News, Noise, and Tests of Present Value Models

Hamidi Sahneh, Mehdi

University of Kent

4 October 2017

News, Noise, and Tests of Present Value Models

Mehdi Hamidi Sahneh [*]

October 19, 2017

I use a present value framework to explore the effects of news (or noisy information) on stock prices and drive theoretical restrictions that link price volatility to noise and information. In particular, I show that market efficiency implies that noise cannot explain more than half of price fluctuations. I propose a novel methodology to decompose stock prices into a value component, related to information about future economic fundamentals, and a noise component. The key observation is that noise by construction cannot change future economic fundamentals, but affects stock prices. The advantage of my approach is that it does not require any particular assumptions on unobserved discount rates and econometricians' information set. Consistent with the predictions of the model, my estimates show that in the prewar period noise explains up to 28% of the S&P 500 index, and 36% in the postwar period. Finally, I find that the U.S. stock market was undervalued during the 1970s and overvalued during the 1990s, but there is no evidence that the market was overvalued before the crash of 1929.

1 Introduction

There is extensive evidence that stock markets react to news about future economic or political events. Obviously, investors ability to interpret the news is constrained by its accuracy. Sometimes news provides precise information, but very often it is imprecise (say, noisy). Based on this noisy information, market participants trade with each other and prices are determined. Due to this ubiquitous role, Fischer Black (1986) dedicated his AFA presidential address to emphasize the role of noise in the financial markets and argued that noise ought to be contrasted with information. But, why stock prices are noisy? How can we measure noise in the stock market? And what are the theoretical restrictions that link price volatility to noise? This paper provides an answer to these questions through the lens of a present value model.

Financial markets react to uncertain news, including public release of information such as macroeconomic forecasts, economic surveys, news about technological innovations, and policymakers' statements. As an example, when Alan Greenspan warned about the irrational exuberance on December 5, 1996, global stock markets declined: Japan's Nikkei dropped 3.2%; Hong Kong's Hang Seng dropped 2.9%; in Germany's DAX dropped 4%, London's FTSE 100 index dropped 4%, and in the United States, the Dow Jones Industrial Average dropped 2.3%. The global markets declined because investors took Greenspan's speech as a signal that the FED might increase the real interest rate to bring down the bullish stock market.[1] In the following weeks, however, stock markets gained back all those loses when there was no sign of tightening monetary policy.

Using a simple Present Value (PV) model, this paper investigates the effects of news on the stock prices and the extent of price changes due to noise and information. Following Blanchard et al. (2013) and Forni et al. (2016), I formalize news by introducing a noisy signal about the future economic fundamentals into a standard PV model. Not dis-

---

[1]See "Greenspan Asks a Question And Global Markets Wobble", by Floyd Norris, Dec 7, 1996. The New York Times.

tinguishing noise from information, market participants react to both on impact. But ex post agents[2] observe the economic fundamentals and distinguish noise from information, and prices adjust gradually to a new level. Therefore, noise generates temporary price fluctuations.

Then, I drive the theoretical restrictions that the Efficient Market Hypothesis (EMH) imposes on the contribution of noise to price volatility.[3] In particular, I show that market efficiency implies that noise cannot explain more than half of price fluctuation. I will show that it is possible to assess this restriction by testing whether the R^2 of a forward-looking regression is bigger than or equal to 0.5. The advantage of my proposal is that it does not rely on any specific assumptions on unobserved discount rates and econometricians' information set. Consistent with the predictions of the model, my estimates show that in the prewar period noise explains up to 28% of the annual Standard and Poor (S&P) 500 index, and 36% in the postwar period.

As a byproduct of the analysis, I propose a novel methodology to decompose stock prices into a value component, related to information about future economic fundamentals, and a noise component.[4] The decomposition relies critically on the fact that noise by construction cannot change the future economic fundamentals, but affects stock prices. The decomposition can be used to date stamp and quantify the overvaluation and undervaluation periods in the stock market. I define overvaluation as periods when price is larger than value component (i.e., the noise is positive), and undervaluation as periods when price is smaller than value component (i.e., the noise is negative). The decomposition, however, is not useful in forming real-time market-timing strategies, given that it is based on the information not available to agents at the time of investment. Although one might think that it is easy to say ex post whether stock prices were overvalued or

---

[2] "market participants", "agents", and "investors" are all the same here.

[3] The EMH, articulated by Fama (1970) asserts that prices fully reflect all available information. See Malkiel (2003) for a recent survey of the literature.

[4] I borrow the value-noise terminology from Black (1986).

undervalued, the answer by no means clear at all. Take the Great Depression for example. Many economists believe that the stock market crash of 1929 followed the burst of a speculative bubble that had taken hold in the late 1920s. See, for instance, Shiller (2014). In contrast, Donaldson and Kamstra (1996) and McGrattan and Prescott (2001) argue that the market was not overvalued before the Wall Street Crash of 1929. Several robustness checks confirm these empirical finding. My estimates support the later view. On the other hand, I find that the U.S. stock market was undervalued by up to 10% during 1970s and overvalued between $10 - 20\%$ during the 1990s.

Literature Review. By touching on the Present Value (PV) models and stock price volatility, this paper connects to a voluminous literature on market efficiency. Closely related are Campbell and Shiller (1987, 1988a,b). These papers argue that stock prices are too volatile to be consistent with standard models of discount rates and to accord with the EMH. The current paper, on the other hand, finds otherwise. The natural question arises is why our conclusions are so different? The answer lies in the assumptions made by Campbell and Shiller on the econometrician's information set. They argue that although agents information set in general exceeds the information set available to the econometrician, but by including the price in the vector autoregression, the inequality becomes an equality since the full information of the market is incorporated in the price.

In my framework, this is not the case. I show that news drives a wedge between the information sets of agents and econometricians, with agents having more information than the econometrician possess. By employing conventional econometric analyses, an outside econometrician might find that stock prices are excessively volatile and wrongly reject the cross-equation restrictions implied by the underlying economic model. Closely connected in this regard is Kasa et al. (2014), who were the first to argue that information heterogeneity can account for many of the PV's apparent empirical shortcomings. They consider a model with persistent heterogeneous beliefs which can give rise to a dynamic

equilibrium solution with decision makers "forecasting the forecasts of others". An econometrician who incorrectly imposes a homogeneous beliefs equilibrium, may find that asset prices are excessively volatile and reject the null hypothesis of market efficiency.

This paper also adds to the growing macroeconomic literature on foresight and business cycle fluctuations. The key feature in the foresight literature is structural shocks with delayed effects, giving rise to equilibrium solutions with non-invertible moving Average representations. See for instance Leeper et al. (2013), Schmitt-Grohé and Uribe (2012), and Forni et al. (2014b), among others. Instead, in my framework, the key feature is a signal (or news) that provides agents with some noisy information about the future economic fundamentals, giving rise to a forward-looking representation.

This paper also builds on and extends Blanchard et al. (2013) and Forni et al. (2014a) who emphasized the key role of noisy news in business cycle fluctuations. Closely related to this paper is the recent paper by Forni et al. (2016), who put forward the idea that if rational agents cannot distinguish between noise and information, they would react to both just the same way. They also propose a novel identification strategy to measure the noise in stock prices. This paper goes beyond Blanchard et al. (2013) and Forni et al. (2016) by driving theoretical and empirical restrictions that the market efficiency imposes on the time series of the data and test those restrictions.

This paper also relates to a large and growing literature in statistical time series on forward-looking (also known as non-causal) autoregressive processes (AR). It is well-known that non-causal processes can display interesting dynamics observed in economic and financial time series.[5] However, to the best of my knowledge, no economic model has been presented that gives rise to such a representation. Therefore, a key contribution of this paper is to propose an economic model that gives rise to a non-causal equilibrium

---

[5]For instance, Gouriéroux and Zakoïan (2016) show that non-causal autoregressive processes allow for local explosive behavior and dynamics that may look like GARCH to an outside econometrician. Lanne and Saikkonen (2013) and Davis and Song (2012) propose estimation procedures for non-causal VAR models.

representation and proving that the optimal forecast of the forward-looking models is nonlinear. In doing so, this paper complements Donaldson and Kamstra (1996) who show that a nonlinear forecasting procedure can rationalize many stock price behaviors.

In sum, this paper lies at the intersection of noisy news in macro-finance literature and econometrics testing of present value models and market efficiency. The rest of the paper is organized as follows. Section II presents a simple analytical example around which the discussion is organized. Section III discusses the implications of the model for the tests of the present value models proposed by Campbell and Shiller. In section IV, I introduce a general model with time-varying discount rate and empirical results. Section V provides some concluding comments. The Appendices provide the proofs, several robustness checks, and a Monte Carlo study.

2 Analytical Example

This section presents a simple Present Value (PV) model with constant discount rates, where the econometric issues can be exposited analytically. Results and conclusions reached in this section extend to more general setups, as Section 4 discusses.

Let me begin with a simple PV model

$$p_t - d_{t-1} = \frac{\kappa}{1-\rho} + \sum_{j=0}^{\infty} \rho^j \mathbb{E}(\Delta d_{t+j} | I_t), \qquad (2.1)$$

where $p_t \equiv \log(P_t)$, $d_t \equiv \log(D_t)$, $\rho \equiv 1/\big(1 + \exp(\overline{d_t - p_t})\big)$, and $\kappa \equiv -\log(\rho) - (1-\rho)\log(1/\rho - 1)$. Moreover, $\mathbb{E}$ denotes expectation conditional on all available information to the market participants at time t, denoted by I_t.

To find the price we need to know how dividends evolve, what investors know, and how they make forecast. Here I make some assumptions. For the most part, I suppress the constants to ease the exposition. For the dividends, I consider a typical unit root

process

$$d_t = d_{t-1} + \epsilon_t, \tag{2.2}$$

where the dividend change, ϵ_t, follows an *independent and identically distributed* (*iid*) process with mean *zero* and variance σ_ϵ^2, that is $\epsilon_t \sim iid(0, \sigma_\epsilon^2)$.

The information set of investors includes all relevant information about future economic fundamentals. News is an important source of such information. To model news, I assume that at the beginning of each period, agents observe some signals about K period ahead dividend changes. But the signal is not perfect -it is contaminated with noise- as in reality news is rarely perfect. In Black's (1986) terminology, noise "makes our observations imperfect" and keeps agents from knowing future dividend changes. Following the convention that $s_{t|t+k}$ denotes the news about time $t + k$ dividend changes, we have

$$s_{t|t+k} = \epsilon_{t+k} + \nu_{t+k}, \ \ k = 0, 1, 2, \cdots, K, \tag{2.3}$$

where $\nu_t \sim iid(0, \sigma_\nu^2)$ denotes the noise which is orthogonal to ϵ_t at all leads and lags. Thus, the agents' information set at the beginning of period t (say I_t), encompass the signals as well as the history of prices and dividends. Notice that at each specific time, agents only observe the signal and cannot distinguish noise from information. But since the signal contains information, rational agents should not ignore it when they make forecast.

Before continuing, I should briefly discuss the length of the signal. The empirical tests and decompositions (discussed later) rely heavily on the assumption about K. Using annual data, I assume $K \leq 10$ to be consistent with the empirical evidence of dividend grwoth (and return) predictability which usually does not extend beyond 10 years.[6] Theoretically, we need an upper bound on K to eliminate issues with perfect signal extraction

---

[6]Moreover, in the real world, news rarely go beyond 10 years. At the time of writing this paper, President Trump is proposing "the largest tax cut in the history of the US", and as a result, the stock market is booming. He has at most two terms (totaling 8 years) to do it.

by ageents using standard filtering techniques.

How agents make forecast? I assume that agents' optimal forecast is linear. The linear optimal forecast is the standard assumption in the literature. See, for instance, Campbell and Shiller (1987, 1988a,b).[7] To see how it works, consider the optimal forecast of future dividend growth $\mathbb{E}[\Delta d_{t+j}|I_t]$, which is simply the projection of Δd_{t+j} onto the agents information set I_t

$$\mathbb{E}[\Delta d_{t+j}|I_t] = \gamma s_{t|t+j},$$
$$= \gamma(\epsilon_{t+j} + \nu_{t+j}), \ j = 0, 1, 2, \cdots, K, \tag{2.4}$$

where $\gamma = \frac{\sigma_\epsilon^2}{\sigma_\epsilon^2 + \sigma_\nu^2}$, is a particularly important constant in our analysis. Notice that, agents discount news by the parameter $\gamma \leq 1$, since $\sigma_\nu^2 \geq 0$. Intuitively, because agents do not know exactly which part of the news is noise, they are cautions and discount the news by the parameter $\gamma \leq 1$, and they discount the news more if the signal is more noisy, that is σ_ν^2 is big.

We can now find equilibrium stock prices. Substituting (2.4) into (2.1), we have

$$p_t - d_{t-1} = \sum_{j=0}^{\infty} \rho^j \mathbb{E}(\Delta d_{t+j}|I_t),$$
$$= \gamma s_{t|t} + \rho \gamma s_{t|t+1} + \cdots + \rho^K \gamma s_{t|t+K}. \tag{2.5}$$

Equation (2.5) illustrates an important point: If agents cannot distinguish between noise and information in the signal, noise affects stock prices. This effect, however, is temporary because when agents observe $\Delta d_{t+k} = \epsilon_{t+k}, \ k = 0, 1, 2, \cdots, K$, they distinguish the noise in the signal and prices adjust.

The next step is to think about how to define noise in stock prices and derive theoretical restrictions that link price price-dividend ratio fluctuations to noise. As (2.2) clearly

---

[7]For more details, the reader should consult Hamilton (1994), Ch. 4.

shows, noise cannot affect dividends, but it affects stock prices. Noise then is then defined as the component of stock prices that is orthogonal to future dividend changes.

To what extent price changes can be attributed to noise? From (2.5) we see that prices can be decomposed into two components, a Value Component (VC) defined as

$$\text{VC} = \gamma\epsilon_t + \rho\gamma\epsilon_{t+1} + \cdots + \rho^K\gamma\epsilon_{t+K},$$

associated with the information about future dividends growth, and a Noise Component (NC) associated with the noise

$$\text{NC} = \gamma\nu_t + \rho\gamma\nu_{t+1} + \cdots + \rho^K\gamma\nu_{t+K}.$$

The size of the NC is determined by σ_ν^2. This is because noise variance affects not only the volatility of the noise shocks, but also the inference problem of investors. In particular, when σ_ν^2 is either too small or too large, noise generates small price fluctuations. In the first case, news is precise. In the second case, news is very imprecise and investors disregard them in their inference.[8]

To see this point, lets consider two interesting limit cases. First consider the case where $\sigma_\nu^2 \to 0$. Then one can show that there is no noise in stock prices as the variance of the NC, $\frac{1-\rho^{2(K+1)}}{1-\rho^2}\frac{\sigma_\epsilon^4\sigma_\nu^2}{(\sigma_\epsilon^2+\sigma_\nu^2)^2}$, vanishes. The intuitive explanation for this result is that when $\sigma_\nu^2 \to 0$, the signal is precise and agents know the future dividend changes. As a result, there is no noise in stock prices. Interestingly, the NC disappears even in the opposite case, when $\sigma_\nu^2 \to \infty$. Intuitively, when σ_ν^2 is very large, the signal is not informative, so that rational agents should ignore it.

The variance of NC is largest for intermediate values of noise variance. For a given value of ρ and σ_ϵ^2 the variance of the NC is maximized at $\sigma_\nu^2 = \sigma_\epsilon^2$, and explains half

---

[8]See also Lorenzoni (2009), Forni et al. (2014a) and Forni et al. (2016).

of price fluctuations. Therefore, market efficiency implies that NC can not explain more than 50% of price fluctuations.

3   Campbell-Shiller Tests of Present Value Models

As noted in the introduction, Campbell and Shiller (1987, 1988a,b) reject the restrictions that PV models with standard discount rates impose on the data. The PV model with constant discount rate presented in Section 2 is useful to show why Campbell-Shiller (henceforth CS) procedure is based on an assumption which is hard to justify. I use this simple environment in order to make my point, but it is straightforward to extend the specification to allow for time-varying discount rates. Since this does not add any new insight, I do not entertain this idea in this section.

3.1   Volatility Bounds

Volatility tests of market efficiency examine whether stock price movements is justified by subsequent changes in discounted future dividend growth. The basic idea of the CS procedure is most easily explained in the context of the model (2.1). More specifically, CS propose to test the restriction

$$\frac{\mathrm{var}(\delta_t)}{\mathrm{var}(\delta_t')} = 1, \tag{3.1}$$

where $\delta_t \equiv p_t - d_{t-1}$ is stationary, and

$$\delta_t' \equiv \mathbb{E}[\sum_{j=0}^{\infty} \rho^j \Delta d_{t+j} | H_t],$$

is the unrestricted VAR forecast of the future changes in dividends, conditional on the econometrician's information set, denoted by H_t. Using the US stock market data, CS find

10

that $\frac{\text{var}(\delta_t)}{\text{var}(\delta_t')}$ is considerably larger than unity and conclude that stock prices are excessively volatile.

The key assumptions underlying the CS testing procedure can be summarized as follows:

(CS1) The transversality condition is satisfied, i.e., $\lim_{i \to \infty} \rho^i \delta_{t+i} = 0$.

(CS2) An assumption about the unobserved discount rates.

(CS3) Both agents and investors use linear projection when they make forecasts.

(CS4) The agents and econometrician's information sets are equivalent.

Assumptions (CS1)-(CS3) have been discussed extensively in the literature. The rational bubble literature, including Blanchard and Watson (1982) and Tirole (1985), argue that the variance bounds may be violated due to the violation of the transversality condition.[9] Assumption (CS2) has elicited the greater attention in the Macro-Finance literature.[10] Since the discount rate is not directly observable, one must use some proxy for the discount rate to test the restriction (3.1). Campbell and Shiller (1987) made the assumption that the real discount rate is constant. In a subsequent work, Campbell and Shiller (1988a) find similar results when dividends are discounted by a consumption-based discount rate, or by an interest rate. Therefore, CS testing procedure is a joint test of market efficiency and a discount rate model, and a rejection should not be interpreted as evidence against market efficiency and rationality. By relaxing (CS3), Donaldson and Kamstra (1996) argue that a nonlinear dividend forecasting model can rationalize the stock price behavior during the 1930s, but they do not provide any formal argument why a nonlinear forecast performs better.

---

[9]There is a vast literature providing evidence against explosive bubbles. See for instance Diba and Grossman (1988), West (1988) and Giglio et al. (2016).

[10]See Cochrane (2011) for a comprehensive review of the litereture.

Assumption (CS4), however, has not received much attention, which is surprising because the restriction (3.1) is completely derived under this assumption. To see this, lets assume that (CS1)-(CS3) are all satisfied. Then, from elementary statistics we know that

$$\frac{\text{var}(\delta_t)}{\text{var}(\delta_t')} \geq 1, \text{if } I_t \subseteq H_t, \tag{3.2}$$

$$\frac{\text{var}(\delta_t)}{\text{var}(\delta_t')} = 1, \text{if } H_t = I_t, \tag{3.3}$$

$$\frac{\text{var}(\delta_t)}{\text{var}(\delta_t')} \leq 1, \text{if } H_t \subseteq I_t. \tag{3.4}$$

Intuitively, whoever has more information makes better forecasts. The natural and important question is which assumption is more realistic. (3.2) is not a realistic assumption, as agents' information set can have a variety of sources, not all of them observed by econometricians. CS acknowledge that the agents information set in general exceeds the econometricians information set, but argue that:[11]

> "*even though we do not observe everything that market participants do, we do observe the log dividend-price ratio, and that variable summarizes the market's relevant information.*" Campbell and Shiller (1988a, p. 207)

Interestingly, this is not the case here. To see how news about future economic fundamentals breaks down (CS4), consider the limit case when agents know exactly all future dividend changes. Suppressing the constant, we can rewrite (2.5) as

$$\delta_t = \epsilon_t + \rho\epsilon_{t+1} + \rho^2\epsilon_{t+2} + \cdots. \tag{3.5}$$

To see how the information sets of investors and econometricians relate to each other, first consider the perspective of a rational agent. The agents' information set is the linear space generated by current and future dividend changes ϵ_t, which is equivalent to the

---

[11]Similar arguments can be found in Campbell and Shiller (1987, p. 1064) and Campbell and Shiller (1988b, p. 669).

linear space generated by current and future δ_t. To see this point, note that (3.5) is invertible in current and future δ_t

$$(1 - \rho F)\delta_t = \epsilon_t, \tag{3.6}$$

where F is the forward operator (*i.e.*, $FX_t = X_{t+1}$). In the time series literature, (3.6) is usually written as

$$(1 - \rho^{-1}L)\delta_t = \epsilon_t, \tag{3.7}$$

where L is the backward operator (*i.e.*, $LX_t = X_{t-1}$). Since the AR polynomial vanishes for $|\rho| < 1$, (3.7) has a stationary forward-looking representation.[12] Therefore, recovering ϵ_t requires not only current and past but also future values of δ_t.

Now consider the perspective of an outside econometrician. By estimating a conventional VAR model, the econometrician "in effect" estimates

$$(1 - \rho L)\delta_t = \epsilon_t^*, \tag{3.8}$$

where $\epsilon_t^* = \frac{1-\rho L}{1-\rho^{-1}L}\epsilon_t$. This representation is obtained by applying the Blaschke factor $\frac{1-\rho L}{1-\rho^{-1}L}$ to both sides of (3.7). See Hansen et al. (1981) and Lippi and Reichlin (1994) for more details. From (3.8) we see that the information set of the econometrician is the linear space generated from current and past values of δ_t, which from the Wold representation is equivalent to the linear space generated by current and past values of ϵ_t^*. But since the Blaschke factor is a two sided filter, recovering ϵ_t requires not only current and past but also future values of ϵ_t^*. Therefore, the linear space generated by current and past values of ϵ_t is at least as large as the linear space generated by current and past ϵ_t^* (or equivalently δ_t). Thus, $H_t \subseteq I_t$.

---

[12]But since $|\rho^{-1}| > 1$, the backward-looking representation is explosive.

Augmenting the econometrician's information set by dividends does not solve the information misalignment. To see this point, suppose the econometrician estimates a VAR that includes the first difference of dividends Δd_t and δ_t. Then, the equilibrium solution is given by

$$\begin{bmatrix} 0 & 1 \\ \frac{1-\rho^{-1}L}{\gamma} & -1 \end{bmatrix} \begin{bmatrix} \delta_t \\ \Delta d_t \end{bmatrix} = \begin{bmatrix} \epsilon_t \\ \nu_t \end{bmatrix},$$

$$\Phi(L)Z_t = \eta_t, \tag{3.9}$$

which is forward-looking since $\det \Phi(z) = \frac{1-\rho^{-1}z}{\gamma}$ vanishes for $z = \rho$. An outside econometrician who estimates a conventional VAR model, effectively estimates the following backward-looking model

$$B(L)W\Phi(L)Z_t = \eta_t^*, \tag{3.10}$$

where

$$B(L) = \begin{bmatrix} \frac{\rho^{-1}-L}{1-\rho^{-1}L} & 0 \\ 0 & 1 \end{bmatrix} \qquad W = \begin{bmatrix} \frac{-1}{\sqrt{1+\rho^2}} & \frac{\rho}{\sqrt{1+\rho^2}} \\ \frac{\rho}{\sqrt{1+\rho^2}} & \frac{1}{\sqrt{1+\rho^2}} \end{bmatrix},$$

and $\eta_t^* = B(L)W\eta_t$. Representation (3.10) is obtained by finding Blaschke matrix $B(L)$ and orthonormal matrix W, such that $B(L)B(L^{-1})' = I$ and $WW' = I$, to flip the zero of $\det \Phi(L)$ outside of the unit circle, but do not alter the covariance generating function of Z_t.

From (3.1) it follows that the agents' information set is the linear space generated from current and past values of η_t, and the econometricians' information set is the linear space generated by current and past values of ϵ_t^*. But since the Blaschke factor is a two sided filter, recovering η_t requires not only current and past but also future values of η_t^*. Therefore, the space generated by η_t is not contained in the space spanned by current and past values of η_t^* (or equivalently Z_t). As a result, $H_t \subseteq I_t$, which implies that the correct

variance bound is given by the inequality (3.4). This is summarized in the following proposition:

Proposition 3.1: Let $\delta_t \equiv \mathbb{E}[\sum_{j=0}^{\infty} \rho^j \Delta d_{t+j} | I_t]$, $\delta_t' \equiv \mathbb{E}[\sum_{j=0}^{\infty} \rho^j \Delta d_{t+j} | H_t]$, and $H_t \subseteq I_t$. Then the present value relation implies that

$$\text{var}(\delta_t') \leq \text{var}(\delta_t)$$

3.2   Cross-Equation Restrictions

A distinguishing characteristic of rational expectations hypothesis is that the parameters describing the stochastic environment that the agents confront appears in the equilibrium solution. CS propose a convenient method for characterizing the cross-equation restrictions that the PV relation imposes on the data. To see how this approach works, consider the PV model

$$\delta_t = \sum_{j=0}^{\infty} \rho^j \mathbb{E}(\Delta d_{t+j} | I_t). \tag{3.11}$$

Testing (3.11) is not an easy task, because $\mathbb{E}(\Delta d_{t+j} | I_t)$ is unobservable. To get around this problem, CS propose to substitute it with the econometrician's forecast obtained from the following VAR model

$$\underbrace{\begin{bmatrix} \delta_t \\ \Delta d_t \end{bmatrix}}_{Z_t} = \underbrace{\begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}}_{\Phi} \underbrace{\begin{bmatrix} \delta_{t-1} \\ \Delta d_{t-1} \end{bmatrix}}_{Z_{t-1}} + \underbrace{\begin{bmatrix} \epsilon_t \\ \nu_t \end{bmatrix}}_{\eta_t}.$$

The vector Z_t has the useful property that to obtain it's k-ahead period forecast, we simply multiply Z_t by the kth power of the matrix Φ:

$$\mathbb{E}[Z_{t+j} | H_t] = \Phi^{j+1} Z_t, \text{ for } j = 0, 1, \cdots, K. \tag{3.12}$$

CS derive the cross-equation restrictions by projecting (3.11) onto the econometrician's information set H_t, which gives

$$\mathbf{e1}'Z_t = \sum_{j=0}^{\infty} \rho^j \mathbf{e2}' \Phi^{j+1} Z_t \tag{3.13}$$

where $\mathbf{e1}' = [1 \ 0]$ and $\mathbf{e2}' = [0 \ 1]$. It follows that (3.11) and (3.13) are the same, *if and only if* (3.12) is satisfied (i.e., the optimal forecast is linear). Since equation (3.13) holds for all realizations of Z_t, we have that

$$\mathbf{e1}' = \sum_{j=0}^{\infty} \rho^j \mathbf{e2}' \Phi^{j+1} = \mathbf{e2}' \Phi (\mathbf{I} - \rho\Phi)^{-1} \tag{3.14}$$

where the second equality follows by evaluating the infinite sum, which must converge because the elements of Z_t are stationary. Postmultiplying both sides of (3.14) by $(\mathbf{I} - \beta\Phi)$, we have

$$\mathbf{e1}'(\mathbf{I} - \rho\Phi) - \mathbf{e2}'\Phi = 0, \tag{3.15}$$

which can be tested using a Wald statistics.

These restrictions are frequently rejected by the data, which has been interpreted as evidence against standard discount rate models or sometimes against market efficiency and rationality. However, note that these restrictions are derived by imposing the arbitrary assumption (CS4). But if stock prices are forward-looking, the optimal forecast in general is nonlinear. This is summarized in the following proposition.

Assumption 3.2. $\eta_t \equiv (\epsilon_t, \nu_t)'$ are *iid*, strictly stationery process with a non-Gaussian distribution such that $(a + 1)$st moment finite with $(a + 1)$st cumulant nonzero for some $a \geq 2$.

Proposition 3.3: Under assumption 1, the best predictor of a non-causal model is non-linear. Therefore, standard cross-equation restriction tests that exclude forward-looking

16

representations, can produce spurious rejections.

This proposition also provides the theoretical link between (CS3) and (CS4). If agents and econometricians have the same information and agents use linear projections to make optimal forecast, then the econometricin's optimal forecast is also linear, i.e.,

$$\mathbb{L}\big[\mathbb{L}[Z_t|I_t]|H_t\big] = \mathbb{L}[Z_t|H_t] = \mathbb{E}[Z_t|H_t],$$

where the first equality is always true by the law of iterated projections and the second equality follows from the fact that conditional expectation is the best possible forecast (in MSE sense).[13] More important, the second equality asserts that the optimality is preserved under (CS4). But, if stock prices are forward-looking, (CS4) will be violated and the optimality disappears

$$\mathbb{L}\big[\mathbb{L}[Z_t|I_t]|H_t\big] = \mathbb{L}[Z_t|H_t] \neq \mathbb{E}[Z_t|H_t].$$

In other words, if $H_t \subseteq I_t$, the econometricians' optimal forecast is in general nonlinear, even if agents' optimal forecast is linear. This result also complements Donaldson and Kamstra (1996), who show that nonlinear forecast outperforms the linear forecast obtained from an unrestricted VAR.

Proposition 3.1 and 3.3 reveal the importance of the assumption (CS4), but do not show whether it matters in practice. In Appendix E, I use the simple model of Section 2 as data generating process to quantify the inferential errors an econometrician might make by imposing wrong assumptions on the econometricinas' information set. Briefly, the Monte Carlo exercise clearly shows that Campbell and Shiller procedure does not control for the discrepancy between the econometricians' and agents' information sets, leading to too frequent rejection of the correct null hypothesis.

---

[13]See Hamilton (1994, Ch. 4).

4 Econometrics Analysis of News

So far, I have assumed that expected stock returns are constant. Although this assumption is analytically convenient, the basic notion of market efficiency does not require constant discount rate assumption. See LeRoy (1973), Rubinstein (1976), and Lucas Jr (1978), among others. I now turn to the full model with time-varying discount rates.

Following the convention that logs of variables are denoted by the lowercase letters, the log-linear approximate relation between prices, dividends, and returns proposed by Campbell and Shiller (1988a) gives

$$\delta_t = \frac{\kappa}{1 - \rho} + \mathbb{E}\Big[\sum_{j=0}^{\infty} \rho^j \big(\Delta d_{t+j} - r_{t+j}\big)|I_t\Big], \tag{4.1}$$

where $\delta_t \equiv p_t - d_{t-1}$ is the log price-dividend ratio, $\rho = 1/\big(1 + \exp(-\overline{\delta})\big)$ and $\kappa = -\log(\rho) - (1 - \rho)\log(1/\rho - 1)$ are a function of the long-run average log dividend-price ratio $\overline{\delta}$.

To complete the description of the model, we must specify a process for dividends, discount rates, and agents information set. Throughout, I assume that agents' optimal forecast is linear.

(A1) Dividends. Dividends follow a unit root process

$$d_t = d_{t-1} + \epsilon_t, \tag{4.2}$$

for $t = 1, 2, \cdots$, and $\epsilon_t \sim iid(0, \sigma_\epsilon^2)$.

(A2) Discount Rates. I assume that discount rates evolve according to a typical random walk

$$r_t = r + \varepsilon_t, \tag{4.3}$$

for $t = 1, 2, \cdots$, and $\varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$.

(A3) Agents' Information Set. At the beginning of each period, agents observe some noisy signals about K_1 period ahead dividend changes

$$s_{t|t+k}^d = \epsilon_{t+k} + \nu_{t+k}, \ k = 0, 1, 2, \cdots, K_1, \tag{4.4}$$

where $\nu_t \sim iid(0, \sigma_\nu^2)$ denotes the noise which is orthogonal to ϵ_t at all leads and lags. Moreover, agents also observe some noisy signals about K_2 period ahead returns

$$s_{t|t+k}^r = \varepsilon_{t+k} + \upsilon_{t+k}, \ k = 0, 1, 2, \cdots, K_2, \tag{4.5}$$

where $\upsilon_t \sim iid(0, \sigma_\upsilon^2)$ denotes the noise which is orthogonal to ε_t at all leads and lags. In general $K_1 \neq K_2$, but for the sake of convenience I set $K_1 = K_2 = K$. Thus, the agents' information set encompasses the signals, as well as the history of prices and dividends.

Although to some readers these assumptions may look special, they are not. It is widely accepted that neither dividend changes nor returns are predictable based on their own past, but both are predictable based on price-dividend ratio.[14] Assumptions (A1)-(A3) formalize these stylized facts, and should also dispel the popular myth that if returns and dividend changes follow a random walk process they are unpredictable. We also know that signals must exists, although we do not know the exact form of them: If signals do not exist, neither dividend changes nor returns are predictable, implying that the price-dividend ratio must be a constant, which obviously is not the case.

Using (A1)-(A3) and the PV model (4.1), we can find equilibrium stock prices

$$\delta_t = \gamma^r s_{t|t}^r + \gamma^d s_{t|t}^d + \rho(\gamma^r s_{t|t+1}^r + \gamma^d s_{t|t+1}^d) + \cdots + \rho^K(\gamma^r s_{t|t+K}^r + \gamma^d s_{t|t+K}^d), \tag{4.6}$$

---

[14]Return and dividend predictability is a controversial topic in financial econometrics. See, for instance, Cochrane (2008), Welch and Goyal (2007), Boudoukh et al. (2006), and Campbell and Thompson (2008). Koijen and Van Nieuwerburgh (2011) provides a recent review of the literature.

where $\gamma^r = \frac{\sigma_\epsilon^2}{\sigma_\epsilon^2 + \sigma_\nu^2}$ and $\gamma^d = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_\upsilon^2}$. From (4.6) it follows that stock prices can be decomposed into a Value Component

$$\text{VC} = \gamma^r \epsilon_t + \gamma^d \varepsilon_t^d + \rho(\gamma^r \epsilon_{t+1}^r + \gamma^d \varepsilon_{t+1}^d) + \cdots + \rho^K(\gamma^r \epsilon_{t+K}^r + \gamma^d \varepsilon_{t+K}^d), \qquad (4.7)$$

associated with the future dividends growth and returns, and a Noise Component

$$\text{NC} = \gamma^r \nu_t + \gamma^d \upsilon_t^d + \rho(\gamma^r \nu_{t+1}^r + \gamma^d \upsilon_{t+1}^d) + \cdots + \rho^K(\gamma^r \nu_{t+K}^r + \gamma^d \upsilon_{t+K}^d). \qquad (4.8)$$

How to measure noise? From (4.6) we see that noise affects stock prices, but noise can not affect future dividend changes or returns. Therefore,

Definition 4.1: Noise is the component of stock prices that is orthogonal to future economic fundamentals.

This definition provides the theoretical framework to measure the noise in the stock market. In the linear regression of δ_t on x_t

$$\delta_t = \gamma_0 + \gamma_1 x_t + u_t, \qquad (4.9)$$

where $x_t \equiv \sum_{j=0}^K \rho^j (\Delta d_{t+j} - r_{t+j})$, the residuals u_t is the component of δ_t that is by construction orthogonal to x_t. Therefore, in the above forward-looking regression, u_t approximates the NC in stock prices, and the fitted value of the regression, denoted by $\hat{\delta}_t$, approximates the VC.[15]

Several comments are in order. First, both the VC and the NC are included in the

---

[15]Note that regression (4.9) is different from the long-run predictability regressions

$$\sum_{j=1}^K y_{t+j} = \gamma_0 + \gamma_1 \delta_t + u_t,$$

where $y_t = r_t$ or Δd_t. The literature almost exclusively focused on the dividend and return predictability. Our focus is instead on the comovement of the price-dividend ratio with future economic fundamentals.

fundamental value, i.e., the right hand side of (4.1). Therefore, although prices deviate from VC due to noise, this cannot be interpreted as a bubble.[16] Second, $\hat{\delta}_t$ is based on the information that is not available to econometricians and agents at time t, and therefore cannot be used for market timing. Finally, to form the right hand side of the regression (4.9), we observe both dividend growth (Δd_t) and returns (r_t) directly from the data. Thus, unlike standard tests in the literature, this is not a joint hypothesis test of the market efficiency and a discount rate model.

What restrictions rationality (or market efficiency) imposes on the size of the NC? From (4.8), variance of the NC is

$$\frac{1 - \rho^{2(K+1)}}{1 - \rho^2} \frac{\sigma_\epsilon^4 \sigma_\nu^2}{(\sigma_\epsilon^2 + \sigma_\nu^2)^2} + \frac{1 - \rho^{2(K+1)}}{1 - \rho^2} \frac{\sigma_\varepsilon^4 \sigma_\upsilon^2}{(\sigma_\varepsilon^2 + \sigma_\upsilon^2)^2}, \tag{4.10}$$

which is maximized when $\sigma_\nu^2 = \sigma_\epsilon^2$ and $\sigma_\upsilon^2 = \sigma_\varepsilon^2$, and explains half of δ_t variation. This proves the following proposition:

Proposition 4.2: Market Efficiency implies that noise component can not explain more than 50% of the δ_t fluctuations.

Proposition 4.2. provides the key theoretical restriction that links the price-dividend ratio fluctuation to noise. The nest step is to think about how to test this restriction empirically. It turn out that the measure of fit, R^2, of the regression (4.9) is particularly useful for this purpose. Noticing that $1 - R^2$ is the percent of the dependent variable (δ_t) variance explained by residuals, proposition 4.2 implies that R^2 must be greater than or equal to 0.5. This proves the following corollary.

Corollary 4.3: Market Efficiency implies that in the linear regression (4.9)

$$R^2 \geq 0.5 \tag{4.11}$$

---

[16]In the asset pricing literature any deviation of price from the right hand side of (4.1) is known as *bubble*. For an accessible survey of the bubble literature see Brunnermeier (2009).

4.1   Empirical Results

This section presents the empirical analysis to measure the noise in the stock market and test the restriction (4.11). I use the annual S&P 500 index data, obtained from Robert Shiller's website (`http://www.econ.yale.edu/~shiller/data.htm`). The data cover 1871-2015 period.

A practical issue in implementing the tests and the historical decomposition is the choice of K and ρ. For the full sample, I set $K = 15$, which is long enough to capture all possible news that investors might have received about the future economic fundamentals. Chen (2009) provides evidence that returns and dividend growth are predictable at 15- and 20-year horizons.[17] In a similar study analyzing stock price volatility, Cochrane (1992) also uses $K = 15$. The results are not particularly sensitive to the choice of K. The parameter of linearizion can be directly constructed from the data. Empirically, over the period 1871 to 2016 the average log price-dividend ratio is 0.04, implying that $\rho = 0.96$.

Table (1) presents the results. The focus of the regressions is the R^2 (in bold), and its standard error. [18] Consistent with the PV model and market efficiency, the R^2 value for the regression (4.9) is 0.52. The correlation between δ_t and $\hat{\delta}_t$ is 0.72, providing strong evidence that agents have some information about the future economic fundamentals. The coefficients are strongly significant, but due to the endogeneity problem the estimates are not reliable. However, this is not of any concern as we do not test any restrictions on the parameters of the regression (4.9).

The $R^2 = 0.52$ does not look that impressive, but in the subsamples the R^2 increases significantly, reaching values 0.72 in the prewar period and 0.63 in the postwar period,

---

[17]See also Cochrane (2008) and Fama and French (1988).

[18]The standard error for the R^2 is calculated as:

$$SE_{R^2} = \sqrt{\frac{4R^2(1 - R^2)^2(n - k - 1)^2}{(n^2 - 1)(n + 3)}},$$

where n is the number of observations, and $k + 1$ counts an intercept plus the number of independent variables. See Olkin and Finn (1995).

Table 1: Testing Market Efficiency against Actual Data Using Corollary 4.3

Sample Period	K	γ_0	γ_1	$\mathbf{R^2}$	Chow 1945
$1871-2015$	15	3.51 (0.04)	0.67 (0.06)	**0.52** **(0.06)**	66.92 (0.000)
$1871-1945$	10	3.28 (0.02)	0.69 (0.05)	**0.72** **(0.05)**	
$1946-2015$	10	3.70 (0.05)	0.80 (0.08)	**0.63** **(0.07)**	

This table reports estimation results for the equation $\delta_t = \gamma_0 + \gamma_1 x_t + u_t$, where $\delta_t \equiv p_t - d_{t-1}$ and $x_t \equiv \sum_{j=0}^{K} \rho^j (\Delta d_{t+j} - r_{t+j})$. The standard errors are reported below each coefficients. The last column reports the F-statistic and associated p-value from a Chow test with null hypothesis of no structural break in 1945. The data are annual for 1871-2015.

both highly significant. One possible explanation for these findings is parameter instability. Fama and French (2001) and Chen et al. (2012) argue that dividend policy has changed significantly in the postwar period, which complicates predictability regressions.[19] Confirming the instability of the parameter estimates, the final column of Table (1) reports the results of a Chow (1960) test, which strongly rejects the null hypothesis of no structural break in 1945. Appendix B provides evidence that our measure of noise is not driven by dividend smoothing.

These empirical findings suggest that noise explains a large portion of price-dividend ratio fluctuations: 48% in the full sample, 28% in the prewar period, and 37% in the postwar period. The next question is what are the other sources of price-divided ratio fluctuations? According to the PV model (4.1), high price-dividend ratio implies that either future dividend growth must be high or future discount rates must be low. To evaluate the contribution of each component, I repeat the exercise but fix the returns or the dividends changes to their respective mean values when constructing the right hand side of (4.9). First, I examine dividend growth. Panel A of Table (2) shows that price-dividend ratio has a very weak (almost zero) relation to the estimated VC of future

---

[19]See also Lettau and Van Nieuwerburgh (2008).

Table 2: Relative Importance of Dividend Growth and Returns

| Sample Period	K	γ_0	γ_1	$\mathbf{R^2}$	Chow 1945
Panel A: $\delta_t = \gamma_0 + \gamma_1 \sum_{j=0}^{K} \rho^j \Delta d_{t+j} + u_t$					
$1871 - 2015$	15	3.08 (0.03)	-0.02 (0.11)	**0.001** **(0.005)**	23.31 (0.000)
$1871 - 1945$	10	0.00 (0.02)	0.03 (0.08)	**0.002** **(0.009)**	
$1945 - 2015$	10	0.00 (0.07)	-0.02 (0.34)	**0.000** **(0.000)**	
Panel B: $\delta_t = \gamma_0 + \gamma_1 \sum_{j=0}^{K} \rho^j r_{t+j} + u_t$					
$1871 - 2015$	15	3.38 (0.05)	-0.37 (0.05)	**0.297** **(0.065)**	49.85 (0.000)
$1871 - 1945$	10	3.09 (0.03)	-0.25 (0.05)	**0.240** **(0.081)**	
$1945 - 2015$	10	3.71 (0.07)	-0.62 (0.08)	**0.485** **(0.085)**	

Panel A reports estimation results for the equation $\delta_t = \gamma_0 + \gamma_1 x_t + u_t$, where $\delta_t \equiv p_t - d_{t-1}$ and $x_t \equiv \sum_{j=0}^{K} \rho^j \Delta d_{t+j}$. In Panel B, we regress δ_t on $x_t \equiv \sum_{j=0}^{K} \rho^j r_{t+j}$. The first column reports the sample period, and the second reports the value for K. The standard errors are reported below each coefficients. The last column reports the F-statistic and associated p-value from a Chow test with null hypothesis of no structural break in 1945. The data are annual for 1871-2015.

dividend growth. Panel B of Table (2) repeats the analysis for the returns. The R^2 increases to 0.30 over the full sample, and 0.24 in the prewar period. But, in the postwar period the R^2 increases to 0.49, and the restriction (4.11) cannot be rejected at the 5% level. Overall, the regressions of (2) suggest that most variation in price-dividend ratio corresponds to varying expected returns and almost nothing from movements in expected dividend growth. See also Cochrane (2008) and the references therein.

The major advantage of my framework is that it can be used to provide a value-noise decomposition of stock prices, and to date stamp overvaluation and undervaluation periods. The right panels of Figure (1) plots the time series of log price-dividend ratio (δ_t), estimated VC ($\hat{\delta}_t$), and estimated NC ($\delta_t - \hat{\delta}_t$). The left panels report the percentage deviation of price-dividend ratio from the VC, i.e., $\frac{(\delta_t - \hat{\delta}_t)}{\hat{\delta}_t} \times 100$. As figures clearly show,

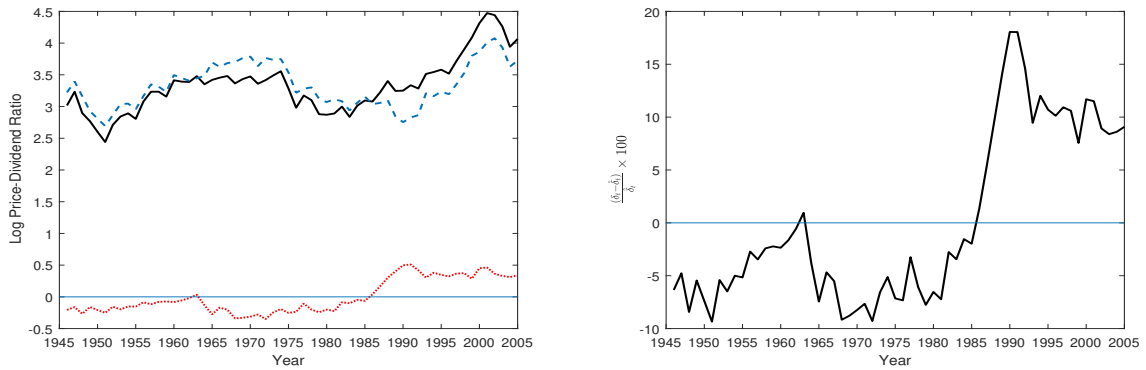Figure 1: Value-Noise Decomposition of Price-Dividend Ratio

(a) Sample Period: 1871-2015
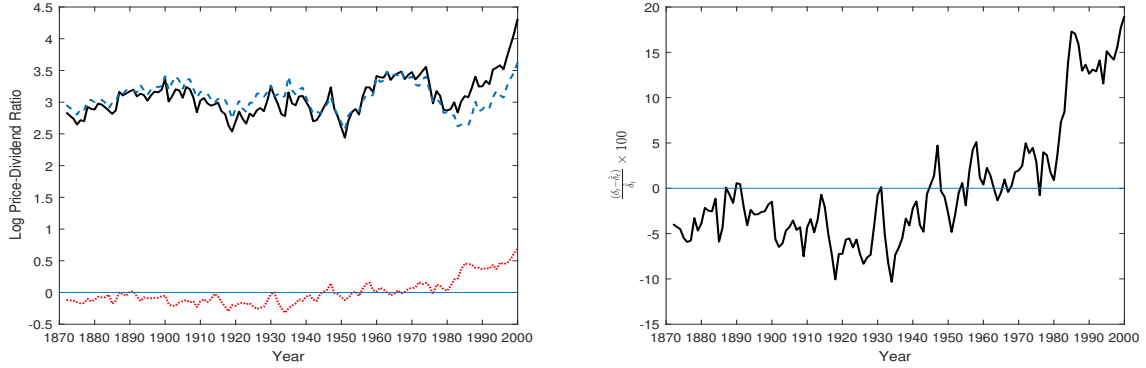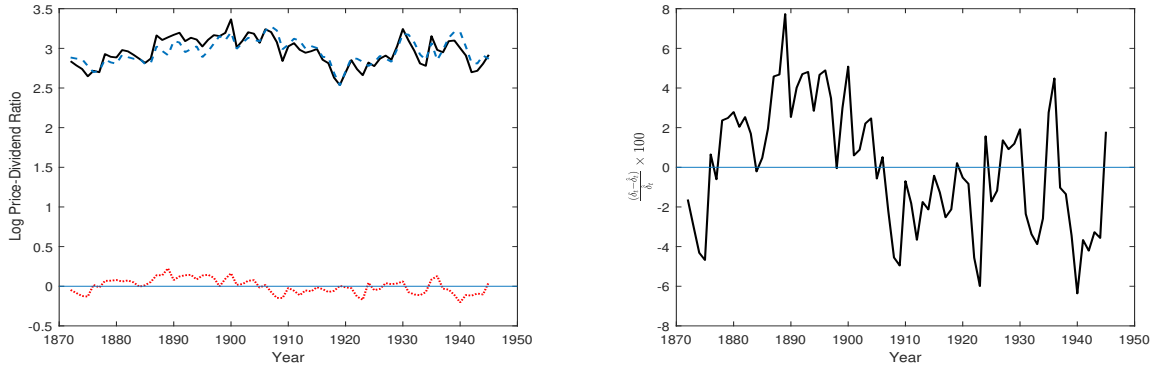
(b) Sample Period: 1871-1945

(c) Sample Period: 1946-2015

The figure plots historical decomposition of stock prices according to the equation (4.9). The top panels plot the decomposition for the full sample (1871-2015), the middle panels is for the prewar period (1871-1945), and the bottom panels gives the decomposition for the postwar period (1946-2015). The left panels plot the value-noise decomposition of stock prices: The solid (black) line is the log price-dividend ratio δ_t, the dashed (blue) line is the estimated VC $\hat{\delta}_t$, and the dotted (red) line is the estimated NC, i.e., $\delta_t - \hat{\delta}_t$. The right panels report the noise as a percentage of VC, i.e., $\frac{(\delta_t - \hat{\delta}_t)}{\hat{\delta}_t} \times 100$. The data are annual for 1871-2015.

noise can be negative as well as positive. Intuitively, a positive NC implies that prices are high compare to what an investor will receive in the form of future dividends and returns, and vice versa.

From visual inspection of Figure (1) we observe that the estimated VC mimics fairly well the broad movement in stock prices. However, there is still significant noise in the market, which implies that prices deviate from VC. However, since the NC is inside the fundamental value, this deviation is not due to a bubble or other irrational explanations.

Some famous historical periods deserve further discussion here. Lets take the *Great Depression* first. Many economists believe that the stock market crash of 1929 followed the burst of a speculative bubble that had taken hold in the late 1920s. See for instance Shiller (2015). Visual inspection of Figure (1) it clearly emerges that stock prices did not significantly deviate from the VC, suggesting that the stock market was not overvalued before the Wall Street Crash of 1929. This leads me to conclude that the boom of the 1920s was not necessarily due to a bubble in the market price. [20] These empirical finding are robust to dividend smoothing, approximation error, and linear regression specification. See Appendixes B, C, and D.

Another key period is the 1990s, where the price of Internet related stocks soared rapidly before plummeting in 2000. From Figure (1) it emerges that the stock prices deviated the VC by about 15%. Thus, the market was overvalued before the Crash and the overvaluation started around 1988. Finally, my estimates show that the U.S. stock market was undervalued during 1965-1985. Modigliani and Cohn (1979) hypothesize that the stock market suffers from *money illusion*, discounting real dividends using nominal discount rates. An implication of such irrational behavior is that when inflation is high (as it was the case during 1970s), the stock market is undervalued. Fama (1981) gives a rational interpretation of this phenomenon: high inflation signals a decline in future

---

[20]Using data on productive capital on stocks and tax rates to estimate the fundamental value, Mc-Grattan and Prescott (2001) also find that stock prices were undervalued, even at their pick. See also Donaldson and Kamstra (1996) and Pástor and Veronesi (2006).

economic activities, and stock prices rationally reflected this information. My proposed method is agnostic as to whether these overvaluation and undervaluation is rational or irrational.

5 Conclusions

Thirty years ago, Black (1986) emphasized the role of noise in the financial markets and argued that noise ought to be contrasted with information. Through the lens of a present value model, this paper explores the effects of news (or noisy information) on stock prices and drives theoretical restrictions that link price volatility to noise and information. In the model, agents receive news about future dividend changes and returns. Not distinguishing noise from information, agents react to both just the same way and as a result noise affects stock prices.

I show that the present value relation imposes testable restrictions on the contribution of noise to the price volatility. In contrast to standard tests in the literature, my approach does not require any particular assumptions about the unobserved discount rate process and the information set of the investors. Using annual data on the U.S. stock market data over the period 1871-2015, I find that consistent with the predictions of market efficiency noise does not explain more than half of price-dividend ratio fluctuations. Moreover, the variation in price-dividend ratio is mostly driven by news about future discount rates and noise, and almost nothing from the news about future dividend growth.

This paper also proposes a decomposition of prices into a value component and noise component. The decomposition provides a simple procedure to date stamp overvaluation and undervaluation. Applying my decomposition to the U.S. stock market data, I find that the market was undervalued during the 1970s and overvalued during the 1990s. Interestingly, I find no evidence the market was overvalued before the Wall Street Crash of 1929.

Appendix A: Proofs

Proof of Proposition 3.3: Since we do not want to take a precise stance on the economic model generating stock prices, lets consider a general VAR model of order p

$$\Phi(L)Z_t = Z_t - \Phi_1 Z_{t-1} - \Phi_2 Z_{t-2} - \cdots - \Phi_p Z_{t-p} = \eta_t, \tag{A.1}$$

where $\Phi(z) = I_m - \Phi_1 z - \Phi_2 z^2 - \cdots - \Phi_p z^p$ is the AR polynomial, and $Z_t := (Z_{t,1}, \cdots, Z_{t,m})^T$ is an m-dimensional stochastic process, including at least prices and dividends, and $\eta_t := (\eta_{t,1}, \cdots, \eta_{t,m})^T$ satisfies Assumption 3.2. The conventional causal VAR models require that $\det \Phi(z) \neq 0$, for all $z \in \mathbb{C}$, such that $|z| \leq 1$. To allow for the forward-looking feature of stock prices, we allow $\det \Phi(z)$ to have roots inside the unit circle. Without loss of generality, assume that $\det \Phi(z)$ has l roots outside the unit circle (the causal roots) and $p - l$ roots inside the unit circle (the non-causal roots).

We want ot show that the optimal forecast of a non-causal process based on its past is nonlinear, i.e., $\mathbb{E}[Z_t|H_t] \neq \mathbb{L}[Z_t|H_t]$. The proof of the univariate case directly follows from Rosenblatt (2000), Theorem 5.4.1 and Corollary 5.4.3. To prove the multivariate case, first note that the optimal forecast in the MSE sense is linear, *if and only if*

$$\mathbb{E}[\eta_t|H_{t-1}] = 0, \tag{A.2}$$

where $\eta_t = Z_t - \mathbb{L}[Z_t|H_{t-1}]$. Therefore, in the following, I show that the reduced form residuals obtained from fitting a linear model do not satisfy (A.2), which implies that conditional mean is nonlinear.

It is well-known that by fitting a conventional VAR model, an econometrician estimates

$$\tilde{\Phi}(L)\Phi(L)Z_t = \underbrace{\tilde{\Phi}(L)\eta_t}_{\eta_t^*}, \tag{A.3}$$

where $\tilde{\Phi}(L)$ is a Blaschke matrix defined as

$$\tilde{\Phi}(L) = B(\phi_1, z)W_1 B(\phi_2, z)W_2 \cdots B(\phi_{p-l}, z)W_{p-l}, \tag{A.4}$$

where W_i, $i = 1, \cdots, p - l$, are orthogonal matrices (i.e., $W_i W_i' = I$) and

$$B(\phi_i, z) = \begin{bmatrix} \frac{1 - \bar{\phi}_i L}{L - \phi_i} & 0 \\ 0 & I_{m-1} \end{bmatrix}$$

with $|\phi_i| < 1$ and $\bar{\phi}_i$ denotes the complex conjugate of ϕ_i. The matrix $\tilde{\Phi}(L)$, in effect, flips the roots of the $\Phi(L)$ from inside the unit circle to outside the unit circle, but preserves the unconditional second moments of the VAR system.

Now let $\eta_{1,t}^*$ denotes the first element of the vector η_t^*. Then, from the definition of η_t^* and Blaschke matrices (5) we have that

$$\eta_{1,t}^* = \sum_{a=1}^{m} \sum_v C_{a,v} \prod_{i \in v} \left(\frac{1 - \bar{\phi}_i L}{L - \phi_i} \right) \eta_{a,t}, \tag{A.5}$$

where the sum in v is over all combinations of indexes $\{1, 2, \cdots, p-l\}$ with no repetition, and $C_{a,v}$, $a = 1, \cdots, m$, functions of the underlying elements of the matrices W_i. Also note that since the Blaschke matrices are full rank matrices, $C_{a,v} \neq 0$ for at least one a and v.

Now from Theorem 5.4.1 and Corollary 5.4.3 of Rosenblatt (2000) we know that the best predictor of each element of the sum is nonlinear, thus

$$\mathbb{E}[\eta_{1,t}^* | \eta_{t-1}^*, \eta_{t-2}^*, \cdots] = \mathbb{E}[\eta_{1,t}^* | H_{t-1}] \neq 0, \tag{A.6}$$

which implies that $\mathbb{E}[\eta_t^* | H_{t-1}] \neq 0$, and the proof is complete.

Appendix B: Dividend Smoothing

The predictability regression is controversial because of dividend smoothing. In their seminal work, Miller and Modigliani (1961) argue that in a perfect capital market dividend policy is irrelevant, and the stock prices should be driven by earning power of a firm's asset. Chen et al. (2012) propose a modified PV model to get around the issue of dividend smoothing. To see how it works, consider the Lintner (1956) partial adjustment model in log form:

$$\Delta d_t = \alpha_0 + \alpha_1 e_t + \alpha_2 d_{t-1} + \vartheta_t, \tag{C.7}$$

where $e_t \equiv \ln(E_t)$, E_t is the earning at time t, and ϑ_t is an error term. Rewrite (C.7) in terms of differences:

$$\Delta d_t = \alpha_1 \Delta e_t + (1 + \alpha_2)\Delta d_{t-1} + \nu_t, \tag{C.8}$$

Dividends are most smoothed if $\alpha_1 = 0$ and $\alpha_2 = 0$, in which case dividends grow at a constant rate plus some noise.

Table (3) reports the estimates from the dividend behavior models (C.7) and (C.8). From Panel A, we see that both α_1 and α_2 have reduced significantly in the postwar period, which implies that dividends are more smoothed. The final column of panel A reports a Chow (1960) test that indicates a significant structural break around 1945. Panel B confirms these empirical findings.

To get around the issue of dividend smoothing, notice that from (C.8) it follows that

$$\sum_{j=0}^{\infty} \rho^j \Delta d_t = \frac{(1 + \alpha_2)}{1 - (1 + \alpha_2)\rho} \Delta d_{t-1} + \frac{\alpha_1}{1 - (1 + \alpha_2)\rho} \sum_{j=0}^{\infty} \rho^j \Delta e_{t+j}$$
$$+ \frac{\alpha_1}{1 - (1 + \alpha_2)\rho} \sum_{j=0}^{\infty} \rho^j \Delta \vartheta_{t+j} \tag{C.9}$$

Suppressing the constant, (4.1) can be rewritten as

$$\delta_t = \mathbb{E}\Big[\sum_{j=0}^{\infty}\rho^j r_{t+j} + \frac{\alpha_1}{1-(1+\alpha_2)\rho}\sum_{j=0}^{\infty}\rho^j\Delta e_{t+j}\Big|I_t\Big]. \tag{C.10}$$

The above implies that, to measure the noise and test the restriction (4.11), one can use the modified versions of (4.9), as follows

$$\delta_t = \gamma_0 + \gamma_1 x_t + u_t, \tag{C.11}$$

where $x_t = \sum_{j=0}^{K}\rho^j r_{t+j} + \frac{\alpha_1}{1-(1+\alpha_2)\rho}\sum_{j=0}^{K}\rho^j\Delta e_{t+j}$.

Table (4) presents the results. The R^2 value for in the prewar period is 0.33, which does not satisfy the restriction (4.11), but this increases to 0.52 for the postwar period. The right panels of Figure (2) plots the time series of log price-dividend ratio (δ_t), value component ($\hat{\delta}_t$), and noise components ($\delta_t - \hat{\delta}_t$), estimated from the regression (C.11). The left panels report the percentage deviation of price-dividend ratio from the value component, i.e., $\frac{(\delta_t - \hat{\delta}_t)}{\hat{\delta}_t} \times 100$. These empirical findings suggest that although dividend smoothing is an important issue in the postwar period, our measure of noise is not driven by dividend smoothing.

Table 3: Dividend Policy Models Using Actual Dividends and Earnings 18712015

Panel A: $\Delta d_t = \alpha_0 + \alpha_1 e_t + \alpha_2 d_{t-1} + \vartheta_t$					
Sample Period	α_0	α_1	α_2	R^2	Chow 1945
$1871 - 2015$	0.01 (0.04)	0.24 (0.02)	-0.32 (0.03)	0.40	5.81 (0.001)
$1871 - 1945$	0.08 (0.11)	0.33 (0.04)	-0.42 (0.06)	0.50	
$1945 - 2015$	0.06 (0.07)	0.12 (0.02)	-0.17 (0.04)	0.26	

Panel B: $\Delta d_t = \beta_0 + \beta_1 \Delta e_t + \beta_2 \Delta d_{t-1} + \nu_t$					
Sample Period	β_0	β_1	β_2	R^2	Chow 1945
$1871 - 2015$	0.01 (0.01)	0.13 (0.03)	0.16 (0.08)	0.11	5.52 (0.001)
$1871 - 1945$	0.00 (0.01)	0.26 (0.06)	0.11 (0.11)	0.21	
$1945 - 2015$	0.01 (0.01)	0.02 (0.03)	0.38 (0.12)	0.14	

This table reports estimation results for two dividend behavior models. The first is the (C.7) model and the second is estimated using the first differences (C.8). The standard errors are reported below each coefficients. The last column reports the F-statistic and associated p-value from a Chow test with null hypothesis of no structural break in 1945. The data are annual for 1871-2015.

Table 4: Testing Market Efficiency against Actual Data Using Corollary 4.3

| $\delta_t = \gamma_0 + \gamma_1 x_t + u_t$					
Sample Period	K	γ_0	γ_1	$\mathbf{R^2}$	Chow 1945
$1871 - 2015$	15	3.40 (0.04)	0.43 (0.05)	**0.34** (0.07)	51.13 (0.000)
$1871 - 1945$	10	3.12 (0.03)	0.32 (0.05)	**0.33** (0.09)	
$1946 - 2015$	10	3.71 (0.06)	0.65 (0.08)	**0.52** (0.08)	

This table reports estimation results for the linear specification (C.11), where $x_t = \sum_{j=0}^{K} \rho^j r_{t+j} + \frac{\alpha_1}{1-(1+\alpha_2)\rho} \sum_{j=0}^{K} \rho^j \Delta e_{t+j}$. The standard errors are reported below each coefficients. The last column reports the F-statistic and associated p-value from a Chow test with null hypothesis of no structural break in 1945. The data are annual for 1871-2015.

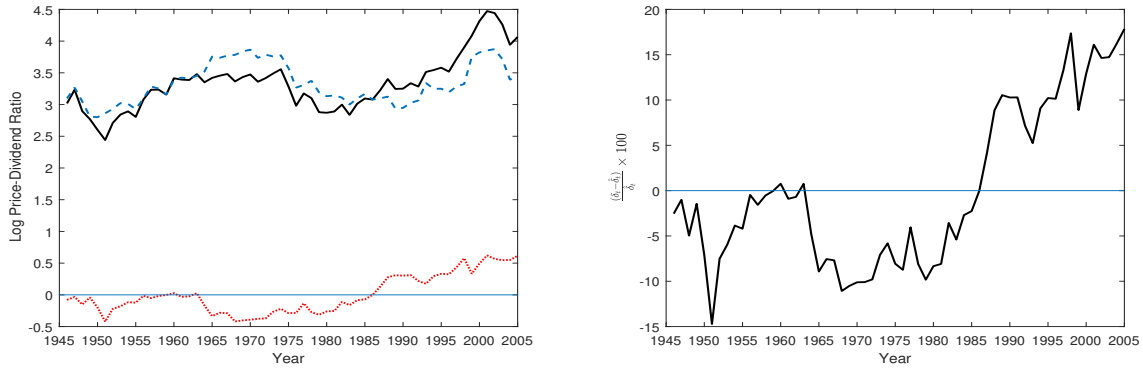# Figure 2: Value-Noise Decomposition of Price-Dividend Ratio: Dividend Smoothing

(a) Sample Period: 1871-2015
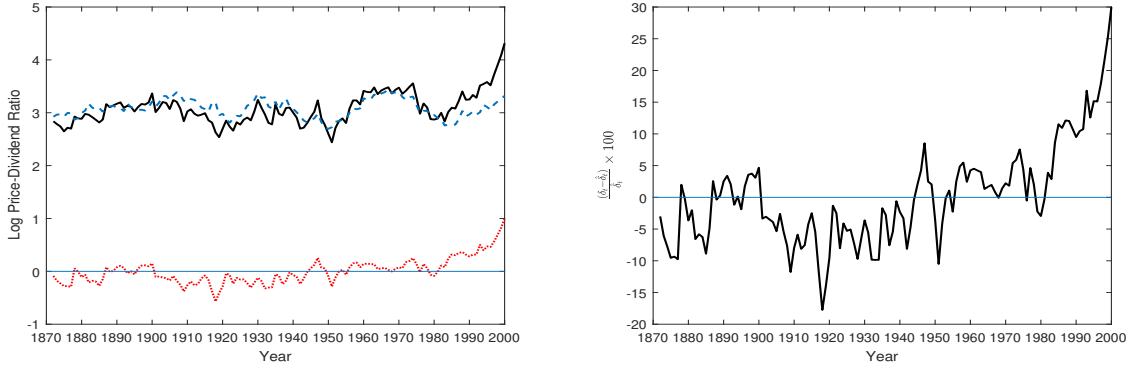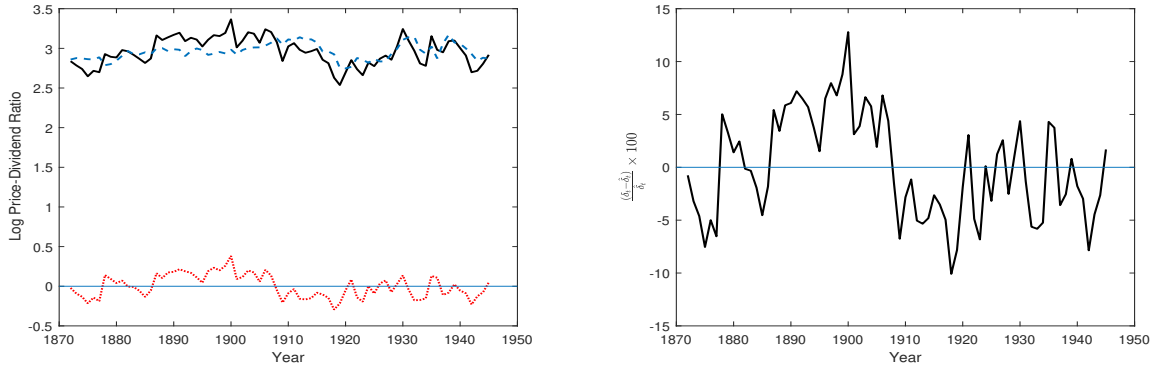
(b) Sample Period: 1871-1945

(c) Sample Period: 1946-2015

The figure plots historical decomposition of stock prices according to the equation (C.11). The top panels plot the decomposition for the full sample (1871-2015), the middle panels is for the prewar period (1871-1945), and the bottom panels gives the decomposition for the postwar period (1946-2015). The left panels plot the value-noise decomposition of stock prices: The solid (black) line is the log price-dividend ratio δ_t, the dashed (blue) line is the estimated value component $\hat{\delta}_t$, and the dotted (red) line is the estimated noise component, i.e., $\delta_t - \hat{\delta}_t$. The right panels report the noise as a percentage of value component, i.e., $\frac{(\delta_t - \hat{\delta}_t)}{\hat{\delta}_t} \times 100$. The data are annual for 1871-2015.

Appendix C: Approximation Error

So far, the empirical results are based on a first-order Taylor approximation of the general Euler equation. A reasonable concern is that the higher-order terms in the Taylor expansion, which are neglected, can create an approximation error. Campbell and Shiller (1988a) show that empirically the error is small and almost constant. In this section, I will argue that the main results of this paper are not sensitive to the approximation error.

Consider a general one-period Euler equation:

$$P_t = \mathbb{E}\big[M_{t+1}(P_{t+1} + D_t)|I_t\big].$$

Iterating forward and imposing the transversality condition, we obtain the PV model with time-varying discount rates

$$P_t = \mathbb{E}\left[\sum_{j=1}^{\infty}\Big(\prod_{i=1}^{j} M_{t+1}\Big)D_{t+j}|I_t\right]. \tag{C.12}$$

Since prices and dividends are not stationary, population moments can not be estimated from sample counterparts. However, if dividend growth, $\frac{D_t}{D_{t-1}} = G_t$, and discount rates are stationary, the PV model implies that price-dividend ratio is stationary.[21] Thus, we can write (C.12) as

$$\frac{P_t}{D_t} = \mathbb{E}\left[\sum_{j=1}^{\infty}\prod_{i=1}^{j} M_{t+i}G_{t+i}|I_t\right], \tag{C.13}$$

where $M_t G_t$ is the discounted dividend growth rate, which can be obtained directly from the data.

We can approximate the noise component by the residuals of the following modified linear regression

$$\frac{P_t}{D_t} = \gamma_0 + \gamma_1 X_t + u_t, \tag{C.14}$$

---

[21]For the formal proof see (Cochrane, 1994).

where $X_t = \sum_{j=1}^{K} \prod_{i=1}^{j} M_{t+i} G_{t+i}$ and u_t is orthogonal to the right hand side by construction, and the fitted value of the regression approximates the fundamental component.

Table (5) presents the results. The R^2 value for in the prewar period is 0.68, but this decreases to 0.51 for the postwar period. This should provide evidence that our measure of noise is not driven by approximation error. The right panels of Figure (3) plots the time series of log price-dividend ratio (δ_t), value component $(\hat{\delta}_t)$, and noise components $(\delta_t - \hat{\delta}_t)$, estimated from the regression (C.14). The left panels report the percentage deviation of price-dividend ratio from the value component, i.e., $\frac{(\delta_t - \hat{\delta}_t)}{\hat{\delta}_t} \times 100$. Surprisingly, the noise appears to explain an even bigger portion of price fluctuations, but the conclusion that the US market was undervalued during 1970s and overvalued during 1990s still holds. Still there is no sign of overvaluation before the crash of 1930s.

Table 5: Testing Market Efficiency against Actual Data Using Corollary 4.3

| $\frac{P_t}{D_t} = \gamma_0 + \gamma_1 X_t + u_t$					
Sample Period	K	γ_0	γ_1	$\mathbf{R^2}$	Chow 1945
$1871 - 2015$	15	1.28 (2.6)	2.08 (0.24)	**0.36** **(0.07)**	31.67 (0.000)
$1871 - 1945$	10	0.02 (1.56)	2.52 (0.20)	**0.68** **(0.06)**	
$1946 - 2015$	10	-6.48 (5.07)	4.64 (0.60)	**0.51** **(0.08)**	

This table reports estimation results for the equation (C.14), where $X_t \equiv \sum_{j=1}^{K} \prod_{i=1}^{j} M_{t+i} G_{t+i}$, $M_t \equiv \frac{P_t + D_t}{P_{t-1}}$ and $G_t \equiv \frac{D_t}{D_{t-1}}$. The first column reports the sample period, and the second reports the value for K. The standard errors are reported below each coefficients. The last column reports the F-statistic and associated p-value from a Chow test with null hypothesis of no structural break in 1945. The data are annual for 1871-2015.

Figure 3: Value-Noise Decomposition of Price-Dividend Ratio: Exact PV Model
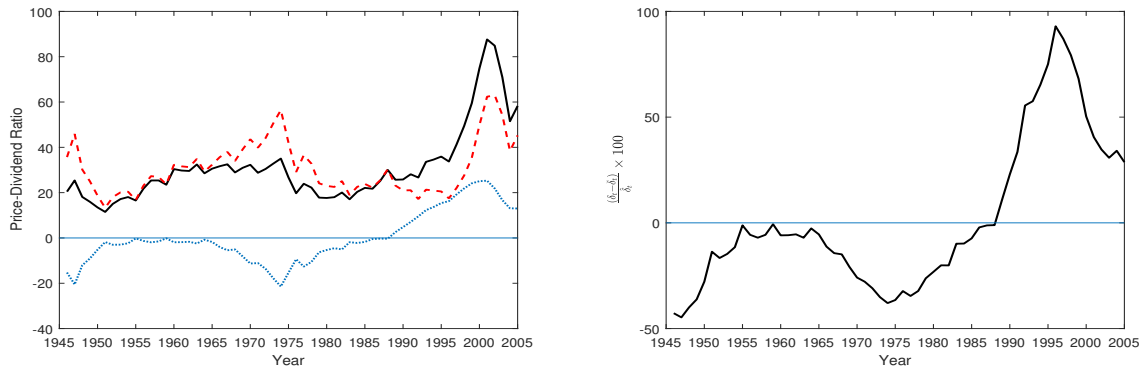
(a) Sample Period: 1871-2015
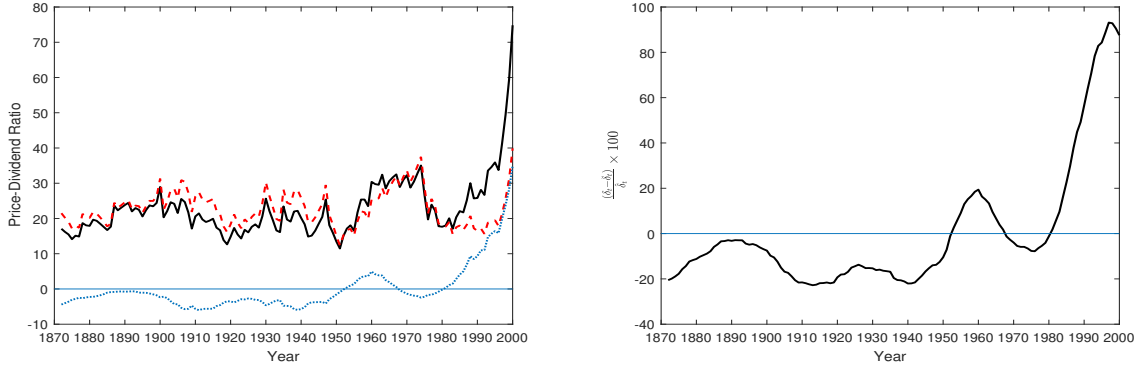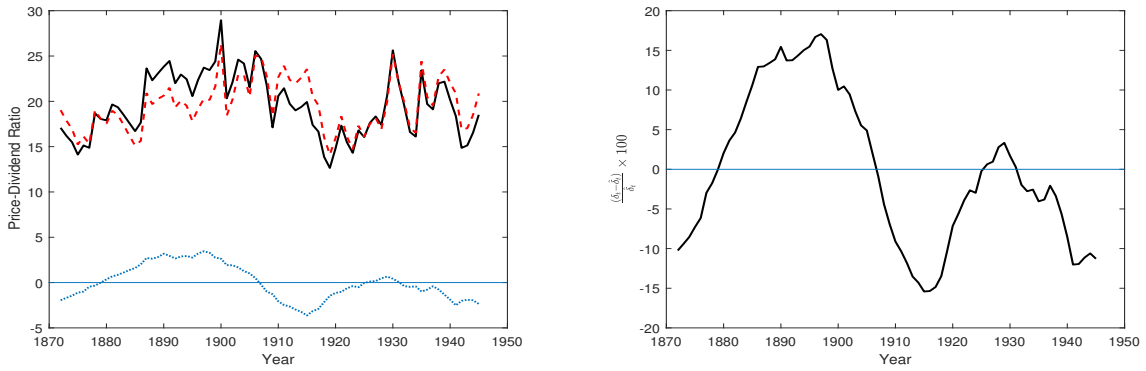
(b) Sample Period: 1871-1945

(c) Sample Period: 1946-2015

The figure plots historical decomposition of stock prices according to the equation (C.14). The top panels plot the decomposition for the full sample (1871-2015), the middle panels is for the prewar period (1871-1945), and the bottom panels gives the decomposition for the postwar period (1946-2015). The left panels plot the value-noise decomposition of stock prices: The solid (black) line is the log price-dividend ratio δ_t, the dashed (blue) line is the estimated value component $\hat{\delta}_t$, and the dotted (red) line is the estimated noise component, i.e., $\delta_t - \hat{\delta}_t$. The right panels report the noise as a percentage of value component, i.e., $\frac{(\delta_t - \hat{\delta}_t)}{\hat{\delta}_t} \times 100$. The data are annual for 1871-2015.

Appendix D: Quadratic Specification

While the linear specification provides important insights, in many settings, it may not hold. A tractable alternative is to formulate the decomposition problem as a linear-quadratic regression,

$$\delta_t = \gamma_0 + \gamma_1 x_t + \gamma_2 x_t^2 + u_t, \tag{D.15}$$

where $x_t = \sum_{j=0}^{K} \rho^j (\Delta d_{t+j} - r_{t+j})$. Panel B of Table (6) reports the results for the quadratic model that fits different sample periods. Again, the R^2 is grater than 0.5 in all periods, which is consistent with market efficiency. Including cubic terms improves the R^2, but not significantly. The right panels of Figure (4) plots the time series of log price-dividend ratio (δ_t), value component ($\hat{\delta}_t$), and noise components ($\delta_t - \hat{\delta}_t$), estimated from the regression (D.15). The left panels report the percentage deviation of price-dividend ratio from the value component, i.e., $\frac{(\delta_t - \hat{\delta}_t)}{\hat{\delta}_t} \times 100$.

Table 6: Testing Market Efficiency against Actual Data Using Corollary 4.3

|		$\delta_t = \gamma_0 + \gamma_1 x_t + \gamma_2 x_t^2 + u_t$				
Sample Period	K	γ_0	γ_1	γ_2	$\mathbf{R^2}$	Chow 1945
$1871 - 2015$	15	3.68 (0.06)	1.39 (0.18)	0.56 (0.14)	**0.57** **(0.05)**	34.50 (0.000)
$1871 - 1945$	10	3.24 (0.04)	0.46 (0.17)	-0.25 (0.18)	**0.73** **(0.05)**	
$1946 - 2015$	10	3.70 (0.05)	1.17 (0.17)	0.43 (0.18)	**0.66** **(0.06)**	

This table reports estimation results for the equation (D.15), where $\delta_t \equiv p_t - d_{t-1}$ and $x_t \equiv \sum_{j=0}^{K} \rho^j (\Delta d_{t+j} - r_{t+j})$. The first column reports the sample period, and the second reports the value for K. The standard errors are reported below each coefficients. The last column reports the F-statistic and associated p-value from a Chow test with null hypothesis of no structural break in 1945. The data are annual for 1871-2015.

Figure 4: Value-Noise Decomposition of Price-Dividend Ratio: Quadratic Specification
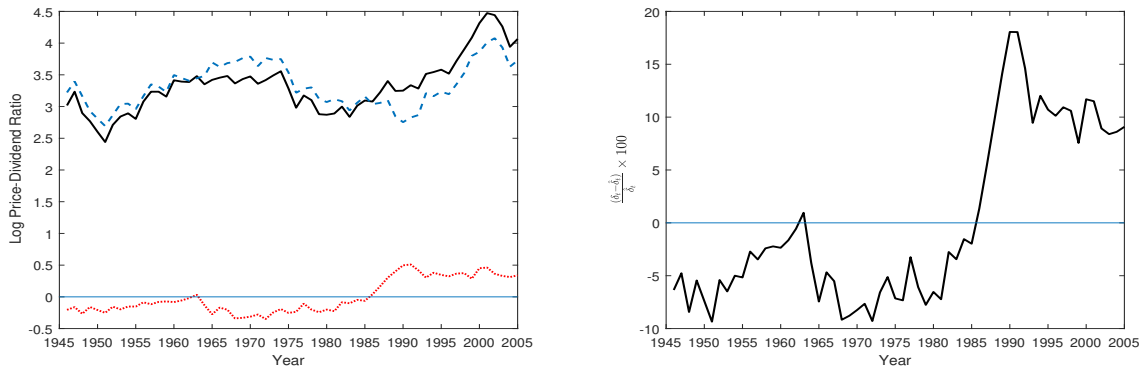
(a) Sample Period: 1871-2015
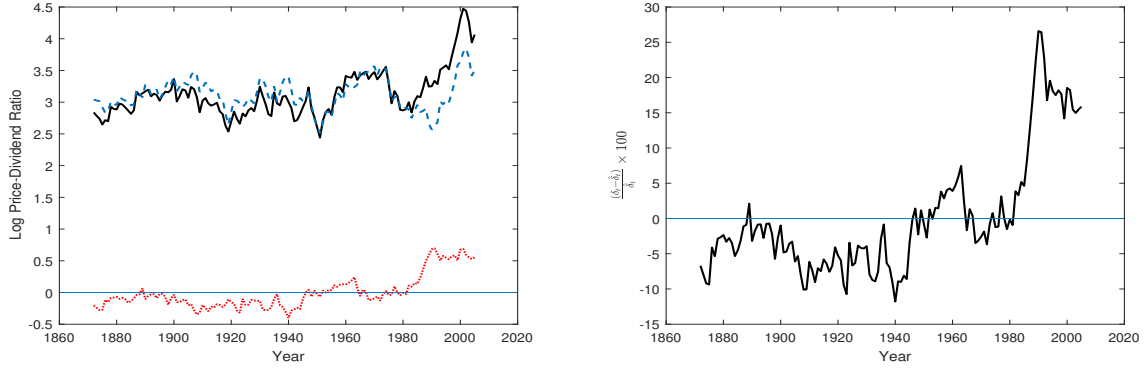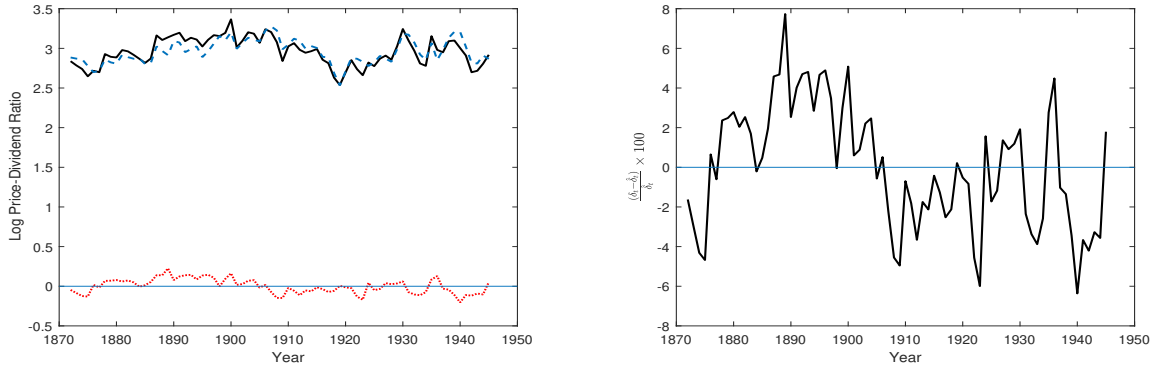
(b) Sample Period: 1871-1945

(c) Sample Period: 1946-2015

The figure plots historical decomposition of stock prices according to the equation (D.15). The top panels plot the decomposition for the full sample (1871-2015), the middle panels is for the prewar period (1871-1945), and the bottom panels gives the decomposition for the postwar period (1946-2015). The left panels plot the value-noise decomposition of stock prices: The solid (black) line is the log price-dividend ratio δ_t, the dashed (blue) line is the estimated value component $\hat{\delta}_t$, and the dotted (red) line is the estimated noise component, i.e., $\delta_t - \hat{\delta}_t$. The right panels report the noise as a percentage of value component, i.e., $\frac{(\delta_t - \hat{\delta}_t)}{\hat{\delta}_t} \times 100$. The data are annual for 1871-2015.

Appendix E: Simulations

This section contains Monte Carlo simulations to shed light on the importance of assumption (CS4). The algorithm for generating data from the model specified in Section 2 is as follows:

1. Obtain dividend changes $\epsilon_t = d_t - d_{t-1}$ and its variance σ_ϵ^2. Calculate $\rho \equiv 1/\big(1 + \exp(\overline{d_t - p_t})\big)$ and $\kappa \equiv -\log(\rho) - (1-\rho)\log(1/\rho - 1)$ from the data. Set $K = 15$ and $itr = 1$.

2. Generate random noise using $\nu_t \sim iidN(0, \sigma_\nu^2)$, and set $\gamma = \frac{\sigma_\epsilon^2}{\sigma_\epsilon^2 + \sigma_\nu^2}$. I consider alternative values of σ_ν^2 below.

3. Generate bootstrap sample itr using the parametric bootstrap: Resample the ϵ_t's, denoting the bootstrap ϵ_t by ϵ_t^*. Set starting value $d_1^* = d_1$. For $t > 1$, construct the bootstrap sample using (2.2) and (2.5):

$$d_t^* = d_{t-1}^* + \epsilon_t^*,$$

$$p_t^* - d_{t-1}^* = \frac{\kappa}{1-\rho} + = \gamma(\epsilon_t^* + \nu_t) + \rho\gamma(\epsilon_{t+1}^* + \nu_{t+1}) + \cdots + \rho^K\gamma(\epsilon_{t+K}^* + \nu_{t+K}).$$

4. Construct $\delta_t^{*\prime}$ from a VAR system including $Z_t^* = (\Delta d_t^*, \delta_t^*)$.

5. If $itr < ITR$ then increase itr by one and return to step two.

Figure (5) compares the sample path of actual stock prices, simulated stock prices, and their VAR forecast based on a VAR with a single lag. The top left panel plots the actual δ_t, along with the VAR forecast δ_t'. Other panels plot the simulated δ_t^* and $\delta_t^{*\prime}$, for three different values of noise variance. From the plot of actual data we see that δ_t is much more volatile than δ_t', which Campbell and Shiller (1987, 1988a,b) interpret as evidence of excessively volatility. However, as the other panels clearly show, such

conclusions are absolutely unfounded, because these figures are based on simulated data that by construction are generated by the rational model (2.1).
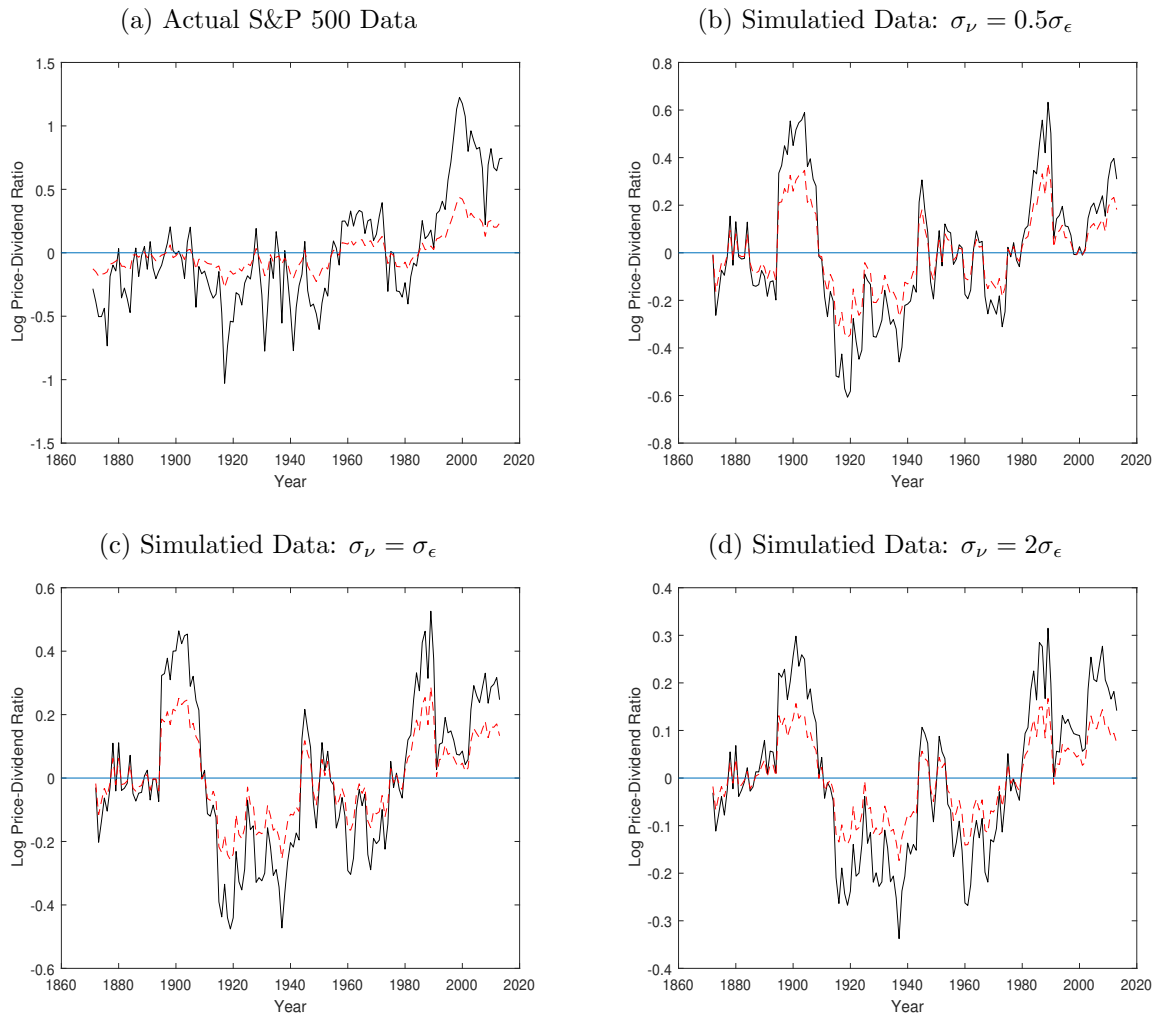
Table (7) reports a formal test of the hypothesis (3.1) and (3.15). Campbell and Shiller (1988a, p. 214) also argue that if the model with constant discount rate were true, the variable δ'_t would place a unit weight on δ_t and a zero weight on Δd_{t-1}. More specifically, in the following linear regression

$$\delta_t = \alpha + \beta_1 \delta'_t + \beta_2 \Delta d_{t-1} + v_t, \qquad \text{(E.16)}$$

CS propose to jointly test whether $(\beta_1, \beta_2) = (1, 0)$.

Table (7) also reports the result of this test. Given that assumptions (CS1)-(CS3) hold by construction, the only explanation for the observed overrejection of the null hypothesis is that assumption (CS4) is not satisfied. In contrast, the R^2 test proposed in Section 4 (column 4) controls the size.

Figure 5: Actual Stock Prices vs VAR Forecasts

(a) Actual S&P 500 Data

(b) Simulatied Data: $\sigma_\nu = 0.5\sigma_\epsilon$

(c) Simulatied Data: $\sigma_\nu = \sigma_\epsilon$

(d) Simulatied Data: $\sigma_\nu = 2\sigma_\epsilon$

This figure plots log price-dividend ratio δ_t (solid-black) versus the VAR forecast δ'_t (dashed-red). The top right panel plots the actual data. Other panels plot the simulated data with different noise variances.

41

Table 7: Tests of Present Value Models

	$\text{var}(\delta_t) = \text{var}(\delta'_t)$	$\chi^2(2)$	Test of $\delta_t = \delta'_t$	$\mathbf{R^2 \geq 0.5}$
	Actual S&P 500 Data			
p-Value	0.00	0.00	0.00	$\mathbf{\leq 0.05}$
	Simulated Data: $\sigma_\nu = 0.5\sigma_\epsilon$			
% of rejections	38.8	100	98.8	**5.6**
	Simulated Data: $\sigma_\nu = \sigma_\epsilon$			
% of rejections	45.1	100	98.2	**5.3**
	Simulated Data: $\sigma_\nu = 2\sigma_\epsilon$			
% of rejections	49.4	100	98.6	**6.0**

This table reports estimation results for the Present Value models. The top row reports the results for the actual data, where $\delta_t = p_t - d_{t-1}$ and δ'_t denotes its VAR forecast. The first column reports the p-Value for the restriction $\text{var}(\delta_t) = \text{var}(\delta'_t)$, against the alternative that $\text{var}(\delta_t) > \text{var}(\delta'_t)$. The second column reports the Wald statistic for the cross-equation restrictions (3.15). Third column reports the F-statistics for the null hypothesis that in the regression (E.16), $(\beta_1, \beta_2) = (1, 0)$. The last column (in bold) reports the R^2 test of restriction (4.11). The second-forth rows report percentage of rejections of the null hypothesis using a nominal size of 5%. All experiments have 1000 replications. The last column reports the mean of the R^2 in 1000 Monte Carlo replications.

References

Black, F., 1986. Noise. *The journal of finance*, 41(3):528–543.

Blanchard, O. J. and Watson, M. W. Bubbles, rational expectations and financial markets. Technical report, National Bureau of Economic Research, 1982.

Blanchard, O. J., L'Huillier, J.-P., and Lorenzoni, G., 2013. News, noise, and fluctuations: An empirical exploration. *American Economic Review*, 103(7):3045–70.

Boudoukh, J., Richardson, M., and Whitelaw, R. F., 2006. The myth of long-horizon predictability. *The Review of Financial Studies*, 21(4):1577–1605.

Brunnermeier, M. K., 2009. Bubbles: Entry in new palgrave dictionary of economics. *ed. Steven Durlauf and Lawrence Blume.*

Campbell, J. Y. and Shiller, R. J., 1987. Cointegration and tests of present value models. *The Journal of Political Economy*, pages 1062–1088.

Campbell, J. Y. and Shiller, R. J., 1988a. The dividend-price ratio and expectations of future dividends and discount factors. *Review of financial studies*, 1(3):195–228.

Campbell, J. Y. and Shiller, R. J., 1988b. Stock prices, earnings, and expected dividends. *The Journal of Finance*, 43(3):661–676.

Campbell, J. Y. and Thompson, S. B., 2008. Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies*, 21(4): 1509–1531.

Chen, L., 2009. On the reversal of return and dividend growth predictability: A tale of two periods. *Journal of Financial Economics*, 92(1):128–151.

Chen, L., Da, Z., and Priestley, R., 2012. Dividend smoothing and predictability. *Management science*, 58(10):1834–1853.

Chow, G. C., 1960. Tests of equality between sets of coefficients in two linear regressions. *Econometrica: Journal of the Econometric Society*, pages 591–605.

Cochrane, J. H., 1992. Explaining the variance of price–dividend ratios. *Review of Financial Studies*, 5(2):243–280.

Cochrane, J. H., 1994. Permanent and transitory components of gnp and stock prices. *The Quarterly Journal of Economics*, pages 241–265.

Cochrane, J. H., 2008. The dog that did not bark: A defense of return predictability. *Review of Financial Studies*, 21(4):1533–1575.

Cochrane, J. H., 2011. Presidential address: Discount rates. *The Journal of Finance*, 66 (4):1047–1108.

Davis, R. and Song, L., 2012. Noncausal vector ar processes with application to economic time series. *DP Columbia University*.

Diba, B. T. and Grossman, H. I., 1988. Explosive rational bubbles in stock prices? *The American Economic Review*, pages 520–530.

Donaldson, R. G. and Kamstra, M., 1996. A new dividend forecasting procedure that rejects bubbles in asset prices: The case of 1929s stock crash. *Review of Financial Studies*, 9(2):333–383.

Fama, E. F., 1970. Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, 25(2):383–417.

Fama, E. F., 1981. Stock returns, real activity, inflation, and money. *The American Economic Review*, 71(4):545–565.

Fama, E. F. and French, K. R., 1988. Dividend yields and expected stock returns. *Journal of financial economics*, 22(1):3–25.

Fama, E. F. and French, K. R., 2001. Disappearing dividends: changing firm characteristics or lower propensity to pay? *Journal of Financial economics*, 60(1):3–43.

Forni, M., Gambetti, L., Lippi, M., and Sala, L., 2014a. Noisy news in business cycles.

Forni, M., Gambetti, L., and Sala, L., 2014b. No news in business cycles. *The Economic Journal*.

Forni, M., Gambetti, L., Lippi, M., and Sala, L., 2016. Noise bubbles. *The Economic Journal*.

Giglio, S., Maggiori, M., and Stroebel, J., 2016. No-bubble condition: Model-free tests in housing markets. *Econometrica*, 84(3):1047–1091.

Gouriéroux, C. and Zakoïan, J.-M., 2016. Local explosion modelling by noncausal process. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. ISSN 1467-9868.

Hamilton, J. D., 1994. *Time series analysis*. Princeton university press Princeton.

Hansen, L. P., Sargent, T. J., et al. Exact linear rational expectations models: Specification and estimation. Technical report, 1981.

Kasa, K., Walker, T. B., and Whiteman, C. H., 2014. Heterogeneous beliefs and tests of present value models. *The Review of Economic Studies*, 81(3):1137–1163.

Koijen, R. and Van Nieuwerburgh, S., 2011. Predictability of returns and cash flows. *Annual Review of Financial Economics*, 3(1):467–491.

Lanne, M. and Saikkonen, P., 2013. Noncausal vector autoregression. *Econometric Theory*, 29(03):447–481.

Leeper, E. M., Walker, T. B., and Yang, S.-C. S., 2013. Fiscal foresight and information flows. *Econometrica*, 81(3):1115–1145.

LeRoy, S. F., 1973. Risk aversion and the martingale property of stock prices. *International Economic Review*, pages 436–446.

Lettau, M. and Van Nieuwerburgh, S., 2008. Reconciling the return predictability evidence. *Review of Financial Studies*, 21(4):1607–1652.

Lintner, J., 1956. Distribution of incomes of corporations among dividends, retained earnings, and taxes. *The American Economic Review*, 46(2):97–113.

Lippi, M. and Reichlin, L., 1994. Var analysis, nonfundamental representations, blaschke matrices. *Journal of Econometrics*, 63(1):307–325.

Lorenzoni, G., December 2009. A theory of demand shocks. *American Economic Review*, 99(5):2050–84.

Lucas Jr, R. E., 1978. Asset prices in an exchange economy. *Econometrica: Journal of the Econometric Society*, pages 1429–1445.

Malkiel, B. G., 2003. The efficient market hypothesis and its critics. *The Journal of Economic Perspectives*, 17(1):59–82.

McGrattan, E. R. and Prescott, E. C. The stock market crash of 1929: Irving fisher was right! Technical report, National Bureau of Economic Research, 2001.

Miller, M. H. and Modigliani, F., 1961. Dividend policy, growth, and the valuation of shares. *the Journal of Business*, 34(4):411–433.

Modigliani, F. and Cohn, R. A., 1979. Inflation, rational valuation and the market. *Financial Analysts Journal*, 35(2):24–44.

Olkin, I. and Finn, J. D., 1995. Correlations redux. *Psychological Bulletin*, 118(1):155.

Pástor, L. and Veronesi, P., 2006. Was there a nasdaq bubble in the late 1990s? *Journal of Financial Economics*, 81(1):61–100.

Rosenblatt, M., 2000. *Gaussian and non-Gaussian linear time series and random fields*. Springer.

Rubinstein, M., 1976. The valuation of uncertain income streams and the pricing of options. *The Bell Journal of Economics*, pages 407–425.

Schmitt-Grohé, S. and Uribe, M., 2012. What's news in business cycles. *Econometrica*, 80(6):2733–2764.

Shiller, R. J., 2014. Speculative asset prices. *The American Economic Review*, 104(6): 1486.

Shiller, R. J., 2015. *Irrational exuberance*. Princeton university press.

Tirole, J., 1985. Asset bubbles and overlapping generations. *Econometrica: Journal of the Econometric Society*, pages 1071–1100.

Welch, I. and Goyal, A., 2007. A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies*, 21(4):1455–1508.

West, K. D., 1988. Bubbles, fads and stock price volatility tests: a partial evaluation. *The Journal of Finance*, 43(3):639–656.