Title
Risk, Return and Dividends

Authors
Ang, Andrew
Liu, Jun

Publication Date
2005-03-01

Risk, Return and Dividends*

Andrew Ang†
Columbia University, USC and NBER

Jun Liu‡
UCLA

First Version: 12 September, 2004

JEL Classification: G12
Keywords: risk-return trade-off, risk premium,
stochastic volatility, predictability

*We especially thank John Cochrane, as portions of this manuscript originated from extensive conversations between John and the authors. We also thank Joe Chen, Chris Jones, Greg Willard, and seminar participants at Columbia University, ISCTE Business School (Lisbon), LSE, Melbourne Business School, UCLA, University of Maryland, University of Michigan, and USC for helpful comments.

†Marshall School of Business, USC, 701 Exposition Blvd, Rm 701, Los Angeles, CA 90089; ph: (213) 740-5615; fax: (213) 740-6650; email: aa610@columbia.edu; WWW: http://www.columbia.edu/∼aa610.

‡C509 Anderson School, UCLA CA 90095. Email: jliu@anderson.ucla.edu, ph: (310) 825-4083, WWW: http://www.personal.anderson.ucla.edu/jun.liu/

Abstract

We characterize the joint dynamics of expected returns, stochastic volatility, and prices. In particular, with a given dividend process, one of the processes of the expected return, the stock volatility, or the price-dividend ratio fully determines the other two. For example, the stock volatility determines the expected return and the price-dividend ratio. By parameterizing one, or more, of expected returns, volatility, or prices, common empirical specifications place strong, and sometimes inconsistent, restrictions on the dynamics of the other variables. Our results are useful for understanding the risk-return trade-off, as well as characterizing the predictability of stock returns.

1 Introduction

We fully characterize the relationship between expected returns, stock volatility and prices by using the dividend process of a stock, and derive over-identifying restrictions on the dynamics of these variables. We show that given the dividend process, it is enough to specify one of the expected return, the stock return volatility, or the price-dividend ratio. Determining one of these variables completely determines the other two. These relations are not merely technical restrictions, but they lend insight into the nature of the risk-return relation and the predictability of stock returns.

Our method of using the dividend process to characterize the risk-return relation requires no economic assumptions other than ruling out asset price bubbles. In particular, we do not require the preferences of agents, equilibrium concepts, or a pricing kernel. This is in contrast to previous work that requires equilibrium conditions, in particular, the utility function of a representative agent, to pin down the risk-return relation. For example, in a standard CAPM or Merton (1973) model, the expected return of the market is a product of the relative risk aversion coefficient of the representative agent and the variance of the market return.

The intuition behind our risk-return relations is a simple observation that, by definition, returns comprise both capital gain and dividend yield components. Hence, the return is a function of price-dividend ratios and dividend growth rates. Thus, given the dividend process, if we specify the expected return process, we can compute price-dividend ratios. The second moment of the return, or equivalently the approximate volatility process, is also a function of price-dividend ratios and dividend growth rates. Thus, using dividends and price-dividend ratios, we can compute the volatility process of the stock. Going in the opposite direction, if dividends are given and we specify a process for stochastic volatility, we can back out the price-dividend ratio, because the second moment of returns is a function of price-dividend ratios and dividend growth. The price-dividend ratio, together with cashflow growth rates, can be used to infer the process for expected returns. In continuous-time, expected returns, stock volatility, and price-dividend ratios are linked by a series of differential equations.

Our risk-return relations are empirically relevant because our conditions impose stringent restrictions on asset pricing models. Many common empirical applications often directly specify one of the expected return, risk, or the price-dividend ratio. Often, this is done without considering the dynamics of the other two variables. Our results show that specifying the expected return automatically pins down the diffusion term of returns and vice versa. Hence, specifying one of the expected return, risk, or the price-dividend ratio makes implicit assumptions about

the dynamics of these other variables. Our relations can be used as checks of internal consistency for empirical specifications that usually concentrate on only one of predictable expected returns, stochastic volatility, or price-dividend ratio dynamics.

We illustrate several applications of our risk-return conditions with popular empirical specifications from the literatures of predictability of expected returns and time-varying volatility. First, a large literature beginning with Fama and French (1988a) forecasts expected returns with dividend yields in a linear regression framework. A large asset allocation literature uses these empirical specifications and parameterize conditional expected returns as linear functions of dividend yields.[1] This specification implies that returns are heteroskedastic, and places strong restrictions on the price process. In particular, the drift of the dividend yield is non-linear and generally not stationary. Conversely, if the dividend yield follows a mean-reverting linear process, like the AR(1) specifications assumed by Stambaugh (1999), Campbell and Yogo (2003), and Lewellen (2003), then expected returns cannot be linear functions of the dividend yield, and linear approximations to the drift of the expected return as a function of the dividend yield can be highly inaccurate.

Second, we investigate the implications of predictable mean-reverting components of returns on return volatility and prices. Poterba and Summers (1986) and Fama and French (1988b) find slow mean-reverting components of returns. Even under IID dividend growth, mean-reverting expected returns implies that the expected return must be a non-linear, increasing function of the dividend yield. However, the stochastic volatility generated by mean-reverting expected returns is several orders too small in magnitude to match the time-varying volatility present in data.

Third, it is well known that volatility is more precisely estimated than first moments (see Merton, 1980). Since Engle (1982), a wide variety of ARCH or stochastic volatility models have been used successfully to capture time-varying second moments in asset prices. However, this literature mostly concentrates on specifying the diffusion components of stock returns without considering the implications for the expected return.[2] If we specify the diffusion of the stock return, then, assuming a dividend process, stock prices and expected returns are fully determined.

[1] See, among many others, Kandel and Stambaugh (1996), Balduzzi and Lynch (1999), Campbell and Viceira (1999), Barberis (2000), and Wachter (2004).

[2] Exceptions to this are the GARCH, or stochastic volatility, models that parameterize time-varying variances of an intertemporal asset pricing model. Harvey (1989), Ferson and Harvey (1991), and Scruggs (1998), among others, estimate models of this type.

The idea of using the dividend process to characterize the relationship between risk and return goes back to at least Grossman and Shiller (1981) and Shiller (1981), who argue that the volatility of stock returns is too high compared to the volatility of dividend growth. Campbell and Shiller (1988a and b) linearize the definition of returns and then iterate to derive an approximate relation for the log price-dividend ratio. They use this relation to measure the role of cashflow and discount rates in the variation of price-dividend ratios. Our approach is similar, in that we use the definition of returns to derive relationships between risk, returns, and prices. However, our relations tie expected returns, stochastic volatility, and price-dividend ratios more tightly and rigorously than the linearized price-dividend ratio formula of Campbell and Shiller. Furthermore, we are able to provide exact characterizations between the conditional second moments of returns and prices (the stochastic volatility of returns, and the conditional volatility of expected returns, dividend growth, and price-dividend ratios) that Campbell and Shiller's framework cannot easily handle.

Our risk-return conditions are most closely related to He and Leland (1993), who show that the drift and diffusion term of the price process must satisfy a partial differential equation and a boundary condition in a pure exchange economy. He and Leland show that the form of the risk-return relation is a function of the curvature of the representative agent's utility. Using dividends, rather than preferences, to pin down the risk-return relationship is advantageous because dividends are observable, allowing a stochastic dividend process to be easily estimated. Indeed, a convenient assumption made by many models is that dividend growth is IID. In comparison, there is still no consensus on the precise form that a representative agent's utility should take.

The remainder of the paper is organized as follows. Section 2 derives the risk-return and pricing relations for an economy with a set of state variables driving the time-varying investment opportunity set. In Section 3, we apply these conditions to various empirical specifications in the literature, covering predictability of expected returns by dividend yields, mean-reverting expected returns, and models of stochastic volatility. Section 4 concludes. We relegate all proofs to the Appendix.

2 The Model

Suppose that the state of the economy is described by a single state variable x_t, which follows the diffusion process:

$$dx_t = \mu_x(x_t)dt + \sigma_x(x_t)dB_t^x, \tag{1}$$

where the drift $\mu_x(\cdot)$ and diffusion $\sigma_x(\cdot)$ are functions of x_t. We assume that there is a risky asset that pays the dividend stream D_t, which follows the process:

$$\frac{dD_t}{D_t} = \left(\mu_d(x_t) + \frac{1}{2}\sigma_d^2(x_t)\right)dt + \sigma_d(x_t)dB_t^d, \tag{2}$$

or equivalently:

$$\frac{D_t}{D_0} = \exp\left(\int_0^t \mu_d(x_s)ds + \sigma_d(x_s)dB_s^d\right).$$

For notational simplicity, we assume that shocks to the state variable x_t and shocks to the dividend process are orthogonal, that is dB_t^x and dB_t^d are independent. However, our results apply in a similar fashion to the case when dB_t^x and dB_t^d are correlated.

By definition, the price of the asset P_t is related to dividends D_t and expected returns μ_r by:

$$\frac{\mathrm{E}_t[dP_t] + D_t dt}{P_t} = \mu_r dt. \tag{3}$$

By iterating equation (3), we can write the price as:

$$P_t = \mathrm{E}_t\left[\int_t^T e^{-(\int_t^s \mu_r du)} D_s\, ds + e^{-(\int_t^T \mu_r du)} P_T\right]. \tag{4}$$

Our goal is to determine the drift $\mu_r(\cdot)$ and diffusion $\sigma_r(\cdot)$ of the return process dR_t:

$$dR_t = \mu_r(x_t)dt + \sigma_r(x_t)dB_t^r, \tag{5}$$

under a no-bubble condition:

Assumption 2.1 *The transversality condition*

$$\lim_{T\to\infty} \mathrm{E}_t\left[e^{-(\int_t^T \mu_r du)} P_T\right] = 0$$

holds almost surely.

Assumption 2.1 rules out specifications like the Black-Scholes (1973) and Merton (1973) models, which specify that the stock does not pay dividends. Equivalently, Black, Scholes, and Merton assume that the capital gain represents the entire stock return, so the stock is a bubble process in these economies. By assuming transversality, we can express the stock price

in equation (4) as:[3]

$$P_t = \mathrm{E}_t\left[\int_t^\infty e^{-(\int_t^s \mu_r du)} D_s \, ds\right]. \tag{6}$$

The following proposition characterizes the relationships between dividend growth, the drift and diffusion of the return process dR_t, and price-dividend ratios, subject to Assumption 2.1:[4]

Proposition 2.1 *Suppose the state of the economy is described by x_t, which follows equation (1), and a stock is a claim to the dividends D_t that are described by equation (2). If the price-dividend ratio P_t/D_t is a function $f(x_t)$ of x_t, then the cumulative stock return process dR_t satisfies the following equation:*

$$dR_t = \left(\frac{\mu_x f' + \frac{1}{2}\sigma_x^2 f'' + 1}{f} + \mu_d + \frac{1}{2}\sigma_d^2\right) dt + \sigma_x (\ln f)' \, dB_t^x + \sigma_d \, dB_t^d. \tag{7}$$

Conversely, if the return R_t follows the following diffusion equation:[5]

$$dR_t = \mu_r(x_t)dt + \sigma_{rx}(x_t)dB_t^x + \sigma_{rd}(x_t)dB_t^d, \tag{8}$$

and the stock dividend process is given by equation (1), then the price-dividend ratio $P_t/D_t = f(x_t)$ satisfies the following relation:

$$\mu_x f' + \frac{1}{2}\sigma_x^2 f'' - \left(\mu_r - \mu_d - \frac{1}{2}\sigma_d^2\right) f = -1, \tag{9}$$

and the diffusion of the stock return is determined from the relations:

$$\sigma_{rx} = \sigma_x (\ln f)' \tag{10}$$

$$\sigma_{rd} = \sigma_d. \tag{11}$$

[3] An alternative way to compute the stock price is to iterate the definition of returns $dR_t = (dP_t + D_t dt)/P_t$ forward under the transversality condition $\lim_{T\to\infty} \exp(-(\int_t^T dR_u - \frac{1}{2}\sigma_r^2 du))P_T = 0$ to obtain:

$$P_t = \int_t^\infty e^{-(\int_t^s dR_u - \frac{1}{2}\sigma_r^2 du)} D_s \, ds.$$

This equation holds path by path. As Campbell (1993) notes, we can take conditional expectations of both the left- and right-hand sides to obtain:

$$P_t = \mathrm{E}_t\left[\int_t^\infty e^{-(\int_t^s dR_u - \frac{1}{2}\sigma_r^2 du)} D_s \, ds\right],$$

which can be shown to be equivalent to equation (6).

[4] Although Proposition 2.1 is stated for a univariate state variable x_t, the equations generalize easily to the case where x_t is a vector of state variables. In this case, the ordinary differential equation (9) becomes a partial differential equation, and μ_x, σ_x, μ_d, σ_d, σ_{rx}, and σ_{rd} represent matrix functions of x.

[5] Since dB_t^x and dB_t^d are independent, the diffusion term $\sigma_r(x_t)$ of the return process in equation (5) is given by $\sqrt{\sigma_{rx}(x_t)^2 + \sigma_{rd}(x_t)^2}$.

There are several implications of Proposition 2.1. Most importantly, given the dividend process, specifying one of the price-dividend ratio, the expected stock return, and the stock return volatility, determines the other two. In other words, suppose the dividend cashflows are given. Then, the dynamics of the price-dividend ratio process f completely determines the expected return μ_r and the stock return volatility σ_{rx} from equation (7). The expected stock return alone determines both the stock price (through equation (9)) and the volatility of the return (through equation (10)). Finally, specifying a process for time-varying stock volatility (σ_{rx}) determines the price of the stock (up to a multiplicative constant from equation (10)), and the expected return of the stock (from equation (9) since σ_{rx} determines f). More generally, if the dividend process can also be specified, then we can choose two out of the dividend, expected return, stochastic volatility, and price-dividend ratio, with our two choices completely determining the dynamics of the other two variables.

The relationships between prices, expected returns, and volatility outlined by Proposition 2.1 arise only through the definition of returns and the imposition of transversality. We have used no equilibrium conditions, or specified a pricing kernel, to obtain risk-return relations. Nor do we impose the full structure of an economic asset pricing model, for example, a utility function with a joint distribution of consumption and asset payoffs, to obtain relations between expected returns and volatility. The conditions (7)-(11) can be easily applied empirically because models often assume a process for one or more of μ_r, σ_{rx}, and f. Proposition 2.1 characterizes what the functional form of the expected return, stochastic volatility, or stock price must take after choosing a parameterization of only one of these variables.

There are two effects if we relax the transversality assumption in Proposition 2.1. First, the transversality Assumption 2.1 ensures that the price-dividend ratio is a function of x by Feynman-Kacs. The requirement that $P_t/D_t = f(x_t)$ is not satisfied in economies that assume geometric Brownian motion processes for the stock process (like Black and Scholes, 1973; Merton, 1973). In these economies, there is no state variable describing time-varying investment opportunities as the mean and variance are constant and the stock dividend is zero. Second, the ordinary differential equation defining the price-dividend ratio in equation (9) may now have additional terms with derivatives with respect to time t, and an additional boundary condition. This is due to the fact that when transversality does not hold, the price-dividend ratio is also potentially a function of time t.

While some empirical studies focus on matching the predictability of total returns (see, for example, Fama and French, 1988a and b; Campbell and Shiller, 1988a) and the volatility of

total returns (see, for example, Lo and MacKinlay, 1988), we often also build economic models to explain time-varying excess returns, rather than total returns. Time-varying total returns may be partially driven by stochastic risk-free rates. Proposition 2.1 involves total returns, rather than excess returns. We can handle excess returns in several ways. First, since Ang and Bekaert (2003) and Campbell and Yogo (2003) show that interest rates predict excess returns, risk-free short rates could be included as a state variable in x_t. Second, it is easy to write down a process for risk-free rates and then subtract the risk-free process from both sides of equation (7) to obtain a relation for excess returns, given stock prices and dividends. Third, since the nominal risk-free rate is known at time t over various horizons, Proposition 2.1 can be adjusted to solve for conditional excess returns. Empirically, as returns are sampled at higher frequencies, the effect of risk-free rates diminishes. For example, for daily or weekly returns, there is little difference between total and excess returns.

Finally, the relations between prices, expected returns, and volatility in Proposition 2.1 must hold in any equilibrium model. In an equilibrium model, prices, returns, and volatility are simultaneously determined after specifying a complete joint distribution of state variables, agent preferences, and technologies. In any equilibrium, the relations in Proposition 2.1 must be satisfied. Similarly, if a pricing kernel is specified, together with the complete dynamics of the state variables in the economy, the relations in Proposition 2.1 must also hold.

An advantage of the set-up of Proposition 2.1 is that many empirical specifications in finance specify models of conditional means or variances of returns (like linear dividend predictability regressions or stochastic volatility models), without specifying a full underlying equilibrium framework. In these situations, Proposition 2.1 implicitly pins down the other characteristics of returns and prices that are not explicitly assumed. In Appendix B, we show that an empirical specification of a particular conditional mean, variance or a price process does not necessarily uniquely determine a pricing kernel. However, there exists at least one (and potentially an infinite number of) pricing kernels that can support the choice of a particular expected return or volatility process. Once an empirical specification for expected returns, volatility, or prices is assumed, Proposition 2.1 completely characterizes the dynamics of the other two variables.

3 Empirical Applications

Proposition 2.1 provides practical guidance about useful empirical specifications of prices, expected returns, or volatility. For example, Campbell and Shiller (1988a), Fama and French

(1988a), Hodrick (1992), and Stambaugh (1999), among others, parameterize the expected (excess) return of the market to be a linear function of the dividend yield, or the price-dividend ratio. Proposition 2.1 shows that this places strong restrictions on the dynamics of the stock price and on the stochastic volatility process of returns. Another example is Merton (1980), who estimates several specifications of the expected return on the market as a function of the market variance, because he seeks to avoid taking a stand on the precise functional relation between expected returns and volatility. According to Proposition 2.1, once a time-series process for the market volatility is assumed, the expected return of the market is a consequence of the choice of the volatility process. Our goal in this section is to illustrate how Proposition 2.1 can be applied to various empirical models that have been specified in the literature to produce sharper predictions of risk-return trade-offs and pricing implications.

We work mainly with the assumption that dividends are IID. This assumption is only for illustrative purposes, and we choose this standard assumption because many papers work with IID dividend growth, including the textbook expositions by Campbell, Lo and MacKinlay (1997) and Cochrane (2001).[6] In Section 3.1, we show that time-varying expected returns and stock return volatility in excess of the dividend volatility are two sides of the same coin. Section 3.2 analyzes the case of the linear predictability of expected returns by dividend yields. We consider the more general case of predictable mean-reverting components of expected returns in Section 3.3. Finally, Section 3.4 investigates the implications for expected returns from various models of stochastic volatility.

3.1 IID Dividend Growth

The assumption that dividends are IID is made in many exchange-based economic models with Lucas (1978) trees. If dividend growth is IID, then time-varying price-dividend ratios can result only from time-varying expected returns. This intuition is used by Cochrane (2001) to demonstrate that small, but persistent, changes in expected returns may result in large price changes. The following corollary shows that under IID dividend growth, time-varying expected returns, price-dividend ratios, and time-varying volatility are different ways of viewing a predictable state variable driving the set of investment opportunities in the economy.

[6] Recently Ang (2002), Ang and Bekaert (2003), and Lettau and Ludvigson (2003), show that dividends are not IID but can be predicted by various state variables including interest rates, dividend yields, and consumption-asset-labor deviations from trend.

Corollary 3.1 *Suppose that dividend growth is IID, so that $\mu_d = \bar{\mu}_d$ and $\sigma_d = \bar{\sigma}_d$ are constant in equation (2). If the state variable describing the economy satisfies equation (1) and stock returns are described by the diffusion process in equation (8), where $\sigma_{rd} = \bar{\sigma}_{rd}$ is a constant, then the following statements are equivalent:*

1. *The price-dividend ratio $f = \bar{f}$ is constant.*
2. *The expected return $\mu_r = \bar{\mu}_r$ is constant.*
3. *The volatility of stock returns is the same as the volatility of dividend growth, or $\sigma_{rx} = 0$ in equation (8).*

We can interpret the term σ_{rx} in equation (8) as the excess volatility of returns that is not due to fundamental cashflow risk. Grossman and Shiller (1981) and Shiller (1981) make the argument that volatility of stock returns is too high compared to the volatility of dividend growth in an environment with constant expected returns. Cochrane (2001) provides a pedagogical discussion of this issue and claims that excess volatility is equivalent to price-dividend variability, if cashflows are not predictable. Corollary 3.1 is the mathematical statement of this claim.

3.2 Linear Dividend Yield Predictability of Returns

3.2.1 Implications of Dividend Yields Linearly Predicting Returns

In an environment with no bubbles, time-varying price-dividend ratios must reflect variation in either discount rates or cashflows, or both. To formally capture this intuition, a large number of empirical researchers have predicted future returns with price-dividend ratios or dividend yields using linear regressions.[7] Indeed, if dividend growth is IID, then the only source of time-variation for stock prices is time-varying discount rates. However, even in the simple case of IID dividend growth, the non-linearity of the present-value formula (6) implies that linear regressions may provide an overly simplistic characterization of how dividend yields capture predictable components of returns.

We show how the assumption of linear predictability of returns by dividend yields imposes strong assumptions on the dynamics of the price process. This has two main implications for the standard practice in the predictability and asset allocation literatures. While the predictability

[7] See, among others, Fama and French (1988), Campbell and Shiller (1988a and b), Hodrick (1992), Goetzmann and Jorion (1993), Stambaugh (1999), Ang and Bekaert (2003), Campbell and Yogo (2003), Engstrom (2003), Goyal and Welch (2003), Lewellen (2003), and Valkanov (2003).

literature has used a regression framework to reject the null hypothesis of no predictability of expected returns, the fact that a linear assumption of predictability implies inconsistent behavior of the price-dividend ratio shows that a linear regression cannot be the most powerful test. We explore potentially more powerful empirical specifications to pick up dividend yield predictability below. Second, a large asset allocation literature has taken the linear predictability of expected returns by dividend yields literally. Our results show that this is a very inappropriate specification for time-varying conditional means.

We can fully characterize the implicit restrictions made on prices and volatility by specifying expected returns to be a linear function of dividend yields by applying Proposition 2.1:

Corollary 3.2 *Assume that dividend growth is IID, so $\mu_d = \bar{\mu}_d$ and $\sigma_d = \bar{\sigma}_d$ are constant in equation (2).*
Case A. *Suppose that the dividend yield $D/P = f^{-1}$ linearly predicts returns in the predictive regression:*

$$dR_t = (\alpha + \beta x)dt + \bar{\sigma}_{rx} dB_t^x + \bar{\sigma}_d dB_t^d, \tag{12}$$

where the predictive instrument $x = 1/f$ is the dividend yield and $\bar{\sigma}_{rx}$ is a constant. Then, the dividend yield x follows the diffusion:

$$dx_t = \mu_x(x_t)dt + \sigma_x(x_t)dB_t^x, \tag{13}$$

where the drift $\mu_x(x)$ and diffusion $\sigma_x(x)$ are given by:

$$\begin{aligned} \mu_x(x) &= (\bar{\mu}_d + \frac{1}{2}\bar{\sigma}_d^2 + \bar{\sigma}_{rx}^2 - \alpha)x + (1-\beta)x^2 \\ \sigma_x(x) &= -\bar{\sigma}_{rx} x. \end{aligned} \tag{14}$$

Case B. *If returns are predicted by log dividend yields, $\ln(D/P)$, in the predictive regression (12), then $x = -\ln f$. The process in equation (13) then represents the dynamics of the log dividend yield, where $\mu_x(x)$ and $\sigma_x(x)$ now satisfy:*

$$\begin{aligned} \mu_x(x) &= \bar{\mu}_d + \frac{1}{2}\bar{\sigma}_d^2 + \frac{1}{2}\bar{\sigma}_{rx}^2 - \alpha - \beta x + e^x \\ \sigma_x(x) &= -\bar{\sigma}_{rx}. \end{aligned} \tag{15}$$

Corollary 3.2 implies that if expected returns are linearly predicted by dividend yields or log dividend yields, then the dividend yield or log dividend yield process cannot be a linear process. In particular, the drift of the dividend yield in equation (14) is generally not stationary

because it is quadratic.[8] Similarly, the drift of the log dividend yield in equation (15) involves both a linear and an exponential term and is also generally not stationary. Whereas dividend yield predictability of returns has often been considered in the context of linear models, such as the VAR systems of Campbell and Shiller (1988a and b) and Hodrick (1992), Corollary 3.2 demonstrates that (log) dividend yields cannot follow linear processes if expected returns are linear functions of (log) dividend yields.

The assumption of homoskedastic returns in equation (12) is not restrictive. If $\sigma_{rx}(x)$ is instead a function of x rather than being constant at $\bar{\sigma}_{rx}$, then $\mu_x(x)$ and $\sigma_x(x)$ would inherit further state-dependence from $\sigma_{rx}(x)$. The sign of σ_x for both level and log dividend yields is negative, indicating that returns and dividend yields are conditionally negatively correlated. Since the relative volatility of dividend shocks (σ_d) is small compared to the total variance of returns, the negative conditional correlation of returns and dividend yields is large in magnitude. This prediction is confirmed by empirical estimates of conditional correlations between dividend yield innovations and innovations in returns. For example, Stambaugh (1999) reports that this correlation is around -0.9 for US returns.

We calibrate the resulting non-linearities of dividend yields, or log dividend yields, by estimating the regression implied from the predictive relation (12). We use aggregate S&P500 market data at a quarterly frequency from 1935 to 2001. This is an updated dataset used by Lamont (1998) and Ang and Bekaert (2003). In Panel A of Table 1, we report summary statistics of log stock returns, both total stock returns and stock returns in excess of the risk-free rate (3-month T-bills), together with dividend growth. From Panel A, we set the mean of dividend growth at $\bar{\mu}_d = 0.05$ and dividend growth volatility at $\bar{\sigma}_d = 0.07$. The volatility of dividend growth is much smaller than the the volatility of total returns and excess returns, which are very similar, at approximately 18% per annum. This allows us to set $\bar{\sigma}_{rd}^2 = (0.18)^2 - (0.07)^2$, or $\bar{\sigma}_{rd} = 0.15$.

In Panel B of Table 1, we report linear predictability regressions of continuously compounded returns over the next year on a constant and dividend yields, expressed in levels or logs. Since the data is at a quarterly frequency, but the regression is run with a 1-year horizon on the LHS, the regression entails the use of overlapping observations that induce moving average error terms. We report Hodrick (1992) standard errors in parentheses, which Ang and Bekaert (2003) show to have good small sample properties with the correct size. Ang and

[8] Constantinides (1992) and Ahn, Dittmar and Gallant (2002) provide sufficient conditions to ensure some stationary quadratic drift processes in the context of quadratic term structure models.

Bekaert (2003) and Goyal and Welch (2003) document that dividend yield predictability declined substantially during the 1990's, so we also report results for a data sample that ends in 1990.

The coefficients in the total return regressions are similar to the regressions using excess returns. For example, over the whole sample, the coefficient for the level dividend yield is 2.97 using total returns, compared to 3.35 using excess returns. In the log dividend yield regressions, the coefficient on the log dividend yield is 0.10 (0.11) for total returns (excess returns). Hence, although we perform our calibrations for total returns, similar conditional relationships also hold for excess returns. The second line of Panel B shows that when the 1990's are removed from the sample, the magnitude of the predictive coefficients increases by approximately two, for both the level and log dividend yield regressions. To emphasize the linear predictive relationship in equation (12), we focus on calibrations using the sample without the 1990's. Nevertheless, we obtain similar qualitative patterns for the implied functional form for the drift of the price process when we calibrate parameter values using data over the whole sample.

We focus first on predicting returns with dividend yields expressed in levels, similar to Fama and French (1988a). Since the predictive regressions are at an annual frequency, the estimated coefficients in Panel B allow us to directly match α and β, since we can discretize the drift in equation (12) as approximately $(\alpha + \beta x)\Delta t$. Hence, we set $\alpha = -0.08$ and $\beta = 4.6$. Together with the calibrated values for $\bar{\mu}_d = 0.05$, $\bar{\sigma}_d = 0.07$ and $\bar{\sigma}_{rd} = 0.15$, we compute the implied drift of the level dividend yields using equation (14), which we plot in Figure 1 in the solid curved line.

Figure 1 shows that the implied drift of the dividend yield is highly non-linear. It becomes strongly mean-reverting at high levels of the dividend yield, but at low levels, the dividend yield behaves as if it is a random walk. For comparison, we plot the linear drift of an approximating Ornstein-Uhlenbeck process fitted to the level dividend yield, assuming that the dividend yield $1/f = x$ follows the process:

$$dx_t = \kappa(\theta - x)dt + \bar{\sigma}_x dB_t^x. \tag{16}$$

In data, the quarterly autocorrelation of the dividend yield is 0.96, so we calibrate κ using the relation $0.96 = \exp(-\kappa\Delta t)$, where $\Delta t = 1/4$. The unconditional mean of the dividend yield is 4.4%, so we set $\theta = 0.044$. The dashed line in Figure 1 represents the approximating linear drift $\kappa(\theta - x)$. For small movements around the unconditional mean of the dividend yield, the implied drift and the approximating AR(1) are very similar, but the discrepancy becomes very large for high or low dividend yields.

We now repeat the exercise using the log dividend yield as a predictor. In this exercise, we set the coefficients $\alpha = 0.81$ and $\beta = 0.22$ in equation (12) from the pre-1990's estimates of the log dividend yield predictability regressions in Panel B of Table 1. We represent the implied drift of the log dividend yield process using a solid line in the bottom plot in Figure 1. The dashed line represents the drift of the AR(1), or Ornstein-Uhlenbeck, process (16) fitted to the log dividend yield in data, with the calibrated values $0.94 = \exp(-\kappa/4)$ and $\theta = -3.16$. These numbers represent the quarterly autocorrelation of the log dividend yield and the mean log dividend yield, respectively.

The implied drift of the log dividend yield, if log dividend yields linearly predict returns, has fewer non-linear features than the level dividend yield case. Nevertheless, the implied log dividend yield drift is still non-linear. In particular, for high levels of the log dividend yield, the log dividend yield becomes less mean-reverting. Since dividend growth is IID, high dividend yields result from low prices, which implies that prices slowly wander back from low levels, whereas prices relatively quickly decline from high levels. For comparison, we overlay the drift of the Ornstein-Uhlenbeck approximation (16) from data to the log dividend yield. The approximating linear drift is steeper than the implied drift, indicating that the log transformation eliminates some of the non-linearity, but does not completely remove the non-linear dependence.

3.2.2 Implications of Mean-Reverting Dividend Yields

So far, we have investigated the implications for the (log) dividend yield process by assuming that returns are linearly predicted by log or level dividend yields. Now, we reverse the question. Many studies, like Stambaugh (1999), Campbell and Yogo (2003), and Lewellen (2003) specify the dividend yield, or log dividend yield, process to be the AR(1) process (16). We now show that if the (log) dividend yield is an AR(1) process, then expected returns cannot be linear in the (log) dividend yield.

Corollary 3.3 *Assume that dividend growth is IID, so $\mu_d = \bar{\mu}_d$ and $\sigma_d = \bar{\sigma}_d$ are constant in equation (2).*
Case A. *Suppose the level dividend yield $x = 1/f$, where $f = P/D$, follows the Ornstein-Uhlenbeck process in equation (16). Then, the drift $\mu_r(x)$ and diffusion $\sigma_{rx}(x)$ of the return*

process dR in equation (8) satisfy:

$$\begin{aligned} \mu_r(x) &= \kappa + \bar{\mu}_d + \frac{1}{2}\bar{\sigma}_d^2 - \frac{\kappa\theta}{x} + \frac{\bar{\sigma}_x^2}{x^2} + x \\ \sigma_{rx}(x) &= -\frac{\bar{\sigma}_x}{x} \end{aligned} \tag{17}$$

Case B. *If the log dividend yield $x = -\ln f$ follows the Ornstein-Uhlenbeck process in equation (16), then $\mu_r(x)$ and $\sigma_{rx}(x)$ satisfy:*

$$\begin{aligned} \mu_r(x) &= \bar{\mu}_d + \frac{1}{2}\bar{\sigma}_d^2 + \frac{1}{2}\bar{\sigma}_x^2 - \kappa\theta + \kappa x + e^x \\ \sigma_{rx}(x) &= -\bar{\sigma}_x \end{aligned} \tag{18}$$

Corollary 3.3 implies that if either the level or log dividend yield follows an AR(1), stock returns are a highly non-linear function of dividend yields. Hence, linear regressions to pick up predictability of returns by dividend yields may have low power. Various studies like Ang and Bekaert (2003) demonstrate the low power of various OLS estimators, but Corollary 3.3 emphasizes that it may be the OLS set-up itself that may make dividend yield predictability hard to find in the data (see also Engstrom, 2003; Menzly, Santos and Veronesi, 2004).

If level dividend yields follow an AR(1) process, Corollary 3.3 shows that returns are heteroskedastic as $\sigma_{rx} = -\bar{\sigma}_x/x$. However, when dividend yields x are high, stock volatility is low. This is the opposite to the behavior of these variables in data, since during recessions or periods of market distress, dividend yields tend to be high, while stock returns tend to be volatile. In contrast, if log dividend yields are described by an AR(1) process, Corollary 3.3 makes the strong (counter-factual) prediction that stock returns are homoskedastic.

Figure 2 graphs the expected stock return as a function of the level dividend yield (top panel), or the log dividend yield (bottom panel), in solid lines. We calibrate the level, or log, dividend yield using the implied quarterly moments from equation (16). For the level dividend yield, we match the quarterly autocorrelation, $0.96 = \exp(-\kappa/4)$; the unconditional mean $\theta = 0.044$; and the unconditional variance $0.0132^2 = \sigma_x^2/(2\kappa)$. For the log dividend yield, the corresponding numbers are $\kappa = 0.24$, $\theta = -3.16$, and $\sigma_x = 0.19$. For comparison, we graph the fitted linear regression of total stock returns at an annual horizon regressed onto a constant and the (log) dividend yield in dashed lines. The approximating linear predictive coefficients are the same as those reported in Table 1.

The top panel of Figure 2 shows a pronounced non-linear function of dividend yields and expected stock returns, caused by the interaction of reciprocal and linear functions in equation (17). In particular, low dividend yields predict extremely high conditional expected returns next

period. This result is driven by the Jensen's term, $\bar{\sigma}_x^2/x^2$, which becomes very large at low dividend yields. The large uncertainty at high prices, or low dividend yields, is similar to the strong Jensen's effects in Pastor and Veronesi (2004). This region of increasing expected returns with lower dividend yields coincides with some empirically relevant ranges for the dividend yield: expected returns increase with the level dividend yield for dividend yields lower than 2%. Note that during the late 1990's, dividend yields were very low, sometimes below 1%, and this is precisely the period where linear dividend yield predictability of returns is hard to detect (see Ang and Bekaert, 2003; Goyal and Welch, 2003).

If we fit a straight line to the solid curve in the top panel of Figure 2, the highest linear dividend yield predictability lies around the average dividend yield (4%). The slope of the drift of the expected return flattens for high dividend yields. Hence, at high dividend yield levels, there is hardly any linear predictability of expected returns by dividend yields. This also makes dividend yield predictability hard to detect by OLS regressions during recessions where dividend yields tend to be high.

In the bottom panel of Figure 2, we plot the implied drift of stock returns assuming that the log dividend yield follows an AR(1) process. The implied drift has a higher slope than the linear predictability captured by an OLS regression of stock returns on log dividend yields. Hence, directly modelling the log dividend yield process may allow sharper estimates of dividend yield predictability relations than running OLS predictive regressions on returns. Although the functional form of equation (18) exhibits non-linearities, these are not as pronounced as the implied non-linearities from assuming an AR(1) process for the level dividend yield (shown in the top panel of Figure 2) because of the small, empirically relevant range of the log dividend yield in data.

3.2.3 Implications of the Stambaugh (1999) Model

As a final example of linear dividend yield predictability, we examine the dividend growth process implied by the Stambaugh (1999) model. Stambaugh assumes that the dividend yield $x = D/P$ is an AR(1) process following equation (16), and the stock return is a linear function of the dividend yield:

$$dR_t = (\alpha + \beta x)dt + \sigma_{rx}(x)dB_t^x + \sigma_d(x)dB_t^d. \tag{19}$$

Stambaugh uses this system to assess the small sample bias in a predictive regression where the dividend yield is an endogenous regressor.

The Stambaugh model specifies two of the expected stock return, dividend yield, and stochastic volatility. Proposition 2.1 implies that only one of these variables determines the other two, once the dividend process is fixed. Hence, this system is over-parameterized if the dividends are already specified. However, if we fix the expected stock return and the dividend yield, then this implies that the dividend process must assume a very specific form. In particular, the dividend yield can follow an AR(1) process and stock return predictability can be linear in the dividend yield only if the dividend process itself is predicted by dividend yields. A further application of Proposition 2.1 implies that the drift of dD_t/D_t in equation (2) can be written as a function of the dividend yield x:

$$\alpha - \kappa + \frac{\kappa\theta}{x} - \frac{\sigma_x^2}{x^2} + (\beta - 1)x. \tag{20}$$

Hence, by assuming that dividend yields are mean-reverting, the linear dividend yield predictability of equation (19) implies that dividend yields must predict dividends.

We graph equation (20) in Figure 3, which shows that dividends are an increasing function of dividend yields. This is consistent with the positive OLS predictive estimates reported by Ang and Bekaert (2003) in the post-1952 Treasury Accord sample. This result is the opposite implied by the intuition of Campbell and Shiller (1988a and b), who claim that high dividend yields must forecast either high future returns, low future dividends, or both. According to Campbell and Shiller's reasoning, since high dividend yields should predict high returns from a large, positive β coefficient in (19), high dividend yields should predict low future dividends, if dividend yields predict dividends in the first place. This reasoning implies a downward sloping drift of dividend growth as a function of dividend yields. In contrast, Figure 3 shows a non-linear, but monotonically increasing drift of dividend growth as a function of dividend yields.

The incomplete reasoning of the Campbell and Shiller intuition is that it takes a static view of the dividend yield being a function of returns and dividends. A high dividend yield today is certainly caused by either high future returns, or low future dividend growth, or both. However, there is an implicit assumption being made that the dividend yield also does not change in the future. If dividend yields are mean-reverting, like in the Stambaugh model, then high dividend yields today mean low dividend yields tomorrow. But, low dividend yields are associated with either low returns or high dividend growth rates. The predictability equation (19) already implies that today's high dividend yield implies high expected returns. Thus, the only way that dividend yields can be low next period with high expected returns is by increasing dividends in the future. Thus, if dividend yields are mean-reverting and expected stock returns are positively linearly predicted by dividend yields, high dividend yields predict high, not low, dividend

growth rates.

3.3 Predictable Mean-Reverting Components of Returns

Mean-reversion of asset returns has been investigated by many authors, including Fama and French (1988b) and Poterba and Summers (1986), and more recently in textbook treatments by Campbell, Lo and MacKinlay (1997) and Cochrane (2001). Dividend yield predictability of stock returns is a special case of a mean-reverting expected return. Our goal is to examine the more general case of the expected stock return being a mean-reverting function of a predictable state variable. With mean-reverting returns, Proposition 2.1 places strong restrictions on the dynamics of prices and stock volatility. To characterize these restrictions, we work with IID dividend growth, so $\mu_d = \bar{\mu}_d$ and $\sigma_d = \bar{\sigma}_d$ are constant in equation (2).

We assume that stock returns have a predictable component x_t:

$$dR_t = x_t dt + \sigma_{rx}(x) dB_t^x + \bar{\sigma}_d dB_t^d, \tag{21}$$

where the single mean-reverting state variable x_t follows an AR(1) process:

$$dx_t = \kappa(\theta - x_t)dt + \bar{\sigma}_x dB_t^x. \tag{22}$$

Note that θ is the unconditional mean of the continuously compounded stock return. The volatility $\sigma_{rx}(x)$ is endogenously determined since we specify the conditional mean of the stock return.

From Proposition 2.1, the price-dividend ratio $P/D = f(x)$ satisfies the following ordinary differential equation:

$$-\kappa(\theta - x)f' + \frac{1}{2}\bar{\sigma}_x^2 f'' - \left(x - \bar{\mu}_d - \frac{1}{2}\bar{\sigma}_d^2\right) f = -1, \tag{23}$$

which represents a perpetuity security in a Vasicek (1977) model with the short rate given by $x - \bar{\mu}_d - \frac{1}{2}\bar{\sigma}_d^2$. Once f is determined, $\sigma_{rx} = \bar{\sigma}_x (\ln f)'$ is given by equation (10). Our goal is to calibrate how much stochastic volatility of returns can be attributed to mean-reverting expected returns. Since we assume dividend growth is IID, time-varying expected returns are also the only source of heteroskedasticity.

Rather than solving equation (23), we can solve for f directly using expression (6) for the price-dividend ratio. We can simplify equation (6) by writing:

$$\begin{aligned} \frac{P_t}{D_t} &= \mathrm{E}_t\left[\int_t^\infty \exp\left(-\int_t^s dx_u\right)\exp\left((\bar{\mu}_d + \frac{1}{2}\bar{\sigma}_d)s\right)\right] \\ &= \int_t^\infty \exp\left(-(\theta - \bar{\mu}_d)s + \frac{1}{2}\bar{\sigma}_d s\right)\mathrm{E}_t\left[\exp\left(-\int_t^s (x_u - \theta)du\right)\right] ds. \end{aligned}$$

It can be verified that the conditional expectation of $\exp\left(-\int_t^s (x_u - \theta)du\right)$ is just the zero-coupon bond price in a Vasicek (1977) model, with a short-rate process centered around zero, and is given by:

$$\mathrm{E}_t\left[\exp\left(-\int_t^s (x_u-\theta)du\right)\right] = \exp\left(-(x_t-\theta)\frac{1-e^{-\kappa s}}{\kappa} + \frac{\bar{\sigma}_x^2}{2\kappa^2}\left(s - 2\frac{1-e^{-\kappa s}}{\kappa} + \frac{1-e^{-2\kappa s}}{2\kappa}\right)\right).$$

Hence, we can write the price-dividend ratio as:[9]

$$\frac{P_t}{D_t} = \int_t^\infty \exp\left(-(\theta-\bar{\mu}_d)s + \frac{1}{2}\bar{\sigma}_d s - (x_t-\theta)\frac{1-e^{-\kappa s}}{\kappa} + \frac{\bar{\sigma}_x^2}{2\kappa^2}\left(s - 2\frac{1-e^{-\kappa s}}{\kappa} + \frac{1-e^{-2\kappa s}}{2\kappa}\right)\right) ds. \quad (24)$$

Equation (24) is a strictly decreasing, concave function of the expected return x. This implies that the conditional expected return is a strictly increasing, convex function of the dividend yield. We illustrate this in the top panel of Figure 4, which plots the expected return x versus the dividend yield using the values $\kappa = 0.15$, $\bar{\sigma}_x = 0.027$, $\theta = 0.125$, $\bar{\mu}_d = 0.05$, and $\bar{\sigma}_d = 0.07$. The value for θ is set from the average total continuously compounded return reported in Table 1. The values for κ and $\bar{\sigma}_x$ imply an annual autocorrelation of 0.86 and a volatility of 0.05 for the conditional expected return. We expect κ to be low (or the persistence of x to be high) because common instruments for predicting expected returns like dividend yields or interest rates are persistent variables. Our values for κ and $\bar{\sigma}_x$ are in line with the implied autocorrelation and volatilities for latent conditional expected returns reported by Brandt and Kang (2003), and Johannes and Polson (2003), among others.

The bottom panel of Figure 4 reports the implied stochastic volatility parameter $\sigma_{rx}(x)$ in equation (21). The sign of σ_{rx} is negative, implying a negative correlation between shocks to expected returns and shocks to actual returns. Note that while σ_{rx} is upward sloping, the negative sign on σ_{rx} indicates that $|\sigma_{rx}|$ is decreasing. This implies that heteroskedasticity decreases as expected returns increase. This is counter-factual, as periods of high expected returns (or market crashes) tend to coincide with periods of very high volatility. However, most notable in the bottom panel of Figure 4 is that the magnitude of σ_{rx} is extremely small, around -0.0013. Hence, the total volatility of returns is effectively the same as the volatility of dividend growth.

[9] Equation (24) can be written as $\int_t^\infty \exp(a(s) + b(s)x_t)ds$, which falls into the class of affine present value models developed by Ang and Liu (2001), Bakshi and Chen (2002), and Bekaert and Grenadier (2002).

The intuition behind this result is that large changes in f (and consequently $\ln f$) are required to produce a large amount of stochastic volatility through the relation $\sigma_{rx} = \bar{\sigma}_x(\ln f)'$ in equation (10) of Proposition 2.1. When expected returns are mean-reverting, only the terms in the sum (6) close to t change dramatically when x changes. One way that large changes in f occur from changes in x is when κ is close to zero, or expected returns are almost non-stationary. This corresponds to the case of permanent changes in expected returns. Another alternative for time-varying expected returns to re-produce the amount of stochastic volatility observed in data is for the volatility of expected returns to be the same order of magnitude as the volatility of returns. Since we observe low volatility of dividend growth, these results suggest that we may need an additional volatility factor to explain the variability of returns.

3.4 Models of Stochastic Volatility

Variances of stock returns vary over time, and their dynamics have been successfully captured by a number of models of stochastic volatility. If the dividend process is specified, Proposition 2.1 shows that the presence of stochastic volatility implies that stock returns must be predictable. Since estimating variances is easier than estimating conditional means in small samples, as Merton (1980) comments, we can use Proposition 2.1 to characterize stock return predictability by parameterizing the variance process. Given recent econometric advances in inferring the volatility process from an observed series of realized volatility (see Andersen et al., 2003), Proposition 2.1 can be used to shed light on the nature of the aggregate risk-return relation, on which there is no theoretical or empirical consensus.[10] This is an entirely different approach from the current approaches to estimating risk-return trade-offs, which use different measures of conditional volatility in predictive regressions involving the conditional mean (see, for example, Glosten, Jagannathan and Runkle, 1996).

We look at two well-known stochastic volatility models, the Gaussian model of Stein and Stein (1991) in Section 3.4.1, and the square root model of Heston (1993) in Section 3.4.2. In both cases, we assume that dividend growth is IID ($\mu_d = \bar{\mu}_d$ and $\sigma_d = \bar{\sigma}_d$ are constant in equation (2)) to focus on the relations between risk and return.

[10] For example, in two recent asset allocation applications involving stochastic volatility, Liu (2001) assumes that the Sharpe ratio is increasing in volatility, following Merton (1973), while Chacko and Viceira (2000) assume that the Sharpe ratio is a decreasing function of volatility.

3.4.1 The Stein-Stein (1991) Model

In the Stein and Stein (1991) model, the time-varying stock volatility is parameterized to be an AR(1) process. The Stein-Stein model in our set-up can be written as:

$$\begin{aligned} dR_t &= \mu_r(x_t)dt + x_t dB_t^x + \bar{\sigma}_d dB_t^d \\ dx_t &= \kappa(\theta - x_t)dt + \bar{\sigma}_x dB_t^x. \end{aligned} \tag{25}$$

The variance of the stock return is $x^2 + \bar{\sigma}_d^2$, so the stock return variance comprises a constant component $\bar{\sigma}_d^2$, from dividend growth, and a mean-reverting component x^2. Empirically, shocks to returns and shocks to volatility dynamics are strongly negatively correlated, which is termed the leverage effect, so $\bar{\sigma}_x$ is negative.

The following corollary details the implicit restrictions are on the expected return of the stock $\mu_r(\cdot)$ by assuming that the stochastic volatility process follows the Stein-Stein model:

Corollary 3.4 *Suppose that dividend growth is IID, so $\mu_d = \bar{\mu}_d$ and $\sigma_d = \bar{\sigma}_d$ are constant in equation (2). If the stock variance is given by $\sigma_{rx}(x) = x$ in equation (8), and x follows the mean-reverting process (25) according to the Stein and Stein (1991) model, then the expected stock return $\mu_r(x)$ is given by:*

$$\mu_r(x) = \bar{\mu}_d + \frac{1}{2}\bar{\sigma}_d^2 + \frac{1}{2}\bar{\sigma}_x + \frac{\kappa\theta}{\sigma_x}x + \left(\frac{1}{2} - \frac{\kappa}{\bar{\sigma}_x}\right)x^2 + C^{-1}\exp\left(-\frac{1}{2}\frac{x^2}{\bar{\sigma}_x}\right), \tag{26}$$

where C is an integration constant $C = f(0)$, where $f(0)$ is the price-dividend ratio at time $t = 0$. Furthermore, we can describe the expected return in terms of the price-dividend ratio $f = P/D$ as:

$$\mu_r(f) = \bar{\mu}_d + \frac{1}{2}\bar{\sigma}_d^2 + \frac{1}{2}\bar{\sigma}_x + \frac{\kappa\theta}{\bar{\sigma}_x}\sqrt{2\left|\bar{\sigma}_x \ln\left(\frac{f}{C}\right)\right|} + (\bar{\sigma}_x - 2\kappa)\ln\left(\frac{f}{C}\right) + \frac{1}{f}. \tag{27}$$

The expected return (26) of the stock is a combination of several functional forms. First, the expected return has a constant term, which is the case in a standard Lucas (1978) model with IID consumption growth. Second, the expected return is proportional to volatility x. This specification is implied by models of first-order risk aversion, developed by Yaari (1987) and parameterized by Epstein and Zin (1990). Third, the expected stock return is proportional to the variance x^2, which is the case in a Merton (1973) model. Finally, the last term $C^{-1}\exp(-\frac{1}{2}x^2/\bar{\sigma}_x)$ can be shown to be the dividend yield in this economy. Since the price-dividend ratio is only one component of equation (26), the Stein-Stein model predicts that expected returns may be non-linear functions of price-dividend ratios or dividend yields. We emphasize that the risk-return

trade-off (26) is not derived using an equilibrium approach. The only economic assumptions behind the risk-return trade-off is the IID dividend growth process and the no-bubble condition necessary to derive Proposition 2.1.

We illustrate the risk-return relation (26) in the top panel of Figure 5. We graph the expected return as a function of volatility x using the parameters $\theta = 0.25$, $\kappa = 8$ and $\bar{\sigma}_x = -0.2$. These parameter values are meant to be illustrative, and are consistent with stochastic volatility models estimated by Chernov and Ghysels (2002), among others. These parameter values imply that the unconditional standard deviation of volatility is 5%. We also choose $C = 24.5$, which represents the average price-dividend ratio in the data.

The risk-return trade-off in Figure 5 is highly non-linear. Expected returns initially decrease as a function of volatility for low volatility values. For volatility values higher than 15%, the expected stock return becomes a sharply increasing function of volatility. Empirically, the risk-return relation is very hard to pin down. French, Schwert and Stambaugh (1987), Bollerslev, Engle and Wooldridge (1988), Scruggs (1998) find only weak support for a positive risk-return trade-off, while Ghysels, Santa-Clara and Valkanov (2003) find a significant and positive relation. On the other hand, Campbell (1987) and Nelson (1991) find a significantly negative relations. Glosten, Jagannathan and Runkle (1993), Scruggs (1998), and Harvey (2001) report that the risk-return trade-off is negative, positive, or close to zero, depending on the specification employed. The conflicting empirical findings of the risk-return trade-off are not surprising in light of the decreasing and then increasing expected return as a function of volatility reported in Figure 5. If the Stein and Stein (1991) volatility model is a reasonable description of the data, Figure 5 suggests that an alternative appropriate risk-return empirical specification would be a bi-linear form, so that the risk-return relation can take different slopes over low or high volatilities.

To understand why the risk-return relation in the top panel of Figure 5 generally slopes upwards, consider the following intuition. The price-dividend ratio f in the Stein-Stein economy is given by $f = C^{-1} \exp(-\frac{1}{2}x^2/\bar{\sigma}_x)$, which is a decreasing function of volatility x because $\bar{\sigma}_x$ is negative (due to the leverage effect).

If x is high (and f is low), because of mean-reversion, x is likely to be lower f is likely to be higher in the next period. The return is composed of a capital gain and a dividend component. Since the dividend is IID, the higher f causes the capital gain component to be large for high enough values of x. Hence, high volatility levels correspond to high expected returns. The opposite intuition occurs for low volatility levels. When x is low, and f is high, f is more

likely to be low in the next period because of mean-reversion, and the expected capital loss causes expected returns to be low or negative for low volatility levels. The non-linearity of the variance term in (26) is responsible for the non-monotonicity of the risk-return trade-off.

The bottom panel of Figure 5 plots the expected return as a function of the dividend yield using equation (27). The expected return is an increasing function of the dividend yield. The kink at the unconditional dividend yield C^{-1} is due to the AR(1) formulation for the volatility dynamics in equation (25), allowing volatility to go negative. This assumption is equivalent to assuming a reflecting barrier at $x = 0$. An alternative model that restricts the volatility to be positive and always smooth is the Heston (1993) model, which we examine next.

3.4.2 The Heston (1993) Model

In the Heston (1993) model, the variance follows a square-root process, similar to Cox, Ingersoll and Ross (1987) that restricts the variance to be always positive. This modest change produces a large change in the behavior of the risk premium, as the following corollary shows:

Corollary 3.5 *Suppose that dividend growth is IID, so $\mu_d = \bar{\mu}_d$ and $\sigma_d = \bar{\sigma}_d$ are constant in equation (2). Suppose that returns are described by the Heston (1993) model:*

$$\begin{aligned} dR_t &= \mu_r(x_t)dt + \sqrt{x_t}dB_t^x + \bar{\sigma}_d dB_t^d \\ dx_t &= \kappa(\theta - x_t)dt + \sigma\sqrt{x_t}dB_t^x \end{aligned} \tag{28}$$

Then, the expected stock return $\mu_r(x)$ is given by:

$$\mu_r(x) = \bar{\mu}_d + \frac{1}{2}\bar{\sigma}_d^2 + \frac{\kappa\theta}{\sigma} + \left(\frac{1}{2} - \frac{\kappa}{\sigma}\right)x + C^{-1}\exp\left(-\frac{x}{\sigma}\right), \tag{29}$$

where C is an integration constant $C = f(0)$, where $f(0)$ is the price-dividend ratio at time $t = 0$. Furthermore, we can describe the expected return in terms of the price-dividend ratio $f = P/D$ as:

$$\mu_r(f) = \bar{\mu}_d + \frac{1}{2}\bar{\sigma}_d^2 + \frac{\kappa\theta}{\sigma} + \left(\frac{\sigma}{2} - \kappa\right)\ln\left(\frac{f}{C}\right) + \frac{1}{f}. \tag{30}$$

Like the Stein-Stein (1989) model, σ in equation (28) for the Heston (1993) model is negative empirically reflecting the leverage effect. We can interpret the risk-return trade-off in equation (29) to have three components: a constant term, a term linear in the variance x, and the third term $f = C^{-1}\exp(-x/\sigma)$ can be shown to be the dividend yield. Unlike the risk-return trade-off in the Stein-Stein (1989) model (see equation (26)), there is no term proportional to volatility $\sqrt{x}$.

If C^{-1} is set to be the average dividend yield, which is approximately 4.4%, then the expected stock return in equation (29) is dominated by the linear term $(\frac{1}{2} - \frac{\kappa}{\sigma})x$. Since empirical estimates of the mean-reversion of the variance, κ, are large and estimates of the magnitude of the volatility of the variance, σ, are small, the risk-return trade-off is upward sloping, for empirically relevant parameters. Figure 6 illustrates this.

In the top panel of Figure 6, we plot the expected return (29) as a function of the stock volatility $\sqrt{x}$, to be comparable to the plots of the Stein-Stein (1991) model in Figure 5. We choose the same calibrated parameters that Heston uses: $\theta = 0.01$, $\kappa = 2$, and $\sigma = -0.1$. Figure 6 shows that the risk-return trade-off from a Heston model is always positive! Mechanically, this is because the expected return in the Heston economy in equation (29) lacks a negative term proportional to volatility that enters the risk-return trade-off in the Stein-Stein model (equation (26)). The term proportional to volatility allows the expected return in the Stein-Stein solution to initially decrease, before increasing. In the Heston model, no such initial decrease can occur. The bottom panel of Figure 6 shows that the expected return is an increasing function of dividend yields, and looks remarkably similar to the corresponding picture for the Stein-Stein model in Figure 5. The expected return as a function of the dividend yield is always smooth because of the square-root process for variance in the Heston economy.

4 Conclusion

We derive conditions on expected returns, stock volatility, and price-dividend ratios that asset pricing models must satisfy. In particular, given a dividend process, specifying only one of the expected return process, the stochastic volatility process, or the price-dividend ratio process, completely determines the other two processes. For example, the dividend stream allows the volatility of stock returns to pin down the expected return. We do not need to specify a complete equilibrium model to characterize these risk-return relations, but instead derive these conditions using only the definition of returns, together with a transversality assumption.

Our conditions between risk and return are empirically relevant because many popular empirical specifications assume dynamics for one, or a combination of, expected returns, volatility, or price-dividend ratios, without considering the implicit restrictions on the dynamics of the other variables. We show that some of these implied restrictions may result in strong, sometimes internally inconsistent, dynamics. Our results point the way to future empirical work that can exploit our over-identifying conditions to create more powerful tests to investigate the risk-

return trade-off, the predictability of expected returns, the dynamics of stochastic volatility, and present value relations in a unifying framework.

Appendix

A Proof of Proposition 2.1

Equation (7) follows from a straightforward application of Ito's lemma to the definition of the return:

$$dR_t = \frac{dP_t + D_t dt}{P_t}, \tag{A-1}$$

which we rewrite as $dR_t = df_t/f_t + dD_t/D_t + 1/f_t dt$. Note that we assume that dB_t^d and dB_t^x are uncorrelated by assumption.

The definition of returns in equation (A-1) allows us to match the drift and diffusion terms in equation (7) for R_t. Hence, the price-dividend ratio f, the expected return μ_r, and the volatility terms σ_{rx} and σ_{rd} are determined by re-arranging the drift, and the dB_t^x and dB_t^d diffusion terms, respectively. If the expected return $\mu_r(\cdot)$ is determined, equation (9) defines a differential equation for f, which determines f. Once f is determined, we can solve for σ_{rx} from equation (10). If the return volatility σ_{rx} is specified, we can solve for f from equation (10) up to a multiplicative constant, and this determines the expected return μ_r in equation (9). ■

B Relation of Proposition 2.1 to Pricing Kernel Formulations

By definition, given the dividend process D_t, the price of the stock is given by:

$$P_t = \mathrm{E}_t\left[\int_t^\infty \Lambda_s D_s \, ds\right], \tag{B-1}$$

under the pricing kernel process Λ_t, together with a transversality assumption. We assume that the pricing kernel follows:

$$\frac{d\Lambda_t}{\Lambda_t} = -r_f(x_t)dt - \xi_x(x_t)dB_t^x - \xi_d(x_t)dB_t^d, \tag{B-2}$$

where $r_f(\cdot)$ is the risk-free rate process, and ξ_x and ξ_d are prices of risk corresponding to shocks to the state variable x_t and dividend growth, respectively. Using equation (B-1), we can express the price-dividend ratio as:

$$\frac{P_t}{D_t} = \mathrm{E}_t\left[\int_t^\infty \exp\left(-\int_t^s (r_f + \frac{1}{2}(\xi_x^2 + \xi_d^2))\, du + \xi_x \, dB_u^x + \xi_d \, dB_u^d\right) \times \exp\left(\int_t^s \mu_d du + \sigma_d dB_u^d\right) ds\right].$$

This can be equivalently written as:

$$\frac{P_t}{D_t} = \mathrm{E}_t^Q\left[\int_t^\infty \exp\left(-\int_t^s (r_f - \mu_d - \frac{1}{2}(\sigma_d - \xi_d)^2)\, du\right) ds\right], \tag{B-3}$$

where the Radon-Nikodym derivative defining the risk-neutral measure Q is given by:

$$\frac{dQ}{dP} = \exp\left(-\int_t^s \frac{1}{2}(\xi_x^2 + (\sigma_d - \xi_d)^2)\, du - \xi_x \, dB_u^x - (\sigma_d - \xi_d)dB_u^d\right). \tag{B-4}$$

Note that equation (B-3) is a function f of x_t.

We show how a particular choice of a return process dR_t, together with assumptions on dividends, places restrictions on the underlying pricing kernel process $d\Lambda_t$ through the following proposition:

Proposition B.1 *Suppose the state of the economy is described by x_t, which follows equation (1), and a stock is a claim to the dividends D_t that are described by equation (2). If the stock return follows equation (7) and the pricing kernel process follows equation (B-2), then the price-dividend ratio $P_t/D_t = f(x_t)$ satisfies the following relation:*

$$(\mu_x - \xi_x\sigma_x)f' + \frac{1}{2}\sigma_x^2 f'' - (r_f - \mu_d - \frac{1}{2}\sigma_d^2 + \xi_d\sigma_d)f = -1, \tag{B-5}$$

which determines the price-dividend ratio f. This implies that the expected return $\mu_r(x_t)$ and volatility $\sigma_{rx}(x_t)$ of the return are given by:

$$\begin{aligned} \mu_r &= r_f + \xi_x \sigma_x (\ln f)' + \xi_d \sigma_d, \\ \sigma_{rx} &= \sigma_x (\ln f)' \end{aligned} \tag{B-6}$$

Proof: Equation (B-5) is the standard Feynman-Kac pricing equation. Once the price-dividend ratio f is obtained from solving equation (B-5), we can derive equation (B-6) by equating terms from the drift term of dR_t and the diffusion term on dB_t^x in equation (7). ■

Proposition B.1 states that, given the dividend stream, the pricing kernel completely determines the price-dividend ratio f, the expected return of the stock μ_r, and the volatility of the stock σ_{rx}. However, if we specify the price of the stock, the expected return, or the volatility of the stock (each one being sufficient to determine the other two from Proposition 2.1), the short rate r_f, the prices of risk ξ_x and ξ_d, or the pricing kernel Λ_t are not uniquely determined. For example, suppose we specify μ_r. There are potentially infinitely many pairs of r_f and $\xi = (\xi_x, \xi_d)$ that can produce the same μ_r. For example, one (trivial) choice of ξ is $\xi = (0, 0)$ corresponding to risk neutrality, and the stock return is the same as the risk-free rate. Whereas Proposition 2.1 shows that specifying μ_r, σ_{rx}, or f completely determines the return process, the result from Proposition B.1 implies that a single choice of μ_r, σ_{rx}, or f does not necessarily determine the pricing kernel.

C Proof of Corollary 3.1

Statements (2) and (3) are equivalent from equation (10) of Proposition 2.1. Assume that $f = \bar{f}$ is a constant. Then, using equation (9), we can show that $\mu_r = \bar{f}^{-1} + \bar{\mu}_d + \frac{1}{2}\bar{\sigma}_d^2$, which is a constant. Hence (2) follows from (1). Finally, to show that (1) follows from (2), suppose that $\mu_r = \bar{\mu}_r$ is a constant. From equation (9), f satisfies the following ODE:

$$\mu_x f' + \frac{1}{2}\sigma_x^2 f'' - \left(\bar{\mu}_r - \bar{\mu}_d - \frac{1}{2}\bar{\sigma}_d^2\right) f = -1. \tag{C-1}$$

Since the term on f is constant, it follows that the price-dividend ratio $P/D = f = (\bar{\mu}_r - \bar{\mu}_d - \frac{1}{2}\bar{\sigma}_d^2)^{-1}$ is the solution. Note that this is just the Gordon formula, expressed in continuous-time. Hence, the price-dividend ratio is constant. ■

D Proof of Corollary 3.2

Using equation (10) of Proposition 2.1, we have $\bar{\sigma}_{rx} = \sigma_x(x)(\ln f)' = -1/x$, since $f = 1/x$. Rearranging, we obtain equation $\sigma_x(x) = -\bar{\sigma}_{rx} x$. From equation (9), we have:

$$\alpha + \beta x = \frac{\mu_x(x) f' + \frac{1}{2}\bar{\sigma}_{rx}^2 x^2 f'' + 1}{f} + \bar{\mu}_d + \frac{1}{2}\bar{\sigma}_d^2. \tag{D-1}$$

Substituting $f' = -1/x^2$ and $f'' = 2/x^3$, and re-arranging this expression for $\mu_x(x)$ yields equation (14). A similar derivation is used for equation (15), except we employ the transformation $x = -\ln f$, or $f = e^{-x}$. ■

E Proof of Corollary 3.3

This is a straightforward application of equation (7) of Proposition 2.1, using $f = 1/x$ for the level dividend yield and $f = \exp(-x)$ for the log dividend yield. ■

F Proof of Corollary 3.4

Using equation (10) of Proposition 2.1, we have: $x = \bar{\sigma}_x (\ln f)'$, which we can solve for the price-dividend ratio f as:

$$f = C \exp\left(\frac{1}{2}\frac{x^2}{\bar{\sigma}_x}\right), \tag{F-1}$$

where C is the integration constant $C = f(0)$. We can derive equation (26) by substituting the expression for f into equation (9) of Proposition 2.1. To derive equation (27), we use the expression for f to substitute $x^2 = 2\bar{\sigma}_x \ln(f/C)$, and $x = \sqrt{2|\bar{\sigma}_x \ln(f/C)|}$. ■

G Proof of Corollary 3.5

The proof is similar to Corollary 3.4, except now the price-dividend ratio f is given by:

$$f = C \exp\left(\frac{x}{\sigma}\right), \tag{G-1}$$

where C is the integration constant $C = f(0)$. ■

References

[1] Ahn, D. H., R. F. Dittmar, and A. R. Gallant, 2002, "Quadratic Term Structure Models: Theory and Evidence," *Review of Financial Studies*, 15, 243-288.

[2] Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys, 2003, "Modeling and Forecasting Realized Volatility," *Econometrica*, 71, 529-626.

[3] Ang, A., 2002, "Characterizing the Ability of Dividend Yields to Predict Future Dividends in Log-Linear Present Value Models," working paper, Columbia Business School.

[4] Ang, A., and G. Bekaert, 2003, "Stock Return Predictability: Is It There?" working paper, Columbia Business School.

[5] Ang, A., and J. Liu, 2001, "A General Affine Earnings Valuation Model," *Review of Accounting Studies*, 6, 397-425.

[6] Bakshi, G., and Z. Chen, 2002, "Stock Valuation in Dynamic Economies," working paper, University of Maryland.

[7] Balduzzi, P., and A. W. Lynch, 1999, "Transaction Costs and Predictability: Some Utility Cost Calculations," *Journal of Financial Economics*, 52, 47-78.

[8] Barberis, N., 2000, "Investing for the Long Run when Returns are Predictable," *Journal of Finance*, 55, 1, 225-264.

[9] Bekaert, G., and S. Grenadier, 2002, "Stock and Bond Pricing in an Affine Equilibrium," working paper, Columbia Business School.

[10] Black, F., and M. S. Scholes, 1973, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81, 637-659.

[11] Brandt, M., and Q. Kang, 2003, "On the Relationship Betweeen the Conditional Mean and Volatility of Stock Returns: A Latent VAR Approach," forthcoming *Journal of Financial Economics*.

[12] Bollerslev, T., R. F. Engle, and J. M. Wooldridge, 1988, "A Capital Asset Pricing Model with Time-Varying Covariances," *Journal of Political Economy*, 96, 116-131.

[13] Campbell, J. Y., 1993, "Intertemporal Asset Pricing without Consumption Data," *American Economic Review*, 83, 487-512.

[14] Campbell, J. Y., 1987, "Stock Returns and the Term Structure," *Journal of Financial Economics*, 18, 373-399.

[15] Campbell, J. Y., A. W. Lo, and A. C. MacKinlay, 1997, *The Econometrics of Financial Markets*, Princeton University Press, New Jersey.

[16] Campbell, J. Y., and R. J. Shiller, 1988a, "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1, 3, 195-228.

[17] Campbell, J. Y., and R. J. Shiller, 1988b, "Stock Prices, Earnings and Expected Dividends," *Journal of Finance*, 43, 3, 661-676.

[18] Campbell, J. Y., and L. M. Viceira, 1999, "Consumption and Portfolio Decisions when Expected Returns are Time Varying," *Quarterly Journal of Economics*, 114, 433-495.

[19] Campbell, J. Y., and M. Yogo, 2003, "Efficient Tests of Stock Return Predictability," working paper, Harvard University.

[20] Chacko, G., and L. Viceira, 2000, "Dynamic Consumption and Portfolio Choice with Stochastic Volatility in Incomplete Markets," working paper, Harvard University.

[21] Chernov, M., and E. Ghysels, 2002, "Towards a Unified Approach to the Joint Estimation of Objective and Risk Neutral Measures for the Purpose of Options Valuation," *Journal of Financial Economics*, 56, 407-458.

[22] Cochrane, J. H., 2001, *Asset Pricing*, Princeton University Press, New Jersey.

[23] Constantinides, G. M., 1992, "A Theory of the Nominal Term Structure of Interest Rates," *Review of Financial Studies*, 5, 531-553.

[24] Cox, J. C., J. E. Ingersoll, Jr. and S. A. Ross, 1985, "A Theory Of The Term Structure Of Interest Rates," *Econometrica*, 53, 2, 385-408.

[25] Engle, R. F., 1982, "Autoregressive Conditional Heteroscedasticity With Estimates Of The Variance Of United Kingdom Inflations," *Econometrica*, 50, 4, 987-1008.

[26] Engstrom, E., 2003, "The Conditional Relationship between Stock Market Returns and the Dividend Price Ratio," working paper, Columbia University.

[27] Epstein, L. G., and S. E. Zin, 1990, " 'First-Order' Risk Aversion and the Equity Premium Puzzle," *Journal of Monetary Economics*, 26, 387-407.

[28] Fama, E., and K. R. French, 1988a, "Dividend Yields and Expected Stock Returns," *Journal of Financial Economics*, 22, 3-26.

[29] Fama, E., and K. R. French, 1988b, "Permanent and Temporary Components of Stock Prices," *Journal of Political Economy*, 96, 246-273.

[30] French, K. R., G. W. Schwert, and R. F. Stambaugh, 1987, "Expected Stock Returns and Volatility," *Journal of Financial Economics*, 19, 3-29.

[31] Ferson, W., and C. R. Harvey, 1991, "The Variation of Economic Risk Premiums," *Journal of Political Economy*, 99, 285-315.

[32] Ghysels, E., P. Santa-Clara, and R. Valkanov, 2003, "There is a Risk-Return Tradeoff After All," working paper, UCLA.

[33] Glosten, L. R., R. Jagannathan, and D. E. Runkle, 1993, "On The Relation Between The Expected Value And The Volatility Of The Nominal Excess Return On Stocks," *Journal of Finance*, 48, 5, 1779-1801.

[34] Goetzmann, W. N., and P. Jorion, 1993, "Testing the Predictive Power of Dividend Yields," *Journal of Finance*, 48, 2, 663-679.

[35] Goyal, A., and I. Welch, 2003, "The Myth of Predictability: Does the Dividend Yield Forecast the Equity Premium?" *Management Science*, 49, 5, 639-654.

[36] Grossman, S. J., and R. J. Shiller, 1981, "The Determinants of the Variability of Stock Market Prices," *American Economic Review*, 71, 2, 222-227.

[37] Harvey, C. R., 1989, "Time-Varying Conditional Covariances in Tests of Asset Pricing Models," *Journal of Financial Economics*, 24, 289-317.

[38] Harvey, C. R., 2001, "The Specification of Conditional Expectations," *Journal of Empirical Finance*, 8, 573-638.

[39] He, H., and H. Leland, 1993, "On Equilibrium Asset Price Processes," *Review of Financial Studies*, 6, 3, 593-617.

[40] Heston, S. L., 1993, "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options," *Review of Financial Studies*, 6, 327-343.

[41] Hodrick, R. J., 1992, "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement," *Review of Financial Studies*, 5, 3, 357-386.

[42] Johannes, M., and N. Polson, 2003, "MCMC Methods for Continuous-Time Financial Econometrics," working paper, Columbia Business School.

[43] Kandel, S., and R. F. Stambaugh, 1996, "On the Predictability of Stock Returns: An Asset-Allocation Perspective," *Journal of Finance*, 51, 2, 385-424.

[44] Lamont, O., 1998, "Earnings and Expected Returns," *Journal of Finance*, 53, 5, 1563-1587.

[45] Lettau, M., and S. Ludvigson, 2003, "Expected Returns and Expected Dividend Growth," working paper, NYU.

[46] Lewellen, J., 2003, "Predicting Returns with Financial Ratios," forthcoming *Journal of Financial Economics*.

[47] Liu, J., 2001, "Dynamic Portfolio Choice and Risk Aversion," working paper, UCLA.

[48] Lo, A. W., and A. C. MacKinlay, 1988, "Stock Market Prices Do Not Follow Random Walks: Evidence From A Simple Specification Test," *Review of Financial Studies*, 1, 1, 41-66.

[49] Lucas, R. E., 1978, "Asset Prices in an Exchange Economy," *Econometrica*, 46, 6, 1429-1445.

[50] Menzly, L., J. Santos, and P. Veronesi, 2004, "Understanding Predictability," *Journal of Political Economy*, 112, 1, 1-47.

[51] Merton, R. C., 1973, "An Intertemporal Capital Asset Pricing Model," *Econometrica*, 41, 867-887.

[52] Merton, R. C., 1980, "On Estimating the Expected Return on the Market: An Exploratory Investigation," *Journal of Financial Economics*, 8, 323-361.

[53] Nelson, D. B., 1991, "Conditional Heteroskedasticity in Asset Returns: A New Approach," *Econometrica*, 59, 347-370.

[54] Pástor, L., and P. Veronesi, 2004, "Was there a NASDAQ Bubble in the Late 1990's?" working paper, University of Chicago.

[55] Poterba, J., and L. Summers, 1986, "The Persistence of Volatility and Stock Market Fluctuations," *American Economic Review*, 76, 1142-1151.

[56] Richardson, M., 1993, "Temporary Components of Stock Prices: A Skeptic's View," *Journal of Business and Economic Statistics*, 11, 2, 199-207.

[57] Scruggs, J. T., 1998, "Resolving The Puzzling Intertemporal Relation Between The Market Risk Premium And Conditional Market Variance: A Two-Factor Approach," *Journal of Finance*, 53, 2, 575-603.

[58] Shiller, R. J., 1981, "Do Stock Prices Move Too Much To Be Justified By Subsequent Changes In Dividends?," *American Economic Review*, 71, 3, 421-436.

[59] Stambaugh, R. F., 1999, "Predictive Regressions," *Journal of Financial Economics*, 54, 375-421.

[60] Stein, J. C., and E. Stein, 1991, "Stock Price Distributions with Stochastic Volatility: An Analytic Approach", *Review of Financial Studies*, 4, 727-752.

[61] Valkanov, R., 2003, "Long-Horizon Regressions: Theoretical Results and Applications," *Journal of Financial Economics*, 68, 2, 201-232.

[62] Vasicek, O., 1977, "An Equilibrium Characterization Of The Term Structure," *Journal of Financial Economics*, 5, 2, 177-188.

[63] Wachter, J., 2002, "Portfolio Consumption Decisions under Mean-Reverting Returns: An Exact Solution for Complete Markets," *Journal of Financial and Quantitative Analysis*, 1, 63-91.

[64] Yaari, M. E., 1987, "The Dual Theory of Choice under Risk," *Econometrica*, 55, 95-115.

Table 1: Dividend Yield Predictability Regressions

Panel A: Summary Statistics

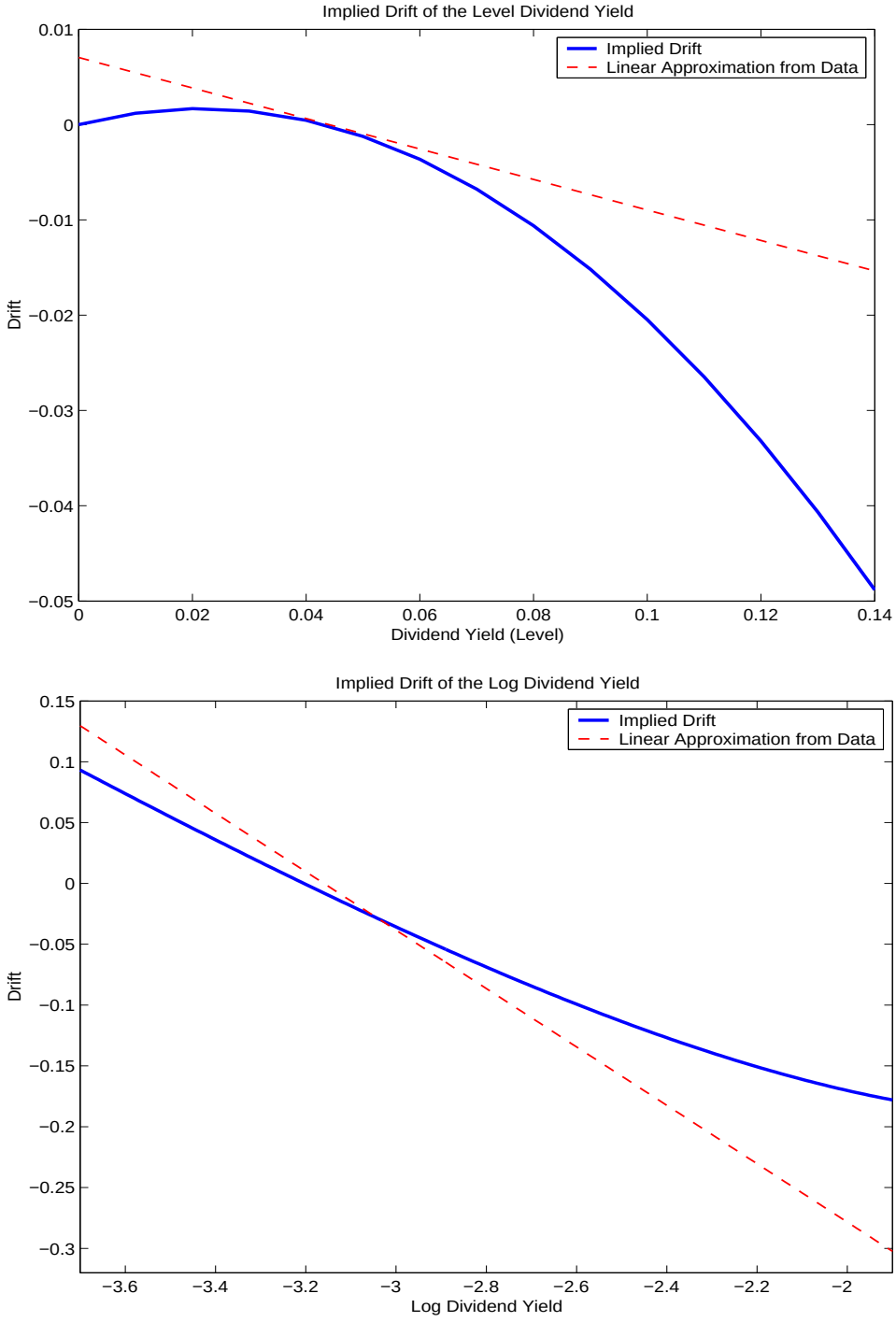
	Total Returns		Excess Returns		Dividend Growth	
	Mean	Stdev	Mean	Stdev	Mean	Stdev
1935:Q1 – 2001:Q4	0.125	0.169	0.070	0.173	0.053	0.066
1935:Q1 – 1990:Q4	0.121	0.173	0.066	0.178	0.059	0.071

Panel B: Dividend Yield Regressions

	Total Returns				Excess Returns			
	Const	Level Div Yield	Const	Log Div Yield	Const	Level Div Yield	Const	Log Div Yield
1935:Q1 – 2001:Q4	0.005	2.967	0.452	0.100	-0.067	3.351	0.439	0.113
	[0.07]	[2.03]	[2.41]	[1.76]	[-1.10]	[2.37]	[2.40]	[2.03]
1935:Q1 – 1990:Q4	-0.082	4.591	0.812	0.219	-0.157	5.041	0.819	0.238
	[-1.09]	[2.71]	[3.25]	[2.80]	[-2.08]	[3.06]	[3.34]	[3.10]

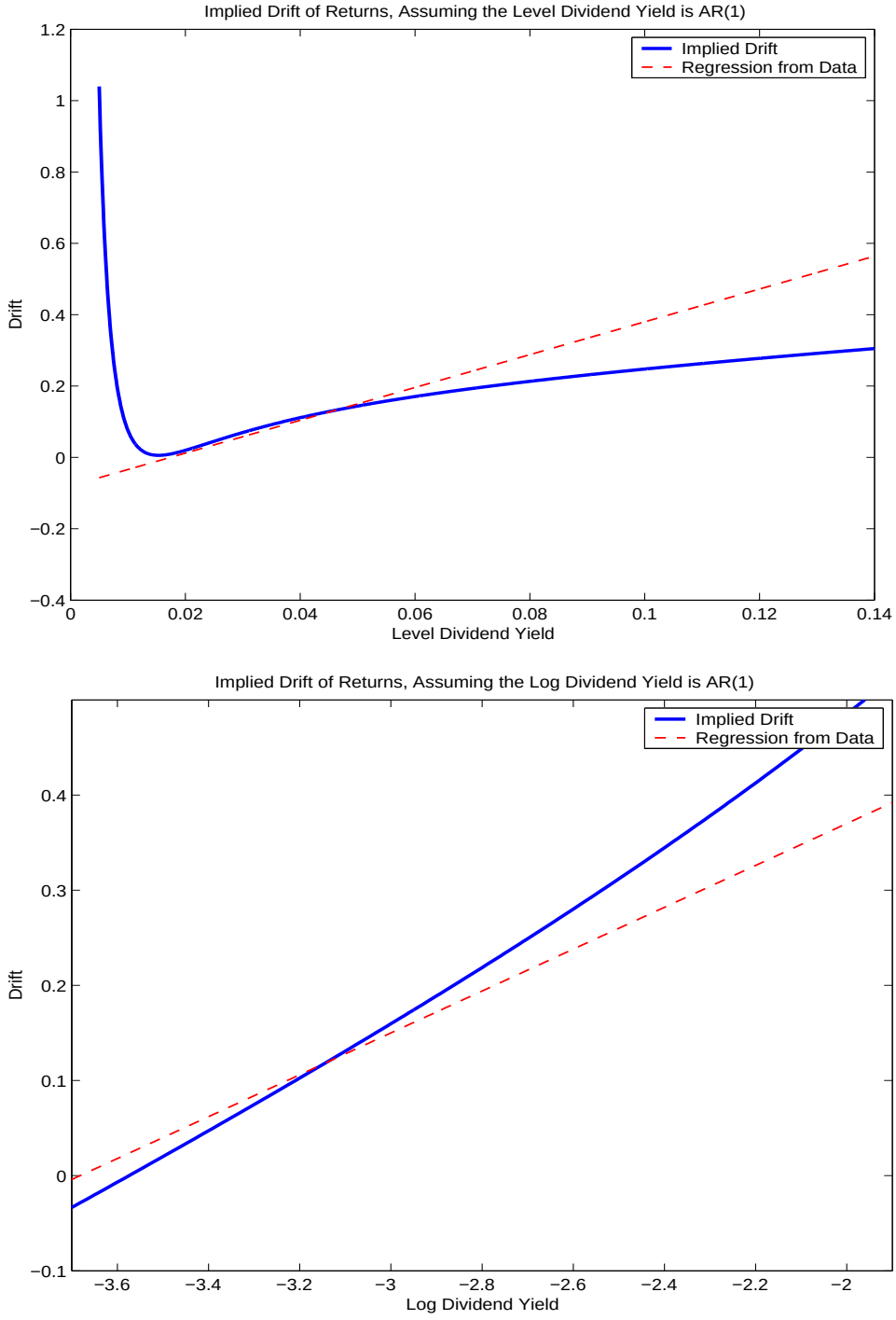
Panel A reports means and standard deviations of total returns, returns in excess of the risk-free rate (3-month T-bills), and dividend growth. All returns and growth rates are continuously compounded. Panel B reports predictive regressions of gross (or excess returns) onto a constant and a predictor. The predictor is either the dividend yield expressed in levels, or the log dividend yield. The regressions are run at an annual horizon of returns on the LHS, but at a quarterly frequency. Hodrick (1992) t-statistics are reported in parentheses. The stock data is the S&P500 from Standard and Poors and the frequency is quarterly. In Panel A, means and standard deviations for quarterly returns or growth rates have been annualized.

Figure 1: Implied Drift of the Level or Log Dividend Yield



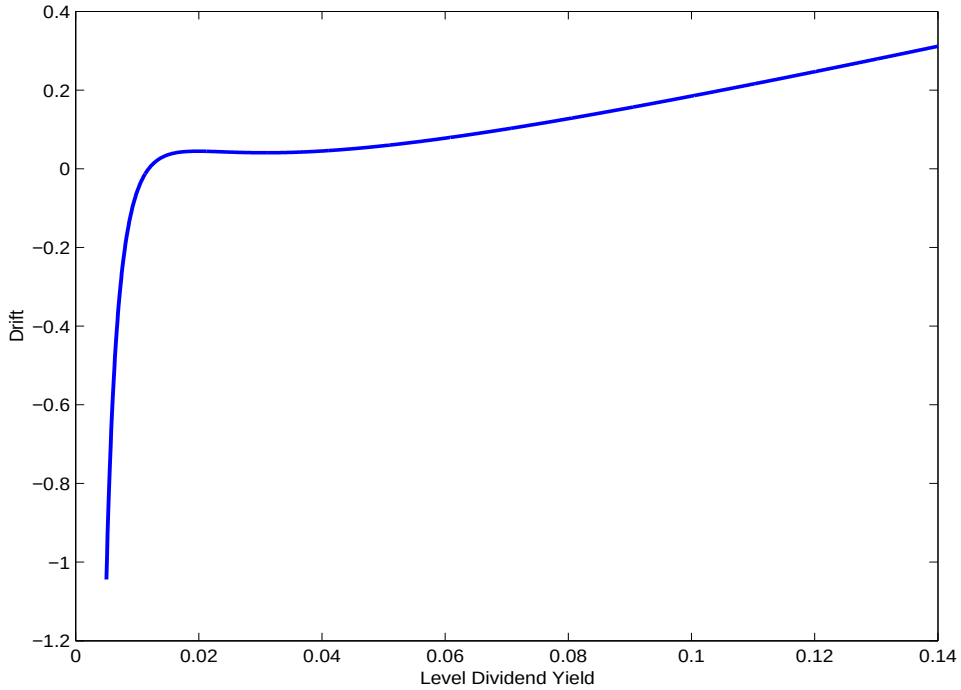
In the top panel, we graph the implied drift of the level dividend yield (equation (14)) using the calibrated parameter values $\bar{\mu}_d = 0.05$, $\bar{\sigma}_d = 0.07$, $\bar{\sigma}_{rd} = 0.15$, $\alpha = -0.08$, and $\beta = 4.6$ in the solid line. The dotted line represents the drift of an AR(1) fitted to the level dividend yield in data $\kappa(\theta - x)$, where $0.96 = \exp(-\kappa/4)$ and $\theta = 0.044$ in equation (16). In the bottom panel, we plot the implied drift of the log dividend yield (equation (15)) using $\alpha = 0.81$ and $\beta = 0.22$. The approximating AR(1) drift is produced by using $0.94 = \exp(-\kappa/4)$ and $\theta = -3.16$ in equation (16). The calibrations are done using quarterly S&P500 data from 1935 to 1990.

Figure 2: Expected Returns as a Function of Level or Log Dividend Yields



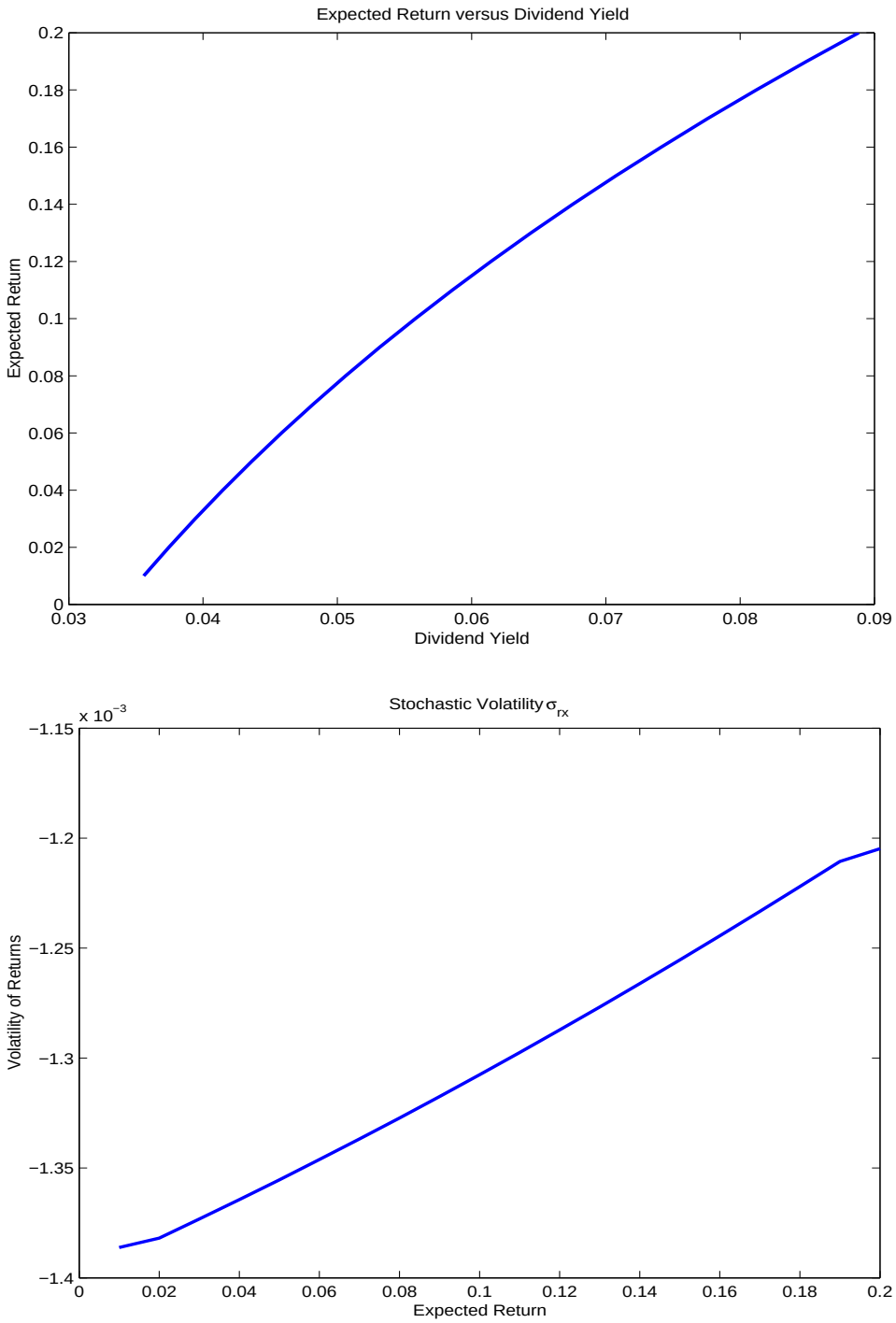
In the top panel, we graph the drift of the total stock return (equation (8)) as a function of the level dividend yield in the solid line, if the level dividend yield follows the Ornstein-Uhlenbeck process in equation (16), using the calibrated parameter values $\bar{\mu}_d = 0.05$ and $\bar{\sigma}_d = 0.07$. For the level dividend yield process, we match the quarterly autocorrelation, $0.96 = \exp(-\kappa/4)$, the long-term mean $\theta = 0.044$, and the unconditional variance $0.0132^2 = \sigma_x^2/(2\kappa)$. The dashed line represents the linear regression of total stock returns at an annual horizon regressed onto a constant and the level dividend yield, using the values from Table 1. In the bottom panel, we repeat the exercise for total stock returns as a function of the log dividend yield, with the corresponding parameters are $\kappa = 0.24$, $\theta = -3.16$, and $\sigma_x = 0.19$. The calibrations are done using quarterly S&P500 data from 1935 to 1990.

Figure 3: Drift of Dividend Growth Implied by the Stambaugh (1999) System



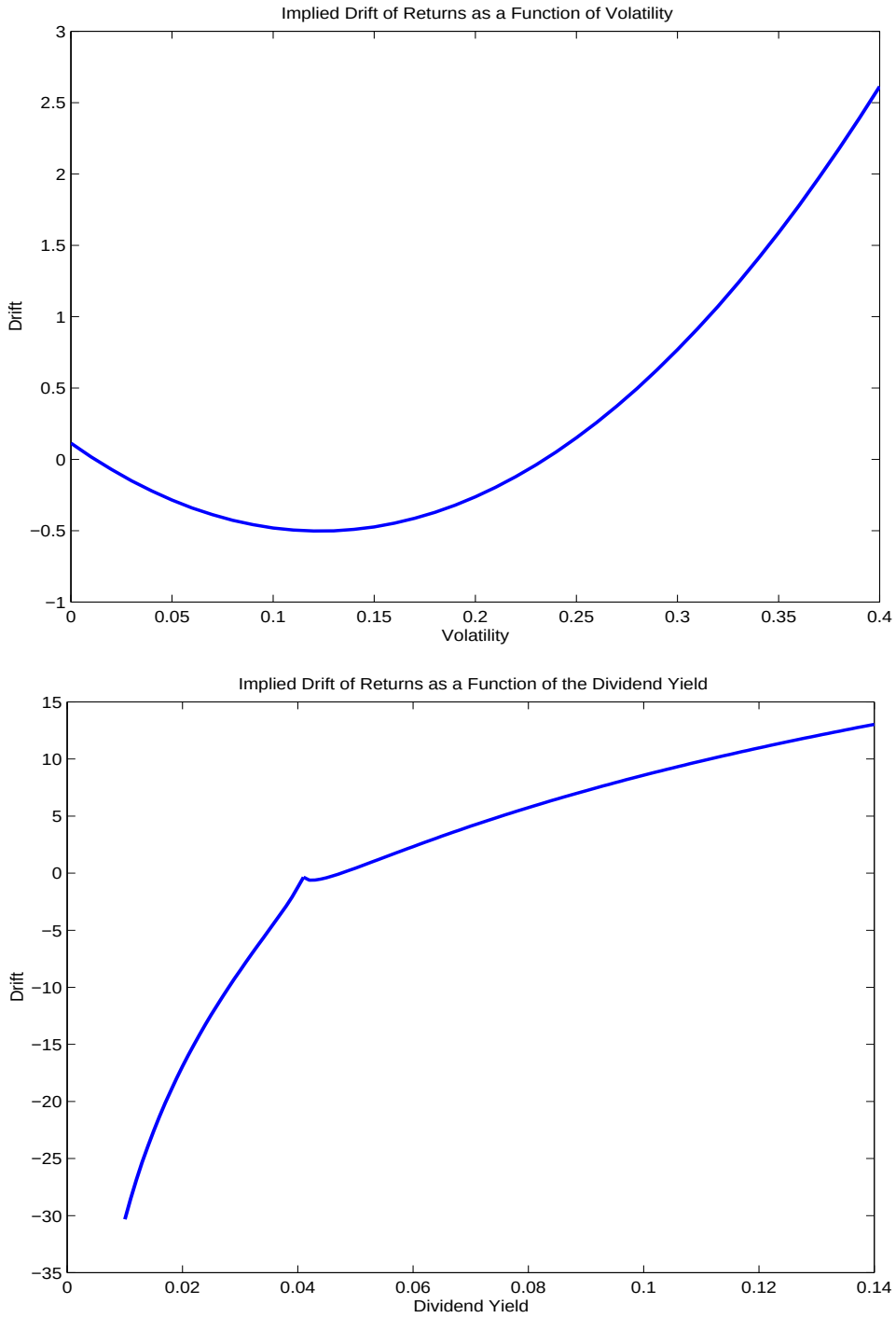
We graph the drift of dividend growth dD_t/D_t (equation (20)) from the Stambaugh (1999) system, where the level dividend yield x is mean-reverting in equation (16), and stock returns are linearly predicted by dividend yields in equation (19). We use the parameter values $\kappa = 0.16$, $\theta = 0.044$, $\sigma_x = 0.0075$, $\alpha = -0.08$ and $\beta = 4.60$. The calibrations are done using quarterly S&P500 data from 1935 to 1990.

Figure 4: Implications for Predictability and Stochastic Volatility from Mean-Reverting Expected Returns



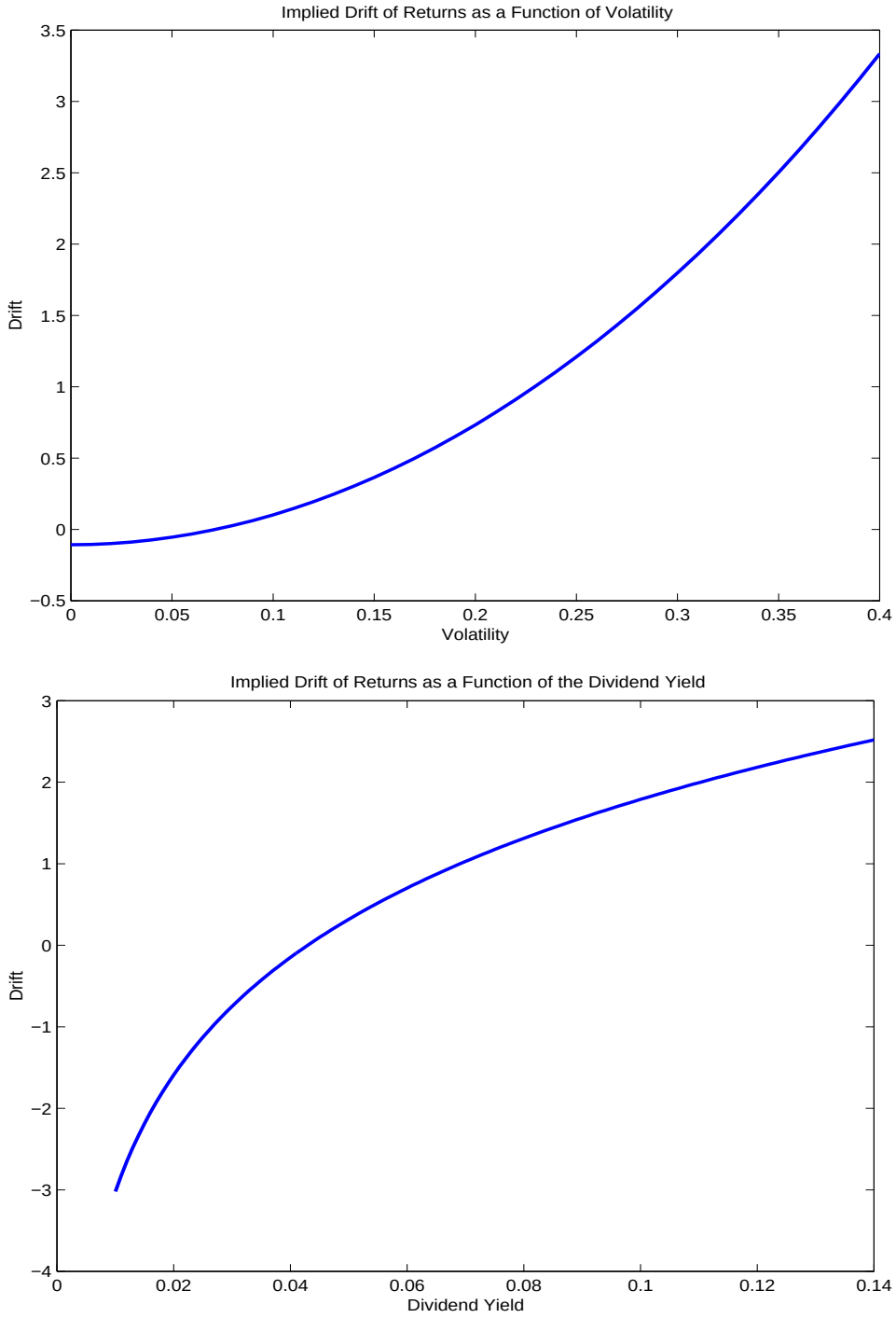
In the top panel, we graph the conditional expected return x versus the dividend yield, obtained from inverting equation (24), using the parameter values $\kappa = 0.15$, $\bar{\sigma}_x = 0.027$, $\theta = 0.125$, $\bar{\mu}_d = 0.05$, and $\bar{\sigma}_d = 0.07$. In the bottom panel, we graph the implied stochastic volatility parameter $\sigma_{rx}(\cdot)$ in equation (21) as a function of x. To produce the plots, we use quadrature to solve the price-dividend ratio in equation (24), and then numerically take derivatives of the log price-dividend ratio to compute $\sigma_{rx}(\cdot)$.

Figure 5: Implied Drift of Returns Implied by the Stein-Stein (1991) Model



In the top panel, we graph the implied drift of the stock return (equation (26)) as a function of the stock volatility implied by the Stein and Stein (1991) model in equation (25). We use the parameters $\theta = 0.25$, $\kappa = 8$, $\bar{\sigma}_x = -0.2$, $C = 24.5$, which represents the average price-dividend ratio, $\bar{\mu}_d = 0.05$, and $\bar{\sigma}_d = 0.07$. In the bottom panel, we graph the implied stock return drift as a function of the dividend yield (equation (27)).

Figure 6: Implied Drift of Returns Implied by the Heston (1993) Model



In the top panel, we graph the implied drift of the stock return (equation (29)) as a function of the stock volatility implied by the Stein and Stein (1991) model in equation (28). We use the parameters $\theta = 0.01$, $\kappa = 2$, $\sigma = -0.1$, which are the parameters used by Heston (1993), $C = 24.5$, which represents the average price-dividend ratio, $\bar{\mu}_d = 0.05$, and $\bar{\sigma}_d = 0.07$. In the bottom panel, we graph the implied stock return drift as a function of the dividend yield (equation (30)).