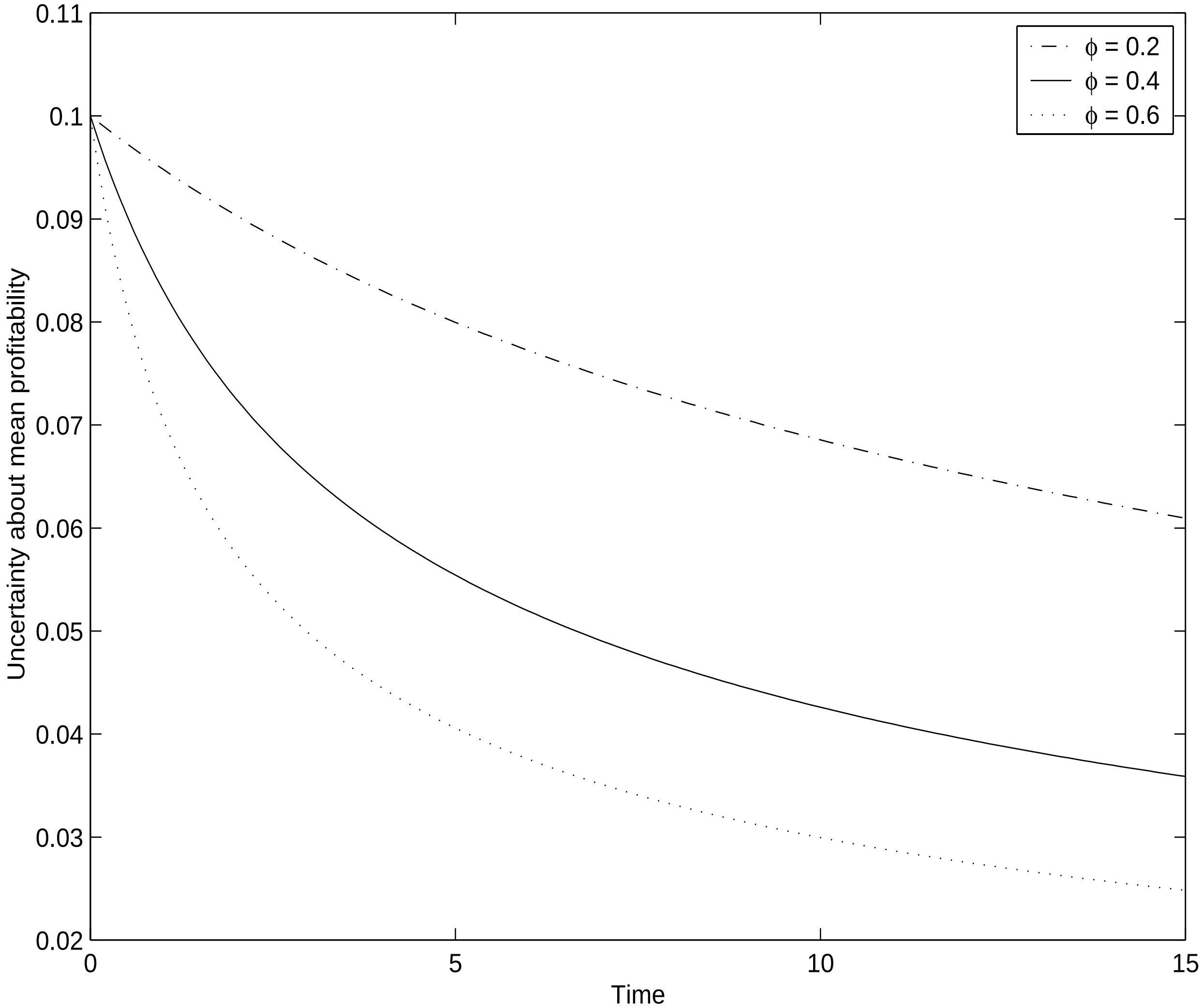

Figure 2. Uncertainty about mean profitability over time.

The figure plots the evolution over time of $\widehat{\sigma}_t$, standard deviation of the posterior distribution of mean profitability $\overline{\rho}$. Three values are considered for the parameter ϕ, which governs the speed of mean reversion in profitability. The prior standard deviation is $\widehat{\sigma}_0 = 0.10$, $c = 0.0434$, and all other parameter values are as in Figure 1.

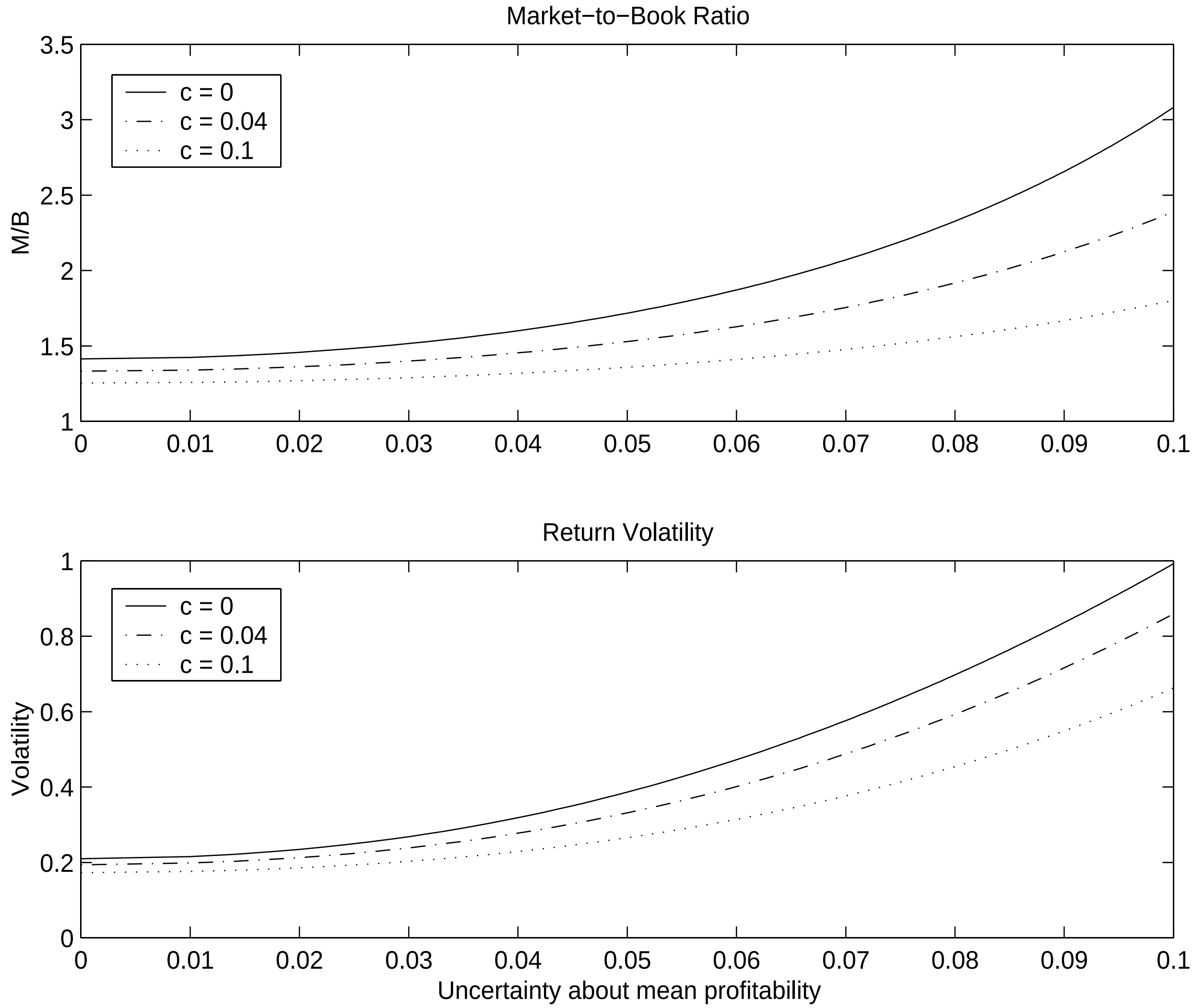

Figure 3. M/B and stock return volatility versus uncertainty about mean profitability. In the top panel, the figure plots the model-implied M/B ratio against $\widehat{\sigma}_t$, standard deviation of the posterior distribution of mean profitability $\overline{\rho}$. In the bottom panel, the figure plots the model-implied return volatility against $\widehat{\sigma}_t$. Three values are considered for the dividend yield c. All other parameter values are as in Figure 1.

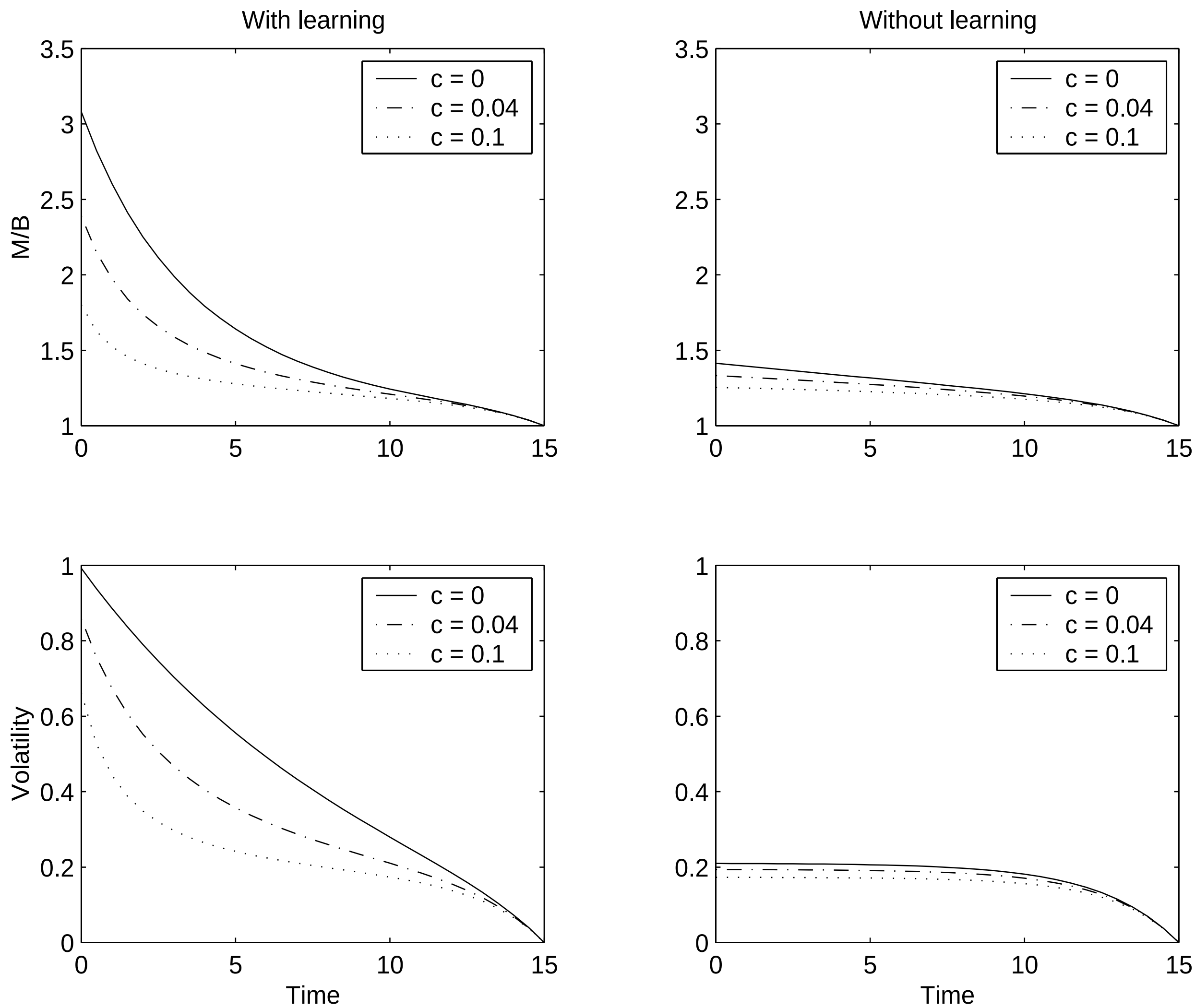

Figure 4. M/B and stock return volatility over time.

The figure plots the evolution over time of the model-implied M/B ratio (top two panels) and the model-implied return volatility (bottom two panels). In the left-hand panels, mean profitability $\overline{\rho}$ is treated as unknown. In the right-hand panels, $\overline{\rho}$ is treated as known. Three values are considered for the dividend yield c. The prior standard deviation is $\widehat{\sigma}_0 = 0.10$, and all other parameter values are as in Figure 1.

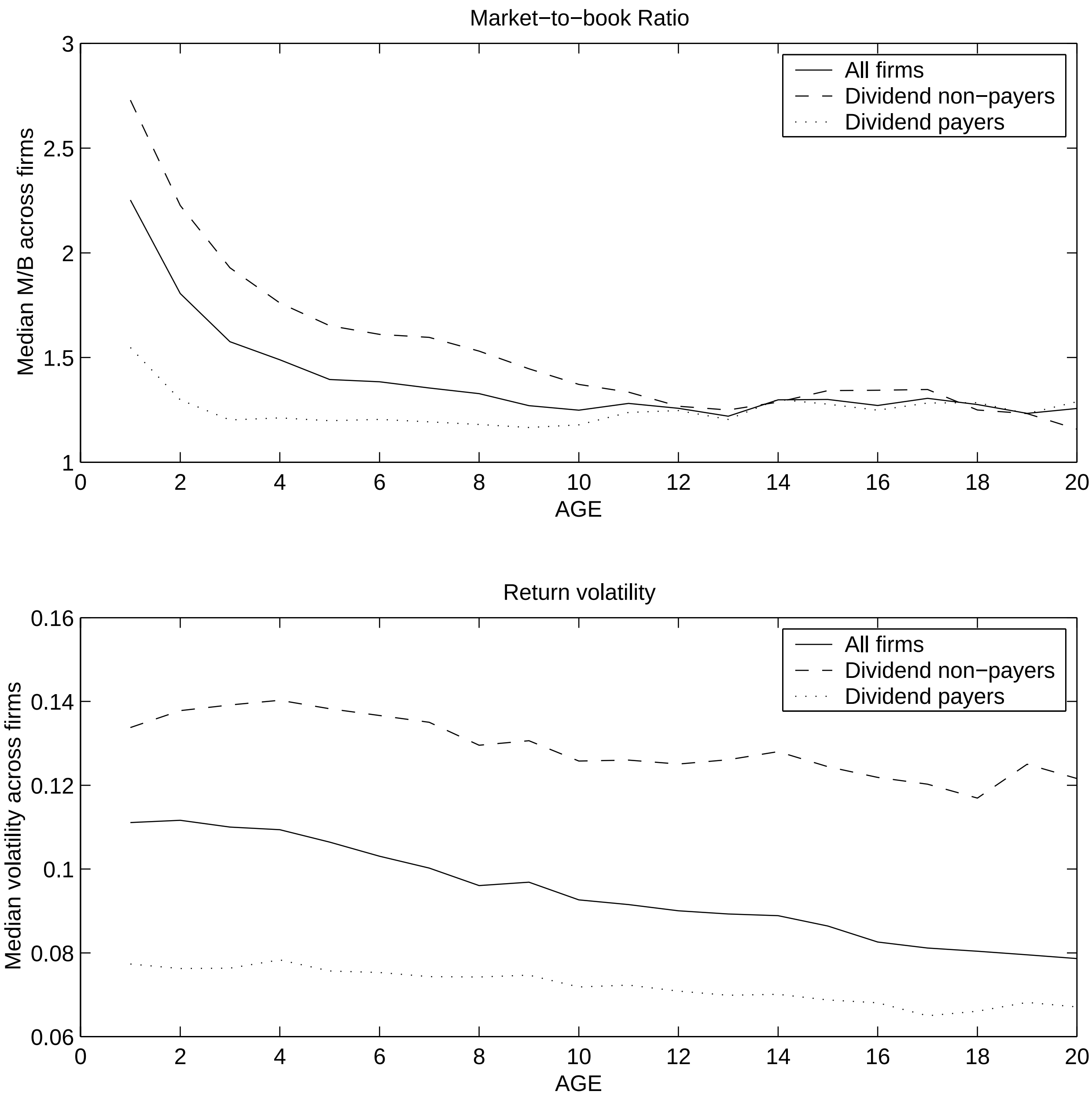

Figure 5. M/B ratio and return volatility in the years after listing.

For each age, the figure plots the median M/B ratio (top panel) and the median idiosyncratic return volatility (bottom panel) across firms of that age, regardless of the calendar year in which that age was reached. The solid line plots the medians across all firms, the dashed line plots the median across dividend non-payers, and the dotted line across dividend payers. Idiosyncratic volatility is estimated from the market model regression.

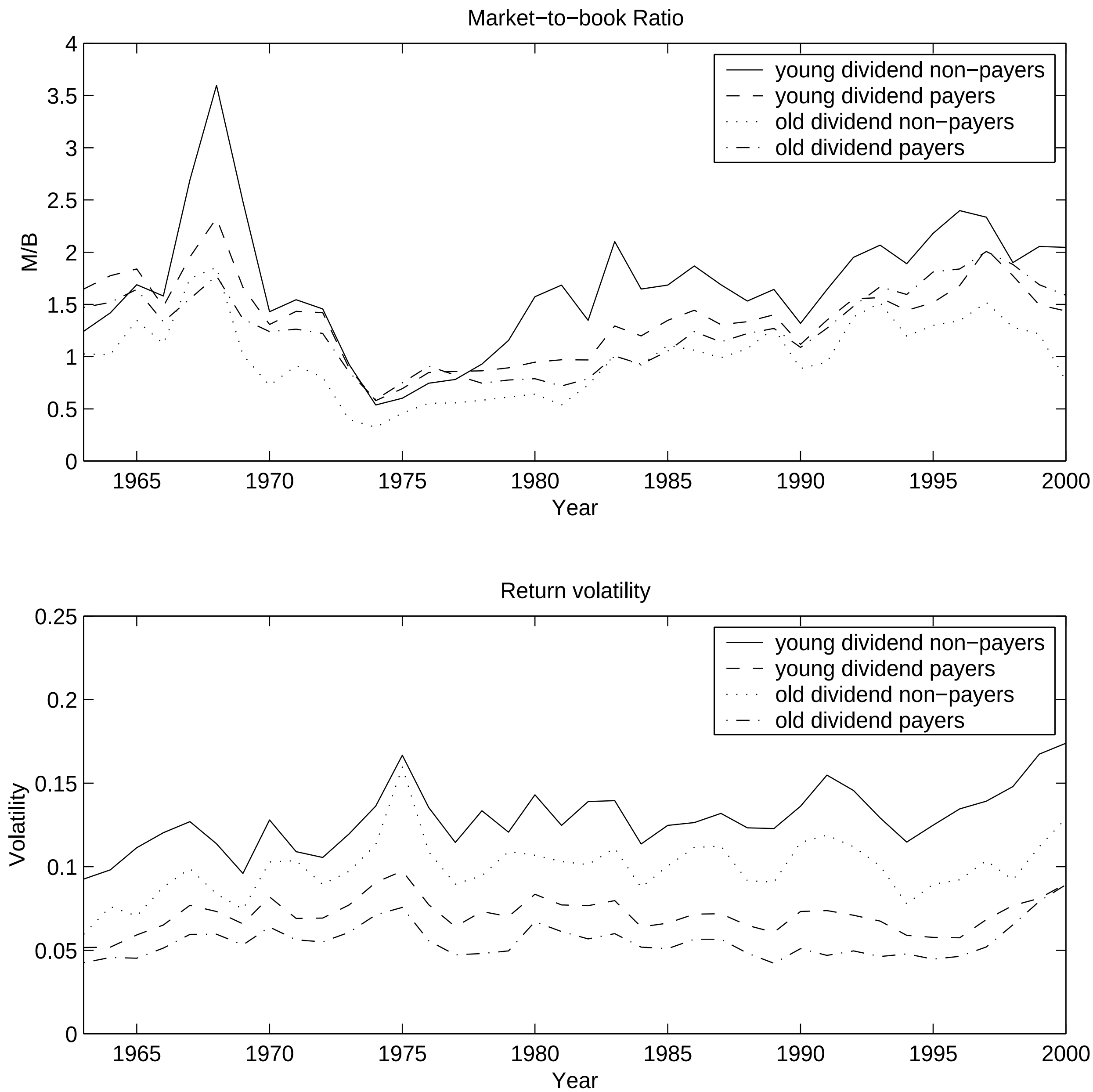

Figure 6. M/B ratios and return volatility in calendar time.

The figure plots the evolution of M/B and idiosyncratic return volatility in calendar time. Each year, all firms are categorized as young or old depending on whether their age exceeds the midpoint between the minimum and maximum age in the cross-section. Firms are also separately sorted into dividend payers and non-payers, depending on whether they paid any common stock dividends in the sorting year. Four groups of stocks are formed each year by intersecting the independent sorts on age and dividends. The top panel plots the median M/B ratios for these four groups, and the bottom panel plots the median return volatilities. Idiosyncratic volatility is estimated from the market model regression.

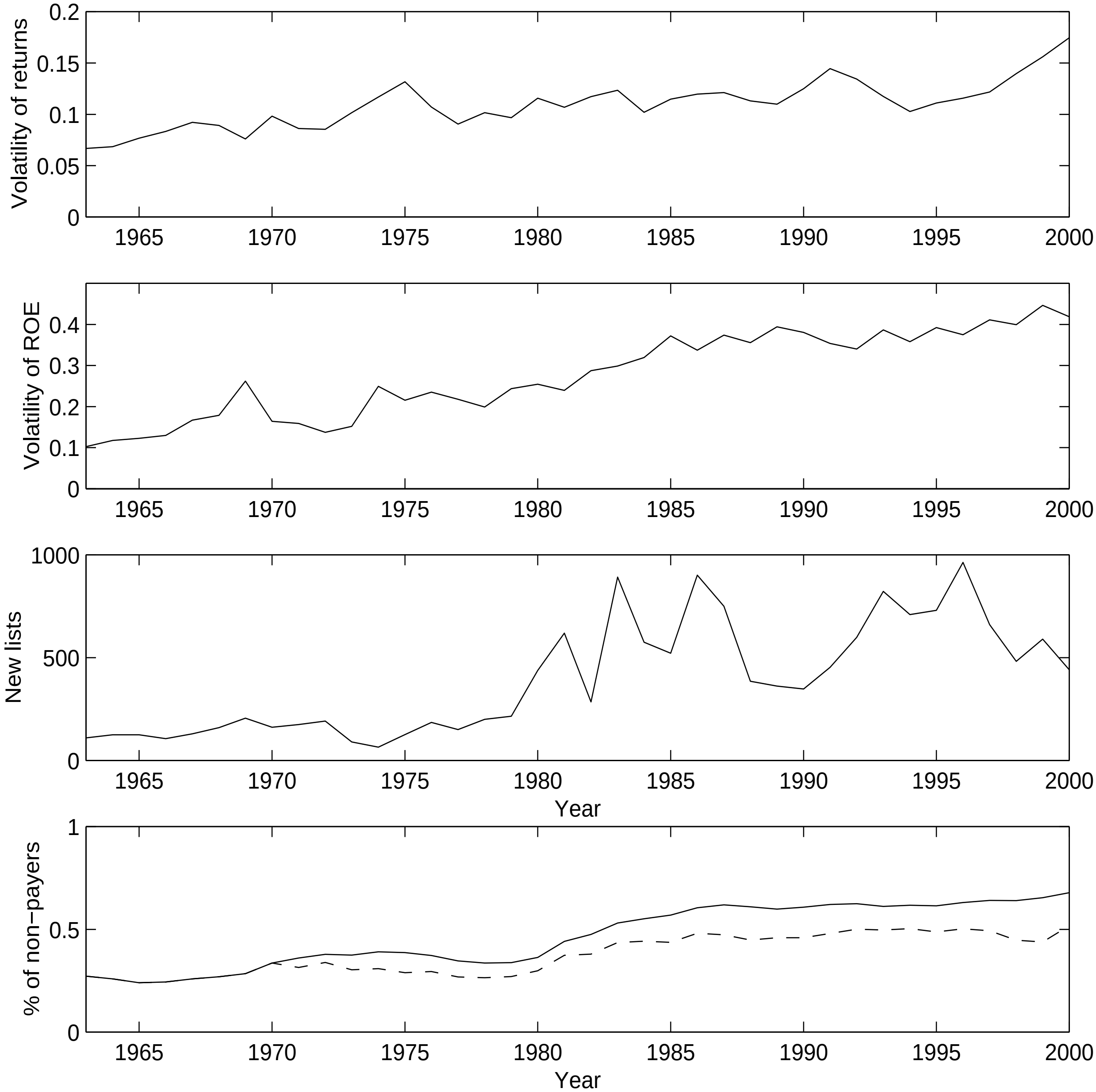

Figure 7. Return volatility, volatility of profitability, the number of new lists, and the fraction of dividend non-payers.

In the top panel, the figure plots the evolution of average idiosyncratic return volatility in calendar time. The evolution of the cross-sectional standard deviation of ROE is plotted in the second panel. The third panel plots the number of new lists. The bottom panel plots the fraction of all firms that pay no dividends (solid line) and the fraction of firms that neither pay dividends nor repurchase any shares in the current year.

References

Ang, Andrew, and Jun Liu, 2001, "A general affine earnings valuation model", *Review of Accounting Studies*, forthcoming.

Bakshi, Gurdip, and Zhiwu Chen, 2001, "Stock valuation in dynamic economies", working paper, University of Maryland.

Beaver, William H., 1970, "The time series behavior of earnings", *Journal of Accounting Research* (Supplement), 62–89.

Berk, Jonathan B., Richard C. Green, and Vasant Naik, 1999, "Optimal investment, growth options, and security returns", *Journal of Finance* 54, 1553–1607.

Bernardo, Antonio E., and Bhagwan Chowdry, 2002, "Resources, real options, and corporate strategy", *Journal of Financial Economics* 63, 211–234.

Brennan, Michael J. and Eduardo S. Schwartz, 1982, "Consistent regulatory policy under uncertainty", *Bell Journal of Economics* 13, 506–521.

Brennan, Michael J. and Yihong Xia, 2001, "Stock return volatility and equity premium", *Journal of Monetary Economics* 47, 249–283.

Campbell, John Y., Martin Lettau, Burton G. Malkiel, and Yexiao Xu, 2001, "Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk", *Journal of Finance* 56, 1–43.

Cohen, Randolph B., Christopher Polk, and Tuomo Vuolteenaho, 2001, "The value spread", NBER working paper No. 8242.

David, Alexander, 1997, "Fluctuating confidence in stock markets: Implications for returns and volatility", *Journal of Financial and Quantitative Analysis* 32, 427–462.

Diether, Karl B., Christopher J. Malloy, and Anna Scherbina, 2002, "Differences of opinion and the cross-section of stock returns", *Journal of Finance*, forthcoming.

Duffie, Darrell, 1996, *Dynamic Asset Pricing Theory*, Princeton University Press, Princeton, NJ.

Fama, Eugene F. and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 2000, "Forecasting profitability and earnings", *Journal of Business* 73, 161–175.

Fama, Eugene F., and Kenneth R. French, 2001a, "Disappearing dividends: Changing characteristics or lower propensity to pay?", *Journal of Financial Economics* 60, 3–43.

Fama, Eugene F., and Kenneth R. French, 2001b, "Newly listed firms: Fundamentals, survival rates, and returns", working paper, University of Chicago.

Fama, Eugene F., and J. D. MacBeth, 1973, "Risk, return, and equilibrium: Empirical tests", *Journal of Political Economy* 81, 607–636.

Feltham, Gerald A., and James A. Ohlson, 1995, "Valuation and clean surplus accounting for operating and financial activities", *Contemporary Accounting Research* 11, 689–731.

Feltham, Gerald A., and James A. Ohlson, 1999, "Residual earnings valuation with risk and stochastic interest rates", *The Accounting Review* 74, 165–183.

Freeman, Robert N., James A. Ohlson, and Stephen H. Penman, 1982, "Book rate-of-return and prediction of earnings changes: An empirical investigation", *Journal of Accounting Research* 20, 639–653.

Gomes, Joao, Leonid Kogan, and Lu Zhang, 2001, "Equilibrium cross-section of returns", working paper, University of Pennsylvania.

Grullon, Gustavo, and Roni Michaely, 2000, "Dividends, share repurchases, and the substitution hypothesis", working paper, Rice University and Cornell University.

Hong, Harrison, Terence Lim, and Jeremy C. Stein, 2000, "Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies", *Journal of Finance* 55, 265–295.

Jain, Bharat A., and Omesh Kini, 1994, "The post-issue operating performance of IPO firms", *Journal of Finance* 49, 1699–1726.

Jovanovich, Boyan, 1982, Selection and the evolution of industry, *Econometrica* 50, 649–670.

Kendall, M.G., 1954, "Note on bias in the estimation of autocorrelation", *Biometrika* 41, 403–404.

Lewellen, Jonathan, and Jay Shanken, 2002, "Learning, asset-pricing tests, and market efficiency", *Journal of Finance*, forthcoming.

Liptser, Robert S., and Albert N. Shiryayev, 1977, *Statistics of Random Processes: I, II*, Springer-Verlag, New York.

Lookabill, L., 1976, "Some additional evidence on the time series properties of accounting earnings", *Accounting Review* (October 1976), 724–738.

Marriott, F. H. C., and J. A. Pope, 1954, "Bias in the estimation of autocorrelations", *Biometrika* 41, 390–402.

Myers, Stewart C., 1984, "The capital structure puzzle", *Journal of Finance* 39, 575–592.

Myers, Stewart C., and Nicholas Majluf, 1984, "Corporate financing and investment decisions when firms have information the investors do not have", *Journal of Financial Economics* 13, 187–221.

Ohlson, James A., 1990, "A synthesis of security valuation theory and the role of dividends, cash flows, and earnings", *Contemporary Accounting Research* 6, 648–676.

Ohlson, James A., 1995, "Earnings, book values, and dividends in equity valuation", *Contemporary Accounting Research* 11, 661–687.

Pakoš, Michal, 2001, "Mean reversion in profitability measures", working paper, University of Chicago.

Penman, Stephen H., 1991, "An evaluation of accounting rate-of-return", *Journal of Accounting, Auditing, and Finance* 6, 233–256.

Rothschild, Michael, and Joseph E. Stiglitz, 1970, "Increasing risk: I. A definition", *Journal of Economic Theory* 2, 225-243.

Schultz, Paul, 2002, "Pseudo Market Timing and the Long-Run Underperformance of IPOs", working paper, University of Notre Dame.

Schwartz, Eduardo S., and Mark Moon, 2000, "Rational pricing of Internet companies", *Financial Analysts Journal*, May/June, 62–75.

Timmerman, Alan G., 1993, "How learning in financial markets generates excess volatility and predictability of stock returns", *Quarterly Journal of Economics* 108, 1135–1145.

Vasicek, Oldrich A., 1977, "An equilibrium characterization of the term structure", *Journal of Financial Economics* 5, 177–188.

Veronesi, Pietro, 1999, "Stock market overreaction to bad news in good times: A rational expectations equilibrium model", *Review of Financial Studies* 12, 975–1007.

Veronesi, Pietro, 2000, "How does information quality affect stock returns?", *Journal of Finance* 55, 807–837.

Vuolteenaho, Tuomo, 2000, "Understanding the aggregate book-to-market ratio and its implications to current equity-premium expectations", working paper, Harvard University.

Appendix A: Data Construction

We use annual data for the years 1962 through 2000 extracted from the CRSP/Compustat database. Following Fama and French (1993), book equity is constructed as stockholders' equity plus balance sheet deferred taxes and investment tax credit (Compustat item 35) minus the book value of preferred stock. Depending on availability, stockholder's equity is computed as Compustat item 216, or 60+130, or 6-181, in that order, and preferred stock is computed as item 56, or 10, or 130, in that order. Market equity is computed by multiplying the common stock price at fiscal year-end (item 199) by common shares outstanding (item 25). Earnings are calculated as income before extraordinary items, available to common stockholders (item 237), plus deferred taxes from the income statement (item 50), plus investment tax credit (item 51). Debt is total long-term debt (item 9), assets are total assets (item 6), and dividends are dividends available to common stockholders (item 21). M/B ratio is computed as market equity divided by book equity, return on equity is earnings divided by last year's book equity, and leverage is debt divided by total assets. We eliminate the values of market equity, book equity, and total assets smaller than $1 million, as well as M/B ratios smaller than 0.01 and larger than 100. Common stock repurchases are calculated (back to 1971) as purchases of common and preferred stock (item 115) minus purchases of preferred stock (item 56 last year minus this year), following Grullon and Michaely (2000).

Appendix B: Proofs

Proof of Proposition 1: **Part (a)**. Assume $t = 0$ without loss of generality. Since $D_s = cB_s$ and c is a constant, the value of the stock in equation (5) can be rewritten as

$$M_0 = \frac{c}{\pi_0} \int_0^T E_0 \left[\pi_s B_s\right] ds + \frac{1}{\pi_0} E_0 \left[\pi_T B_T\right], \tag{27}$$

where the inversion of the integral with the expectation is justified by the Fubini theorem. We now compute $E_0 [\pi_T B_T]$. Let $p_t = \log (\pi_t B_t)$. Using Ito's Lemma, we obtain

$$dp_t = \left(\rho_t - c - r - \frac{1}{2}\sigma_\pi \sigma_\pi'\right) dt - \sigma_\pi d\mathbf{W}_t.$$

Let $\mathbf{Z}_t = (p_t, \rho_t)'$, so that we can compactly write the joint process for p_t and ρ_t as

$$d\mathbf{Z}_t = (\mathbf{A} + \mathbf{B}\mathbf{Z}_t)\, dt + \mathbf{\Sigma} d\mathbf{W}_t, \tag{28}$$

where

$$\mathbf{A} = \begin{pmatrix} -r - c - \frac{1}{2}\sigma_\pi \sigma_\pi' \\ \phi\overline{\rho} \end{pmatrix}, \ \mathbf{B} = \begin{pmatrix} 0 & 1 \\ 0 & -\phi \end{pmatrix}, \text{ and } \mathbf{\Sigma} = \begin{pmatrix} -\sigma_\pi \\ \sigma_\rho \end{pmatrix}.$$

This is a standard multi-dimensional linear process whose solution is known in closed form (see for example Duffie, 1996, page 293) as

$$\mathbf{Z}_T | \mathbf{Z}_0 \sim \mathcal{N}\left(\mu\left(\mathbf{Z}_0; T\right), \mathbf{\Sigma}_Z\left(T\right)\right), \tag{29}$$

where

$$\begin{aligned} \mu\left(\mathbf{Z}_0, T\right) &= \mathbf{\Psi}\left(T\right) Z_0 + \int_0^T \mathbf{\Psi}\left(T-t\right) \mathbf{A} dt \\ \mathbf{\Sigma}_Z\left(T\right) &= \int_0^T \mathbf{\Psi}\left(T-t\right) \mathbf{\Sigma}\mathbf{\Sigma}' \mathbf{\Psi}\left(T-t\right)' dt. \end{aligned}$$

Above, $\mathbf{\Psi}(T) = \mathbf{U}\exp(\mathbf{\Lambda}\cdot T)\mathbf{U}^{-1}$, where $\mathbf{\Lambda}$ is a diagonal matrix with the eigenvalues of the matrix $\mathbf{B}$ in (28) along the principal diagonal, $\mathbf{U}$ is the matrix of the associated eigenvectors, and $\exp(\mathbf{\Lambda}\cdot T)$ denotes the diagonal matrix with $e^{\lambda_i T}$ in its $ii-$th position. In our setup,

$$\mathbf{\Psi}(T) = \begin{pmatrix} 1 & \frac{1}{\phi} - \frac{1}{\phi}e^{-\phi T} \\ 0 & e^{-\phi T} \end{pmatrix}. \tag{30}$$

The normality of $\mathbf{Z}_T$ implies that $p_T = \mathbf{e}_1\mathbf{Z}_T$, with $\mathbf{e}_1 = (1,0)$, is also normally distributed:

$$p_T|_{Z_0} \sim \mathcal{N}\left(\mathbf{e}_1\mu(\mathbf{Z}_0,T), \mathbf{e}_1\mathbf{\Sigma}_Z(T)\mathbf{e}_1'\right). \tag{31}$$

Using the properties of the lognormal distribution, we obtain

$$E_0[\pi_T B_T] = E_0[e^{p_T}] = E_0[\exp(\mathbf{e}_1\mathbf{Z}_T)] = \exp\left(\mathbf{e}_1\mu(\mathbf{Z}_0,T) + \frac{1}{2}\mathbf{e}_1\mathbf{\Sigma}_Z(T)\mathbf{e}_1'\right). \tag{32}$$

Using (30) in the formulas for $\mu(\mathbf{Z}_0,T)$ and $\mathbf{\Sigma}_Z(T)$, tedious algebra leads to

$$E_0[\pi_T B_T] = B_0\pi_0 Z(\overline{\rho},\rho_0,T),$$

where $Z(\overline{\rho},\rho_0,T)$ is defined in (7). (Detailed steps are available from the authors upon request.) Finally, since the above holds for any T, we obtain

$$M_0 = \frac{c}{\pi_0}\int_0^T E_0[\pi_s B_s]\,ds + \frac{1}{\pi_0}E_0[\pi_T B_T] = B_0\left[c\int_0^T Z(\overline{\rho},\rho_0,s)\,ds + Z(\overline{\rho},\rho_0,T)\right].$$

Part (b). For convenience, rewrite the pricing function as $M_t = B_t G(\rho_t,t)$, where $G(\rho_t,t) = G(\overline{\rho};\rho_t,\tau)$ is defined in equation (10) and $\tau = T-t$. Then we have

$$\begin{aligned}\frac{\partial G}{\partial \rho_t} &= c\int_t^T \frac{\partial Z(\overline{\rho},\rho_t,s)}{\partial \rho_t}ds + \frac{\partial Z(\overline{\rho},\rho_t,\tau)}{\partial \rho_t} \\ &= c\int_t^T \frac{1}{\phi}\left(1-e^{-\phi s}\right)Z(\overline{\rho},\rho_t,s)\,ds + \frac{1}{\phi}\left(1-e^{-\phi\tau}\right)Z(\overline{\rho},\rho_t,\tau),\end{aligned}$$

using the definition of $Z(\overline{\rho},\rho_0,s)$ in equation (7). From Ito's Lemma, we obtain $dM_t = M_t\mu_t dt + M_t\sigma_M d\mathbf{W}_t$, where the volatility vector is given by

$$\sigma_M = \frac{1}{G(\rho_t,t)}\frac{\partial G(\rho_t,t)}{\partial \rho_t}\sigma_\rho = \frac{1}{\phi}F(\rho_t,\tau,c)\,\sigma_\rho,$$

where

$$F(\rho_t,\tau,c) = \frac{c\int_0^\tau\left(1-e^{-\phi s}\right)Z(\overline{\rho},\rho_t,s)\,ds + \left(1-e^{-\phi\tau}\right)Z(\overline{\rho},\rho_t,\tau)}{c\int_0^\tau Z(\overline{\rho},\rho_t,s)\,ds + Z(\overline{\rho},\rho_t,\tau)}. \tag{33}$$

Finally, to obtain the process for excess stock returns, $dR_t = (dM_t + D_t dt)/M_t - r dt$, recall the basic pricing condition (e.g. Duffie, 1996, page 106)

$$E_t[dR_t] = -Cov\left(dR_t, \frac{d\pi_t}{\pi_t}\right)dt = \sigma_R\sigma_\pi' dt.$$

Since the volatility of returns is the same as that of dM/M, $\sigma_{R,t} = \sigma_M$, we immediately see that $dR_t = \mu_{R,t}dt + \sigma_{R,t}d\mathbf{W}_t$ with $\mu_{R,t} = \frac{1}{\phi}F(\rho_t,\tau,c)\sigma_\rho\sigma_\pi'$ and $\sigma_{R,t} = \sigma_M = \frac{1}{\phi}F(\rho_t,\tau,c)\sigma_\rho$. ■

Proof of Lemma 1: This lemma is a simple application of the Kalman-Bucy filter (e.g. Liptser and Shiryayev, 1977). Specifically, the process $\mathbf{Z}_t = (p_t, \rho_t)'$ in (28) can be rewritten as

$$d\mathbf{Z}_t = [\mathbf{A}_0 + \mathbf{A}_1\overline{\rho} + \mathbf{B}\mathbf{Z}_t]\, dt + \mathbf{\Sigma} d\mathbf{W}_t, \tag{34}$$

where $\mathbf{A}_0 = (-r - c - \frac{1}{2}\sigma_\pi \sigma_\pi', 0)'$ and $\mathbf{A}_1 = (0, \phi)'$. This is a signal process for $\overline{\rho}$. Let $\widehat{\rho}_t = E_t[\overline{\rho}]$ be the expectation of $\overline{\rho}$ conditional on the information set $\mathcal{F}_t = \{\mathbf{Z}_\tau : 0 \le \tau \le t\}$. Define the orthogonalized expectation error,

$$d\widetilde{\mathbf{W}}_t = \mathbf{\Sigma}^{-1}[d\mathbf{Z}_t - E_t(d\mathbf{Z}_t)] = \mathbf{\Sigma}^{-1}[d\mathbf{Z}_t - (\mathbf{A}_0 + \mathbf{A}_1\widehat{\rho}_t + \mathbf{B}\mathbf{Z}_t)\, dt]. \tag{35}$$

Theorem 10.3 in Liptser and Shiryayev (1977) then implies that $\widetilde{\mathbf{W}}_t$ is a standard Wiener process with respect to $\{\mathcal{F}_t\}$. Given a prior distribution at time $t = 0$ of $\overline{\rho} \sim \mathcal{N}(\widehat{\rho}_0, \widehat{\sigma}_0^2)$, the conditional expectation $\widehat{\rho}_t$ satisfies the stochastic differential equation

$$d\widehat{\rho}_t = \widehat{\sigma}_t^2 \mathbf{A}_1' (\mathbf{\Sigma}')^{-1} d\widetilde{\mathbf{W}}_t = \widehat{\sigma}_t^2 \frac{\phi}{\sigma_{\rho,2}} d\widetilde{W}_{2,t}.$$

Above, $\widehat{\sigma}_t^2 = E_t\left[(\rho_t - \overline{\rho})^2\right]$ satisfies the Riccati differential equation

$$\frac{d(\widehat{\sigma}_t^2)}{dt} = -(\widehat{\sigma}_t^2)^2 \mathbf{A}_1' (\mathbf{\Sigma}\mathbf{\Sigma}')^{-1} \mathbf{A}_1 = -(\widehat{\sigma}_t^2)^2 \left(\frac{\phi}{\sigma_{\rho,2}}\right)^2,$$

whose solution is

$$\widehat{\sigma}_t^2 = \frac{1}{\frac{1}{\widehat{\sigma}_0^2} + \left(\frac{\phi}{\sigma_{\rho,2}}\right)^2 t}.$$

Note that the information filtration generated by $\mathbf{Z}_t = (\log(\pi_t B_t), \rho_t)'$ is the same as that generated by $\mathbf{Z}_t^* = (\log(\pi_t), \rho_t)'$, because $\log(B_t)$ is an $\mathcal{F}_t$-predictable process. Formally,

$$\begin{aligned} d\widetilde{\mathbf{W}}_t^* &= \mathbf{\Sigma}^{-1}[d\mathbf{Z}_t^* - E_t(d\mathbf{Z}_t^*)] = \mathbf{\Sigma}^{-1}\left[\left(d\mathbf{Z}_t^* + (\rho_t - c, 0)' dt\right) - \left(E_t(d\mathbf{Z}_t^*) + (\rho_t - c, 0)' dt\right)\right] \\ &= \mathbf{\Sigma}^{-1}[d\mathbf{Z}_t - E_t(d\mathbf{Z}_t)] = d\widetilde{\mathbf{W}}_t, \end{aligned}$$

so that $d\widetilde{\mathbf{W}}_t^*$ is (almost surely) the same as $d\widetilde{\mathbf{W}}_t$. This concludes the proof of Lemma 1. ■

For later reference, it is convenient to rewrite the original process in $d\mathbf{Z}_t$ in (28) in terms of the new Brownian motion $d\widetilde{\mathbf{W}}_t$:

$$d\mathbf{Z}_t = [\mathbf{A}_0 + \mathbf{A}_1\widehat{\rho}_t + \mathbf{B}\mathbf{Z}_t]\, dt + \mathbf{\Sigma} d\widetilde{\mathbf{W}}_t. \tag{36}$$

Proof of convexity of $G(\overline{\rho}, \rho_t, \tau)$ in $\overline{\rho}$. Let $Q_\rho(s) = \left(1 - e^{-\phi s}\right)/\phi$. First note that

$$\frac{\partial Z(\overline{\rho}, \rho_t, s)}{\partial \overline{\rho}} = (s - Q_\rho(s))\, Z(\overline{\rho}, \rho_t, s) \quad \text{and} \quad \frac{\partial^2 Z(\overline{\rho}, \rho_t, s)}{\partial \overline{\rho}^2} = (s - Q_\rho(s))^2 Z(\overline{\rho}, \rho_t, s).$$

From the definition of $G(\overline{\rho}, \rho_t, \tau)$ in equation (10), we have

$$\begin{aligned} \frac{\partial G(\overline{\rho}, \rho_t, \tau)}{\partial \overline{\rho}} &= c\int_0^\tau \frac{\partial Z(\overline{\rho}, \rho_t, s)}{\partial \overline{\rho}} ds + \frac{\partial Z(\overline{\rho}, \rho_t, \tau)}{\partial \overline{\rho}} \\ &= c\int_0^\tau (s - Q_\rho(s))\, Z(\overline{\rho}, \rho_t, s)\, ds + (\tau - Q_\rho(\tau))\, Z(\overline{\rho}, \rho_t, \tau) \\ \frac{\partial^2 G(\overline{\rho}, \rho_t, \tau)}{\partial \overline{\rho}^2} &= c\int_0^\tau (s - Q_\rho(s))^2 Z(\overline{\rho}, \rho_t, s)\, ds + (\tau - Q_\rho(\tau))^2 Z(\overline{\rho}, \rho_t, \tau) > 0. \end{aligned}$$ ■

Proof of Proposition 2. Part (a). The pricing equation is

$$M_t = E_t\left[\int_t^T \frac{\pi_\tau}{\pi_t} D_\tau d\tau + \frac{\pi_T}{\pi_t} B_T\right] = E_t\left[E_t\left[\int_t^T \frac{\pi_\tau}{\pi_t} D_\tau d\tau + \frac{\pi_T}{\pi_t} B_T | \overline{\rho}\right]\right] = B_t E_t\left[G\left(\overline{\rho}, \rho_t, \tau\right)\right],$$

with $G\left(\overline{\rho}, \rho_t, \tau\right)$ given in (10). The function $Z\left(\overline{\rho}, \rho_t, \tau\right)$ can be rewritten as

$$Z\left(\overline{\rho}, \rho_t, \tau\right) = Z_1\left(\rho_t, \tau\right) Z_2\left(\overline{\rho}, \tau\right),$$

where $Z_1\left(\rho_t, \tau\right) = \exp\left[-\left(r + c\right)\tau + Q\left(\tau\right) + Q_\rho\left(\tau\right)\rho_t\right]$ and $Z_2\left(\overline{\rho}, \tau\right) = \exp\left[\overline{\rho}\left(\tau - Q_\rho\left(\tau\right)\right)\right]$. Hence

$$\begin{aligned} E_t\left[G\left(\overline{\rho}, \rho_t, \tau\right)\right] &= c\int_0^\tau E_t\left[Z\left(\overline{\rho}, \rho_t, s\right)\right] ds + E_t\left[Z\left(\overline{\rho}, \rho_t, \tau\right)\right] \\ &= c\int_0^\tau Z_1\left(\rho_t, s\right) E_t\left[Z_2\left(\overline{\rho}, s\right)\right] ds + Z_1\left(\rho_t, \tau\right) E_t\left[Z_2\left(\overline{\rho}, \tau\right)\right]. \end{aligned}$$

Since $\overline{\rho}$ is normally distributed (Lemma 1), $\overline{\rho}\left(\tau - Q_\rho\left(\tau\right)\right)$ is normally distributed with mean $\widehat{\rho}_t\left(\tau - Q_\rho\left(\tau\right)\right)$ and variance $\widehat{\sigma}_t^2\left(\tau - Q_\rho\left(\tau\right)\right)^2$. Using the properties of the lognormal distribution,

$$Z_2\left(\widehat{\rho}_t, \widehat{\sigma}_t^2, \tau\right) \equiv E_t\left[Z_2\left(\overline{\rho}, \tau\right)\right] = \exp\left(\widehat{\rho}_t\left(\tau - Q_\rho\left(\tau\right)\right) + \frac{1}{2}\widehat{\sigma}_t^2\left(\tau - Q_\rho\left(\tau\right)\right)^2\right).$$

Finally, defining

$$\begin{aligned} Z^U\left(\widehat{\rho}_t, \widehat{\sigma}_t^2, \rho_t, \tau\right) &= Z_1\left(\rho_t, \tau\right) Z_2\left(\widehat{\rho}_t, \widehat{\sigma}_t^2, \tau\right) && (37) \\ &= \exp\left\{-\left(r + c - \widehat{\rho}_t\right)\tau + Q\left(\tau\right) + Q_\rho\left(\tau\right)\left(\rho_t - \widehat{\rho}_t\right) + \frac{1}{2}\widehat{\sigma}_t^2\left(\tau - Q_\rho\left(\tau\right)\right)^2\right\}, && (38) \end{aligned}$$

we obtain

$$M_t = B_t G^U\left(\widehat{\rho}_t, \widehat{\sigma}_t^2, \rho_t, \tau\right), \qquad (39)$$

where

$$G^U\left(\widehat{\rho}_t, \widehat{\sigma}_t^2, \rho_t, \tau\right) = c\int_0^\tau Z^U\left(\widehat{\rho}_t, \widehat{\sigma}_t^2, \rho_t, s\right) ds + Z^U\left(\widehat{\rho}_t, \widehat{\sigma}_t^2, \rho_t, \tau\right). \qquad (40)$$

Part (b). For simplicity, denote $G^U\left(t\right) = G^U\left(\widehat{\rho}_t, \widehat{\sigma}_t^2, \rho_t, \tau\right)$. Then (39) and (36) imply

$$\begin{aligned} dM_t &= G^U\left(t\right) dB_t + B_t dG^U\left(t\right) \\ &= \left\{\begin{array}{l} G^U\left(t\right)\ \left(\rho_t - c\right) B_t + B_t \frac{\partial G^U}{\partial \rho_t}\phi\left(\widehat{\rho}_t - \rho_t\right) + B_t \frac{\partial G^U(t)}{\partial t} - B_t \frac{\partial G^U(t)}{\partial \widehat{\sigma}_t^2}\left(\widehat{\sigma}_t^2 \frac{\phi_\rho}{\sigma_{\rho,2}}\right)^2 \\ \quad + \frac{1}{2} B_t\left(\frac{\partial^2 G^U(t)}{\partial \rho_t^2}\sigma_\rho \sigma_\rho' + \frac{\partial^2 G^U(t)}{\partial \widehat{\rho}_t^2}\widehat{\sigma}_{\rho,t}\widehat{\sigma}_{\rho,t}' + 2\frac{\partial^2 G^U(t)}{\partial \widehat{\rho}_t \partial \rho_t}\sigma_\rho \widehat{\sigma}_{\rho,t}'\right) \end{array}\right\} dt \\ &\quad + B_t\left(\frac{\partial G^U\left(t\right)}{\partial \rho_t}\sigma_\rho + \frac{\partial G^U\left(t\right)}{\partial \widehat{\rho}_t}\widehat{\sigma}_{\rho,t}\right) d\widetilde{\mathbf{W}}_t, \end{aligned}$$

where $\widehat{\sigma}_{\rho,t}$ is given in equation (22). From the definition of $G^U\left(t\right)$ in equation (40), we have

$$\begin{aligned} \frac{\partial G^U\left(t\right)}{\partial \rho_t} &= c\int_0^\tau \frac{\partial Z^U\left(\widehat{\rho}_t, \widehat{\sigma}_t^2, \rho_t, s\right)}{\partial \rho_t} ds + \frac{\partial Z^U\left(\widehat{\rho}_t, \widehat{\sigma}_t^2, \rho_t, \tau\right)}{\partial \rho_t} \\ &= \frac{1}{\phi}\left(c\int_0^\tau \left(1 - e^{-\phi s}\right) Z^U\left(\widehat{\rho}_t, \widehat{\sigma}_t^2, \rho_t, s\right) ds + \left(1 - e^{-\phi\tau}\right) Z^U\left(\widehat{\rho}_t, \widehat{\sigma}_t^2, \rho_t, \tau\right)\right), \\ \frac{\partial G^U\left(t\right)}{\partial \widehat{\rho}_t} &= c\int_0^\tau \frac{\partial Z^U\left(\widehat{\rho}_t, \widehat{\sigma}_t^2, \rho_t, s\right)}{\partial \widehat{\rho}_t} ds + \frac{\partial Z^U\left(\widehat{\rho}_t, \widehat{\sigma}_t^2, \rho_t, \tau\right)}{\partial \widehat{\rho}_t} \\ &= \frac{1}{\phi}\left(c\int_t^T \left(\phi s - 1 + e^{-\phi s}\right) Z^U\left(\widehat{\rho}_t, \widehat{\sigma}_t^2, \rho_t, s\right) ds + \left(\phi\tau - 1 + e^{-\phi\tau}\right) Z^U\left(\widehat{\rho}_t, \widehat{\sigma}_t^2, \rho_t, \tau\right)\right). \end{aligned}$$

It is useful to define the functions

$$F\left(\rho_t;\widehat{\rho}_t,\widehat{\sigma}_t^2,\tau,c\right) = \frac{c\int_0^\tau\left(1-e^{-\phi s}\right)Z^U\left(\widehat{\rho}_t,\widehat{\sigma}_t^2,\rho_t,s\right)ds+\left(1-e^{-\phi\tau}\right)Z^U\left(\widehat{\rho}_t,\widehat{\sigma}_t^2,\rho_t,\tau\right)}{c\int_0^\tau Z^U\left(\widehat{\rho}_t,\widehat{\sigma}_t^2,\rho_t,s\right)ds+Z^U\left(\widehat{\rho}_t,\widehat{\sigma}_t^2,\rho_t,\tau\right)} \quad (41)$$

$$F_1\left(\rho_t;\widehat{\rho}_t,\widehat{\sigma}_t^2,\tau,c\right) = \frac{c\int_0^\tau\left(\phi s-1+e^{-\phi s}\right)Z^U\left(\widehat{\rho}_t,\widehat{\sigma}_t^2,\rho_t,s\right)ds+\left(\phi\tau-1+e^{-\phi\tau}\right)Z^U\left(\widehat{\rho}_t,\widehat{\sigma}_t^2,\rho_t,\tau\right)}{c\int_0^\tau Z^U\left(\widehat{\rho}_t,\widehat{\sigma}_t^2,\rho_t,s\right)ds+Z^U\left(\widehat{\rho}_t,\widehat{\sigma}_t^2,\rho_t,\tau\right)} \quad (42)$$

Then

$$\begin{aligned}\sigma_{R,t} &= \frac{1}{G^U(t)}\left(\frac{\partial G^U(t)}{\partial\rho_t}\sigma_\rho+\frac{\partial G^U(t)}{\partial\widehat{\rho}_t}\widehat{\sigma}_{\rho,t}\right)\\ &= F\left(\rho_t;\widehat{\rho}_t,\widehat{\sigma}_t^2,\tau,c\right)\frac{1}{\phi}\sigma_\rho+F_1\left(\rho_t;\widehat{\rho}_t,\widehat{\sigma}_t^2,\tau,c\right)\frac{1}{\phi}\widehat{\sigma}_{\rho,t}.\end{aligned}$$

Hence, using the same argument as in the proof of Proposition 1, expected return is given by

$$\begin{aligned}\mu_R &= \sigma_R\sigma_\pi' = F\left(\rho_t;\widehat{\rho}_t,\widehat{\sigma}_t^2,\tau,c\right)\frac{1}{\phi}\sigma_\rho\sigma_\pi'+F_1\left(\rho_t;\widehat{\rho}_t,\widehat{\sigma}_t^2,\tau,c\right)\frac{1}{\phi}\widehat{\sigma}_{\rho,t}\sigma_\pi'\\ &= F\left(\rho_t;\widehat{\rho}_t,\widehat{\sigma}_t^2,\tau,c\right)\frac{1}{\phi}\sigma_\rho\sigma_\pi',\end{aligned}$$

since $\sigma_\pi=(\sigma_{\pi,1},0)$. ■

Proofs of Corollaries 3 and 4: From the definition of $Z^U\left(\rho_t;\widehat{\rho}_t,\widehat{\sigma}_t^2,\tau\right)$ in (37) , we have

$$\frac{\partial Z^U\left(\rho_t;\widehat{\rho}_t,\widehat{\sigma}_t^2,\tau\right)}{\partial\widehat{\sigma}_t^2}=\frac{1}{2}\left(\tau-Q_\rho(\tau)\right)^2Z^U\left(\rho_t;\widehat{\rho}_t,\widehat{\sigma}_t^2,\tau\right).$$

Using the definition of $G^U\left(\widehat{\rho}_t,\widehat{\sigma}_t^2,\rho_t,\tau\right)$ in (40),

$$\begin{aligned}\frac{\partial\log\left(G^U\left(\widehat{\rho}_t,\widehat{\sigma}_t^2,\rho_t,\tau\right)\right)}{\partial\widehat{\sigma}_t^2} &= \frac{1}{G^U\left(\widehat{\rho}_t,\widehat{\sigma}_t^2,\rho_t,\tau\right)}\frac{\partial G^U\left(\widehat{\rho}_t,\widehat{\sigma}_t^2,\rho_t,\tau\right)}{\partial\widehat{\sigma}_t^2}\\ &= \frac{c\int_0^\tau\frac{1}{2}\left(s-Q_\rho(s)\right)^2Z^U\left(\widehat{\rho}_t,\widehat{\sigma}_t^2,\rho_t,s;c\right)ds+\frac{1}{2}\left(\tau-Q_\rho(\tau)\right)^2Z^U\left(\widehat{\rho}_t,\widehat{\sigma}_t^2,\rho_t,\tau;c\right)}{c\int_0^\tau Z^U\left(\widehat{\rho}_t,\widehat{\sigma}_t^2,\rho_t,s\right)ds+Z^U\left(\widehat{\rho}_t,\widehat{\sigma}_t^2,\rho_t,\tau\right).}>0,\end{aligned}$$

proving Corollary 3. Turning to Corollary 4, note that by setting $c=0$ we have

$$\frac{\partial\log\left(G^U\left(\widehat{\rho}_t,\widehat{\sigma}_t^2,\rho_t,\tau\right)\right)}{\partial\widehat{\sigma}_t^2}|_{c=0}=\frac{1}{2}\left(\tau-Q_\rho(\tau)\right)^2.$$

Simple algebra then shows

$$\begin{aligned}\frac{\partial\log\left(G^U\left(\widehat{\rho}_t,\widehat{\sigma}_t^2,\rho_t,\tau\right)\right)}{\partial\widehat{\sigma}_t^2}|_{c=0} &- \frac{\partial\log\left(G^U\left(\widehat{\rho}_t,\widehat{\sigma}_t^2,\rho_t,\tau\right)\right)}{\partial\widehat{\sigma}_t^2}|_{c>0}=\\ &\frac{c\int_0^\tau\frac{1}{2}\left[\left(\tau-Q_\rho(\tau)\right)^2-\left(s-Q_\rho(s)\right)^2\right]Z^U\left(\widehat{\rho}_t,\widehat{\sigma}_t^2,\rho_t,s\right)ds}{c\int_0^\tau Z^U\left(\widehat{\rho}_t,\widehat{\sigma}_t^2,\rho_t,s\right)ds+Z^U\left(\widehat{\rho}_t,\widehat{\sigma}_t^2,\rho_t,\tau\right)}>0.\end{aligned}$$

The inequality stems from the fact that $f(s)=\left(s-Q_\rho(s)\right)^2$ is strictly increasing, and hence $\left(\tau-Q_\rho(\tau)\right)^2-\left(s-Q_\rho(s)\right)^2>0$ for every $\tau>s$. (Note that $f'(s)=2\left(s-Q_\rho(s)\right)\left(1-e^{-\phi s}\right)>0$, because $\left(s-Q_\rho(s)\right)>0$ for all s.) ■

Derivation of expression (25). Since $c = 0$, the pricing equation from (19) is $M_t/B_t = Z^U\left(\rho_t; \widehat{\rho}_t, \widehat{\sigma}_t^2, \tau\right)$. Taking logs and substituting for $Z^U\left(\rho_t; \widehat{\rho}_t, \widehat{\sigma}_t^2, \tau\right)$ from equation (37), we find

$$\begin{aligned}\log\left(\frac{M_t}{B_t}\right) &= -\left(r-\widehat{\rho}_t\right)\tau + \frac{1}{\phi}\left(1-e^{-\phi\tau}\right)\left(\rho_t-\widehat{\rho}_t\right) + \left(\frac{1}{2\phi^2}\widehat{\sigma}_t^2\left(\phi\tau-1+e^{-\phi\tau}\right)^2\right) \\ &\quad + \frac{\sigma_\rho\sigma_\rho'}{2\phi^3}\left(\frac{1-e^{-2\phi\tau}}{2}+\phi\tau-2\left(1-e^{-\phi\tau}\right)\right) + \frac{\sigma_\pi\sigma_\rho'}{\phi^2}\left(1-e^{-\phi\tau}-\phi\tau\right).\end{aligned}$$

From Proposition 2, we have $\mu_{R,t} = \left(1-e^{-\phi\tau}\right)\frac{\sigma_\rho\sigma_\pi'}{\phi}$, so we can substitute for $\sigma_\pi\sigma_\rho'$ to obtain

$$\log\left(\frac{M_t}{B_t}\right) = \alpha_0(\tau) + \alpha_1(\tau)\widehat{\rho}_t + \alpha_2(\tau)\rho_t + \alpha_3(\tau)\mu_{R,t} + \alpha_4(\tau)\sigma_\rho\sigma_\rho' + \alpha_5(\tau)\widehat{\sigma}_t^2,$$

where, since $\tau\phi > 1-e^{-\phi\tau} > 0$ for all $\tau > 0$,

$$\begin{aligned}\alpha_0(\tau) &= -r\tau < 0\ ; \quad \alpha_1(\tau) = \left(\tau - \frac{1}{\phi}\left(1-e^{-\phi\tau}\right)\right) > 0\ ; \quad \alpha_2(\tau) = \frac{1}{\phi}\left(1-e^{-\phi\tau}\right) > 0; \\ \alpha_3(\tau) &= \frac{\left(\left(1-e^{-\phi\tau}\right)\frac{1}{\phi}-\tau\right)}{\left(1-e^{-\phi\tau}\right)} < 0\ ; \quad \alpha_4(\tau) = \frac{1}{2\phi^3}\left(\frac{1-e^{-2\phi\tau}}{2}+\phi\tau-2\left(1-e^{-\phi\tau}\right)\right) > 0; \\ \alpha_5(\tau) &= \frac{1}{2}\left(\tau-\frac{1}{\phi}\left(1-e^{-\phi\tau}\right)\right)^2 > 0. \blacksquare\end{aligned}$$

Proof of Corollary 6: For notational convenience, denote $Z^U\left(\widehat{\rho}_t, \widehat{\sigma}_t^2, \rho_t, s\right)$ by $Z^U(t,s)$ and $F\left(\rho_t; \widehat{\rho}_t, \widehat{\sigma}_t^2, \tau, c\right)$ by $F(t,\tau,c)$. From equation (21), the return volatility vector is given by

$$\sigma_{R,t} = F(t,\tau,c)\ \frac{1}{\phi}\sigma_\rho + F_1(t,\tau,c)\ \frac{1}{\phi}\widehat{\sigma}_\rho,$$

where $F(t,\tau,c)$ and $F_1(t,\tau,c)$ are given in equations (41) and (42), respectively. Setting $c = 0$, we obtain $F(t,\tau,0) = \left(1-e^{-\phi\tau}\right)$ and $F_1(t,\tau,0) = \left(\phi\tau-1+e^{-\phi\tau}\right)$. Since $F(.)$ and $F_1(.)$ are strictly positive functions, the claim of the corollary is proved if we show

$$F(t,\tau,0) - F(t,\tau,c) > 0 \text{ and } F_1(t,\tau,0) - F_1(t,\tau,c) > 0.$$

Note that

$$F(t,\tau,0) - F(t,\tau,c) = \frac{c\int_0^\tau\left(e^{-\phi s}-e^{-\phi\tau}\right)Z^U(t,s)\,ds}{c\int_0^\tau Z^U(t,s)\,ds + Z^U(t,\tau)} > 0,$$

as $\phi > 0$ implies $e^{-\phi s} - e^{-\phi\tau} > 0$ for $s < \tau$. Similarly,

$$F_1(t,\tau,0) - F_1(t,\tau,c) = \frac{c\int_0^\tau\left(\left(\phi\tau+e^{-\phi\tau}\right)-\left(\phi s+e^{-\phi s}\right)\right)Z^U(t,s)\,ds}{c\int_0^\tau Z^U(t,s)\,ds + Z^U(t,\tau)} > 0,$$

because $f(s) = \left(\phi s + e^{-\phi s}\right)$ is strictly increasing in s, and hence $f(\tau) - f(s) > 0$ for all $s < \tau$. $\blacksquare$