The Learning Premium

Maxim Bichuch* Paolo Guasoni †

November 16, 2019

Abstract

We find equilibrium stock prices and interest rates in a representative-agent model where dividend growth is uncertain, but gradually revealed by dividends themselves, while asset prices reflect current information and the potential impact of future knowledge. In addition to the usual premium for risk, stock returns include a learning premium, which reflects the expected change in prices from new information. In the long run, the learning premium vanishes, as prices and interest rates converge to their counterparts in the standard setting with known dividend growth. If both relative risk aversion and elasticity of intertemporal substitution are above one, the model reproduces the increase in price-dividend ratios observed in the past century, and implies that – in the long run – price-dividend ratios may increase a further forty percent above current levels.

JEL Classification: G11, G12.

MSC (2010): 91G10, 91G80.

Keywords: Equilibrium, Asset Pricing, Filtering, Learning.

*Department of Applied Mathematics and Statistics, Whiting School of Engineering, Johns Hopkins University, 3400 North Charles St., Baltimore, MD 21218, USA `mbichuch@jhu.edu`. Work is partially supported by NSF (DMS-1736414), and by the Acheson J. Duncan Fund for the Advancement of Research in Statistics.

†Dublin City University, School of Mathematical Sciences, Glasnevin, Dublin 9, Ireland, and Boston University, Department of Mathematics and Statistics, 111 Cummington Street, Boston, MA 02215, USA, email paolo.guasoni@dcu.ie. Partially supported by the ERC (279582), NSF (DMS-1412529), and SFI (16/SPP/3347 and 16/IA/4443).

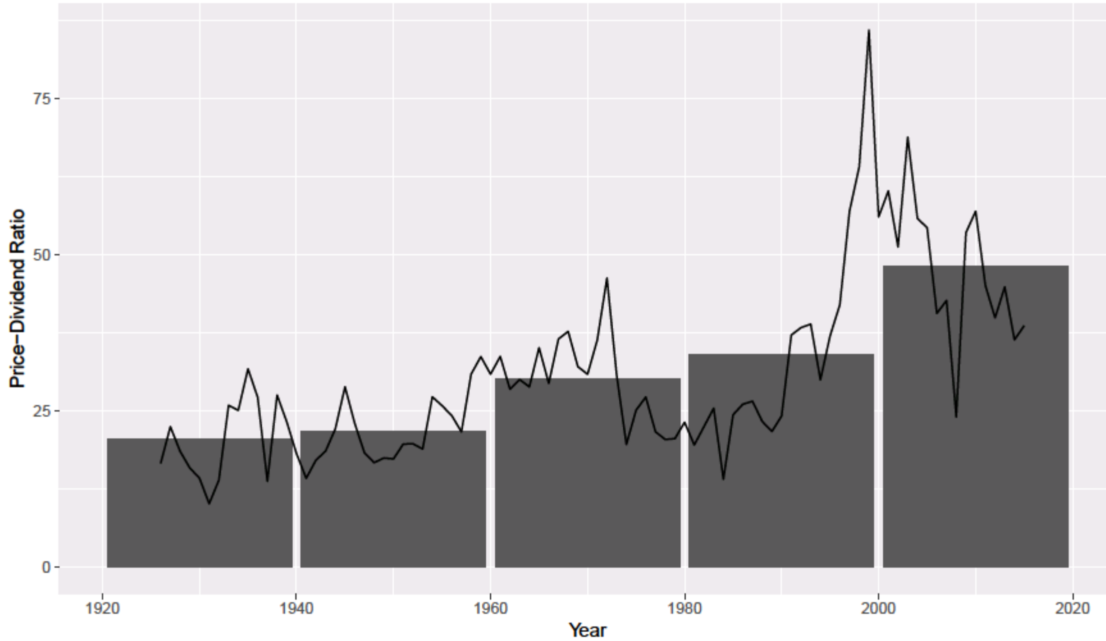

Figure 1: Price to Dividend ratio (vertical, twenty-year averages) over time (horizontal). Source: CRSP monthly data 1926-2015.

1 Introduction

In the past century, equity prices have increased on average relative to dividends (Figure 1). Such an increase is inconsistent with both the standard Lucas' model, which implies a constant price-dividend ratio, and with more recent models, designed to resolve asset pricing puzzles, in which price-dividend ratios are time-varying but randomly fluctuate around a long-term mean.[1] Although some of these models were designed to reproduce time-varying dividend-price ratios that may reproduce the in-sample predictability of stock returns, a reexamination of empirical evidence has led some authors (Goyal and Welch, 2003; Lettau and Ludvigson, 2005; Welch and Goyal, 2007) to question the predictive power of this variable out-of-sample, while others remain convinced of its relevance (Cochrane, 1992, 2007; Campbell and Thompson, 2007). In particular, Lettau and Van Nieuwerburgh (2007) propose to reconcile the inconsistencies between in-sample and out-of-sample results by positing shifts in steady states, which in turn call for new theoretical developments.

Typical asset pricing models assume that investors know exactly the model's parameters and act accordingly. Yet, a hundred years ago equity investors could not rely on the wealth of available data used today to estimate, often imperfectly, the parameters of even the simplest models. Thus, even if past investors had agreed with today's valuation models, they may have required an additional expected return for holding stocks in view of the looming uncertainty of their dynamics – a *learning premium* – which might explain in part their lower valuations at the time.

This paper develops a model in which such a premium can indeed arise. Because the learning premium vanishes by construction, as parameters are gradually revealed, disentangling its value

[1] For example, see Campbell and Shiller (1988), Breen et al. (1989), Fama and French (1993), Glosten et al. (1993), Lamont (1998), Baker and Wurgler (2000), Lettau and Ludvigson (2001), Campbell and Vuolteenaho (2004), Polk et al. (2006), Ang et al. (2007), Binsbergen et al. (2010), Chen et al. (2013), Kelly and Pruitt (2013), Van Binsbergen et al. (2013), Li et al. (2013), Da et al. (2014), and Martin (2013).

from the stationary equity premium is crucial to understand to which extent stocks can reproduce in the future returns that are comparable to their historical averages.

Though it had long been pointed out by Modigliani (1977) and Lucas and Sargent (1981) that the assumption of known parameters was a simplification, to be relaxed at some analytical cost, this area of research has gained momentum only in the past decade, with Hansen (2007) asking explicitly "(a) how can we burden the investors with some of the specification problems that challenge the econometrician, and (b) when would doing so have important quantitative implications".

In a model with multiple macroeconomic states, Johannes et al. (2016) observe that parameter learning improves the model's ability to reproduce stylized facts such as counter-cyclical volatility and expected returns. In a bounded rationality model with long-run risks, Croce et al. (2014) show that limited information generates a downward-sloping equity term structure and a large equity premium. Jagannathan and Liu (2015) develop a latent-variable model for dividend growth which reproduces out-of-sample predictability in stock returns. Collin-Dufresne et al. (2016) show how Bayesian learning becomes a source of long-run risks, and find price-dividend ratios explicitly in the case of unit elasticity of intertemporal substitution.

Learning has a dual impact on asset prices: First, prices experience larger shocks than in models with known parameters. For example, a positive dividend shock generates a proportional shock to prices in the Lucas model, but a more than proportional shock in a model with uncertain dividend growth, as a positive shock also updates the growth rate upwards. Second, investors recognize that each future shock will also affect the value of subsequent shocks, while acknowledging that they will be less informative than present ones of similar magnitude, as parameter uncertainty declines over time. (For example, while estimating the probability of a head in a sequence of coin tosses, the weight of each new outcome in the running estimate declines as the number of tosses increases.)

With the exception of Collin-Dufresne et al. (2016), the literature focuses on the first type of impact, assuming that investors update their beliefs about parameter values, but do not account for future updates in evaluating current prices. Pioneered by Kreps (1998) and further explored by Piazzesi and Schneider (2010) and Cogley and Sargent (2009), such an approach is known as *anticipated utility*, and its main appeal is tractability, as it yields prices obtained by substituting current estimates in the formulas obtained under the assumption of known parameters. The limit of this approach is that it does not reflect the impact on prices of the demand for stocks that stems from hedging against future parameter updates. For example, if a negative dividend shock implies both a lower dividend and lower dividend growth, then stocks are more risky than in a model with known parameters, and investors may change their demand, hence valuation, accordingly.

Exploring parameter learning in a fully rational model quickly leads to some stumbling blocks, as appealing assumptions yield appalling results. The first counterintuitive observation, noted by Veronesi (2000) in a Markov switching model for dividend growth, is that the familiar assumptions of time-additive utility with constant risk aversion do not lead to a learning premium, but a learning *discount* – investors have higher valuations when parameters are uncertain. Intuitively, such a result arises because a fair lottery on a growth rate translates into a *favorable* lottery for its corresponding payoff at long horizon, in view of the convexity that results from compounding the rate over several periods, and implying – counterfactually – a secular decline for the aggregate price-dividend ratio. Brevik and d'Addona (2010) resolve this paradox in the same Markov switching model by reproducing a learning premium through Epstein and Zin (1989) utility with a preference for early resolution of uncertainty.

Importantly, in typical Markov switching models agents strive to estimate an unobservable and ever-fluctuating state of the economy. Thus, learning is *transitory*, as information obtained about the current state becomes less relevant in the future, when reversion to the steady-state distribution unfolds. By contrast, in our model learning is *permanent*, because the unobserved state is constant, and the relevance of information increases over time. As new information is relevant at any future horizon, it has a larger impact on asset prices than if learning were transitory. Indeed, assuming an unknown but constant growth rate with a normal prior, the aforementioned issue of lower price-dividend ratios with learning worsens dramatically – the price-dividend ratio becomes infinite. This surprising phenomenon, which motivates the assumptions made in the paper, is described in detail in the next section, while the appendix describes a model of transitory learning, in which prices are finite.

We specify a discrete-time model for dividend growth in which at each period the dividend can either increase or decrease by fixed factors, as in the binomial model of Cox et al. (1979). The investor's uncertainty is on the relative probability of an upward or downward move, and is resolved over time as dividend growth unfolds. The representative investor is fully rational, hence updates the stock price to reflect both the current probability estimate and its potential future changes.

To identify the posterior distributions at each time, we observe that they coincide with those arising from a Polya-urn scheme, and hence yield posteriors in the Beta-Binomial class. Formally, one can identify ups and downs of the dividend as draws of balls of two different colors from an initial pair of different balls. After each draw, the selected ball is replaced, along another one of the same color, before the next draw.

Then we find in closed form the stock price and its implied equilibrium rate when the representative investor has time-additive utility. In this case, we still observe a counterfactual learning discount, which leads us to investigate the recursive preferences of Epstein and Zin (1989), as to explicitly embed the aversion to later resolution of uncertainty and separate risk aversion from the elasticity of intertemporal substitution.

Although Epstein-Zin preferences do not lead to closed-form solutions for stock prices and interest rates, we find an asymptotic approximation and verify its accuracy first though a convergence result, and then through a direct comparison to the time-additive formula and a numerical computation. Indeed, we find that the Epstein-Zin preferences are able to reproduce realistic values of the price-dividend ratios and their increase over time toward a long-term value that reflects a model with known parameters.

The message of the calibration of the Epstein-Zin model is twofold: First, while the increase in price-dividend ratios over the past ninety years has been significant, the model suggests that, if it is due to the gradual resolution of parameter uncertainty, further increases are likely to continue. Indeed, our calibration implies that the rise of the price-dividend ratio from 30 to 46 over ninety years is consistent with a long-term value of such ratio of 64, i.e., 40% more than 46. Second, the model implies that the resolution of parameter uncertainty, and its consequent increase in the price-dividend ratios, is going to be very slow: even if parameter estimates remained at the current levels fifty years from now, the price-dividend ratio would increase from 46 to 50, well below its long term value. Thus, the model suggests that parameter uncertainty still looms large in asset prices, and that its effects are likely to persist for decades if not centuries.

The rest of this paper is organized as follows: Section 2 motivates the problem and illustrates the pitfalls that ensnare ostensibly natural approaches. Section 3 describes the model and obtains the posterior distribution of parameters. Section 4 contains the main result, which consists of the explicit formulas for equilibrium stock prices and interest rates in the time-additive case, and their asymptotic expansion in the Epsten-Zin case. It then proceeds to discuss their calibration. Concluding remarks are in Section 5. All proofs are in the appendix.

2 Prologue

Consider the familiar Lucas Jr (1978) model, assuming that the dividend process D_t is

$$dD_t = \mu^D D_t dt + \sigma^D D_t dW_t,$$

where μ^D and $\sigma^D > 0$ are two constants. The economy is governed by a representative agent, who maximizes expected power utility from future consumption $E\left[\int_t^\infty \mathrm{e}^{-\beta(s-t)}\, u(C_s)ds\right]$, where $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$ for $\gamma > 0$. (The case $\gamma = 1$ corresponds to logarithmic utility, omitted for brevity, which leads to the same foregoing pricing formulas.) Under the optimality and market-clearing conditions, in this market the asset price S_t and the safe rate r_t are characterized by

$$\begin{aligned}\frac{D_t}{S_t} &= r_t - \mu^D + \gamma(\sigma^D)^2, \\ r_t &= \beta + \gamma\mu^D - \gamma(1+\gamma)\frac{(\sigma^D)^2}{2}.\end{aligned}$$

which imply that both the price-dividend ratio and the safe rate are constant. Of course, a limit of this model is that it assumes that parameters are known with infinite precision from the beginning.

A natural approach to relax this assumption is to assume instead that the dividend growth rate μ^D is learned over time, based on the realizations of the dividend process D_t. Assuming for simplicity that μ^D is a normal variable and that its prior has an independent normal distribution $\mu_0^D \sim N(\mu_0, \sigma_0^2)$, standard filtering results (Liptser and Shiryaev, 2013) yield that the conditional expectation $\mu_t^D = E[\mu^D | (D_s)_{s\le t}]$ is

$$\widehat{\mu^D}_t = \frac{\frac{\mu_0}{\sigma_0^2} + \frac{R_t}{(\sigma^D)^2}}{\frac{1}{\sigma_0^2} + \frac{t}{(\sigma^D)^2}}.$$

where $R_t \triangleq \int_0^t \frac{dD_s}{D_s}$, so that $dR_t = \mu^D dt + \sigma^D dW_t$. Simplifying further, let $\sigma_0 \to \infty$ (which corresponds to a vague prior for μ^D), which yields

$$\widehat{\mu^D}_t = \frac{R_t}{t}, \tag{1}$$

an intuitive formula that identifies the best predictor for the dividend growth rate μ^D as the average of realized growth to date.

In this context, *anticipative utility* assumes that at time t the agent considers the current estimate exact (thereby contradicting himself immediately afterwards by re-estimating the same parameter with additional information). As the estimate is considered final, the same pricing formulas as the Lucas' model apply, with μ^D replaced by its estimate μ_t^D.

To resolve the time-inconsistency of anticipative utility, the agent must remain aware that current estimates are imperfect and hence future dividends will change the estimates of subsequent dividends – even further in the future. In this spirit, denote by $\mathcal{F}_t = \sigma(D_u\colon 0 \le u \le t)$ the observable filtration, equivalently generated by D or R. As $\widehat{\mu^D}_t$ is a martingale in this filtration (by the tower property of conditional expectation), the predictable representation property entails that

$$d\widehat{\mu^D}_t = \frac{\sigma^D}{t} d\widehat{W^D}_t, \tag{2}$$

where $\widehat{W^D}$ is a Brownian Motion under $\mathcal{F}_t$, and the diffusion coefficient σ/t is determined by (1).[2] Thus, the observation dynamics of D is:

$$dD_t = \widehat{\mu^D}_t D_t dt + \sigma^D D_t d\widehat{W^D}_t,$$

the optimality condition prescribes that the state-price density M_t is proportional to the marginal utility of consumption $\mathrm{e}^{-\beta t} D_t^{-\gamma}$, and therefore the asset price is

$$S_t = \frac{1}{M_t}\mathbb{E}\left[\int_t^\infty M_s D_s ds|\mathcal{F}_t\right]. \tag{3}$$

To compute this price, write

$$D_s M_s = D_t M_t \,\mathrm{e}^{-\beta(s-t)+(1-\gamma)\int_t^s \widehat{\mu^D}_u du-(1-\gamma^2)\frac{(\sigma^D)^2}{2}(s-t)+(1-\gamma)\sigma^D\left(\widehat{W^D}_s-\widehat{W^D}_t\right)}.$$

and note that, in view of (2),

$$\begin{aligned}\int_t^s \widehat{\mu^D}_u du &= \int_t^s \left(\widehat{\mu^D}_t + \int_t^u d\widehat{\mu^D}_y\right) du = \widehat{\mu^D}_t(s-t) + \int_t^s \left(\int_t^u d\widehat{\mu^D}_y\right) du \\ &= \widehat{\mu^D}_t(s-t) + \int_t^s \left(\int_y^s du\right) d\widehat{\mu^D}_y = \widehat{\mu^D}_t(s-t) + \sigma^D \int_t^s \frac{s-y}{y} d\widehat{W^D}_y.\end{aligned}$$

where the last equality follows from (2). Therefore,

$$\int_t^s \widehat{\mu^D}_u du + \sigma^D\left(\widehat{W^D}_s - \widehat{W^D}_t\right) = \widehat{\mu^D}_t(s-t) + \sigma^D \int_t^s \frac{s}{y} d\widehat{W^D}_y$$

Thus $\int_t^s \widehat{\mu^D}_u du + \sigma^D\left(\widehat{W^D}_s - \widehat{W^D}_t\right)$ is normally distributed with mean $\widehat{\mu^D}_t(s-t)$ and variance $(\sigma^D)^2\left(\frac{s^2}{t} - s\right)$. Hence,

$$\mathbb{E}\left[M_s D_s|\mathcal{F}_t\right] = D_t M_t \,\mathrm{e}^{\left(-\beta+(1-\gamma)\widehat{\mu^D}_t-(1-\gamma^2)\frac{(\sigma^D)^2}{2}\right)(s-t)+(1-\gamma)^2\frac{(\sigma^D)^2}{2}\left(\frac{s^2}{t}-s\right)}.$$

This equality in turn implies that

$$\begin{aligned}S_t &= \frac{1}{M_t}\mathbb{E}\left[\int_t^\infty M_s D_s ds|\mathcal{F}_t\right] = \frac{1}{M_t}\int_t^\infty \mathbb{E}\left[M_s D_s|\mathcal{F}_t\right] ds \\ &= D_t \int_t^\infty \mathrm{e}^{\left(-\beta+\rho\widehat{\mu^D}_t-(1-\gamma^2)\frac{(\sigma^D)^2}{2}\right)(s-t)+(1-\gamma)^2\frac{(\sigma^D)^2}{2}\left(\frac{s^2}{t}-s\right)} ds.\end{aligned}$$

Crucially, unless $\gamma = 1$ the last integral diverges for any combination of parameters (the factor $e^{(1-\gamma)^2 s^2/t}$ overwhelms all others for s large enough)[3], reflecting that parameter uncertainty vanishes so slowly over time that a forward-looking investor, keen to hedge against potentially low consumption growth, would be willing to pay an arbitrarily high price for the asset, which is the only source or consumption. This finding is reminiscent of the work on parameter uncertainty

[2]In detail, $d\widehat{W^D}_t = \frac{\mu^D - \widehat{\mu^D}_t}{\sigma^D} dt + dW_t$.

[3]The trivial exception is $\gamma = 1$, which leads to $S_t = D_t/\beta$, whence learning has no effects on prices, both with anticipative utility and with rational expectations.

of Pástor and Stambaugh (2012), who note that stocks are *substantially more volatile over long horizons from an investors perspective.*

The message of this negative observation is twofold: first, it shows that anticipative utility is not an accurate approximation of asset prices in a rational-expectations model – the price difference is infinite. Second, it shows that the ostensibly most tractable setting for filtering – a normal prior with normal innovations – implies a level of uncertainty that is fundamentally incompatible with isoelastic preferences.[4]

Thus, the ostensibly natural assumptions of, on one hand, a normal prior with normal updates, and, on the other hand, isoelastic preferences, leads to the unnatural outcome of infinite prices. The model below preserves isoleastic preferences, and faces the challenge of a tractable filtering setting outside of the familiar normality assumptions.

3 Model Definition

The model is based on a Lucas' tree economy with one unit of a risky asset, which yields at time t a perishable dividend D_t that starts at D_0 and follows the discrete-time process:[5]

$$D_t = D_{t-1} e^{\eta + s X_t} \qquad t \in \mathbb{N}, \tag{4}$$

where $\eta \in \mathbb{R}, s > 0$ and $(X_t)_{t \geq 1}$ denotes a sequence of i.i.d. random variables with common Bernoulli distribution with parameter $P \in [0,1]$, that is $\mathbb{P}(X_t = 1|P) = 1 - \mathbb{P}(X_t = 0|P) = P$ for all $t \geq 1$. However, P is unknown to the agent, who initially assumes that $P \sim U[0,1]$, and then gradually learns (i.e. filters) it from the realization of X through Bayesian[6] updating with respect to the natural filtration $\mathcal{F}_t = \sigma(X_1, .., X_t),\ t \geq 0$.

The next lemma finds explicitly the posterior distribution of X_t with respect to its past history. Recall that the beta-binomial distribution $BetaBin(n, \eta, \beta)$ is described by the probability mass function $f(k) = \binom{N}{k} \frac{B(k+\eta, n-k+\beta)}{B(\eta,\beta)}$, $0 \leq k \leq n$, where B is the Beta function (cf. Georgii (2013)). Moreover, recall that the Beta distribution is conjugate to the binomial distribution, and the compound distribution is the Beta-binomial distribution. Put differently, the Beta-binomial is equivalently described as the distribution of $X \sim BetaBin(n, \eta, \beta)$ or as the distribution of $X \sim Bin(n, P)$, where $P \sim Beta(\eta, \beta)$ (Robert, 2007, Chapter 3.3).

Lemma 3.1. *$X_n|\mathcal{F}_{n-1} \sim BetaBin(1, 1 + \sum_{i=1}^{n-1} X_i, n - \sum_{i=1}^{n-1} X_i)$ for $n \geq 1$. Moreover,*

$$\hat{p}_{n-1} = \int_0^1 \mathbb{P}(X_n = 1|p) f_P(p|(X_i)_{1 \leq i \leq n-1}) dp = \frac{1 + \sum_{i=1}^{n-1} X_i}{n+1}.$$

Thus, in the investor's filtration $\sigma((X_i)_{1 \leq i \leq n-1})$ the distribution of X_n is Bernoulli with probability $\hat{p}_{n-1}$, $X_n|(X_i)_{1 \leq i \leq n-1} \sim B(\hat{p}_{n-1})$.

---

[4]A similar but more technical calculation with Epstein-Zin isoelastic preferences confirms that the prices still diverge, except in the case of unit EIS (elasticity of intertemporal substitution) that nests logarithmic utility and implies that $S_t = D_t/\beta$.

[5]Formally, consider a measurable space $(\Omega, \mathcal{F}, \mathbb{P})$ supporting a uniform random variable $P \sim U[0,1]$ and an IID sequence $(X_t)_{t \geq 1}$, $X_t \sim B(P)$, where $B(P)$ denotes the Bernoulli distribution with parameter P. Additionally, define the filtration generated by the observations of X_t. Let $\mathcal{F}_t = \sigma(X_1, .., X_t),\ t \geq 0$, which is the filtration used for Bayesian updating.

[6]This assumption can be relaxed to $P \sim Beta(\eta_0, \beta_0)$, with $\eta_0, \beta_0 > 0$.

The interpretation of the model (4) is as follows. The parameters η and s identify the mean and the standard deviation of the dividend growth rate. Specifically:

$$\mathbb{E}_{n-1}\left[\log\left(\frac{D_n}{D_{n-1}}\right)\right] = \eta + s\hat{p}_{n-1}, \ \ \mathrm{Var}\left(\log\left(\frac{D_n}{D_{n-1}}\right)\Big|(X_i)_{1\leq i\leq n-1}\right) = s^2\hat{p}_{n-1}(1-\hat{p}_{n-1}),$$

where, for brevity, henceforth $\mathbb{E}_t[\cdot]$ denotes conditional expectation with respect to $\mathcal{F}_t$. In view of the previous lemma, the distribution of D_n from the investor's viewpoint is

$$D_n \sim D_0\,\mathrm{e}^{\eta n + sY_n},$$

where $Y_n \sim \mathrm{BetaBin}(n,1,1)$, $n \geq 1$, and in general

$$D_m \sim D_n\,\mathrm{e}^{\eta(m-n)+sY_m^{(n)}},$$

where the superscript n stands for time, so that at time n, D_n, and hence p_n are all known. Thus $Y_m^{(n)} \sim \mathrm{BetaBin}(m-n,(n+2)\hat{p}_n, n-(n+2)p_n+2)$, $m \geq n \geq 0$, so that $Y_n = Y_n^{(0)}$.

Define the (ex-dividend) price S_t, as the price of a security that entitles the holder to the dividend D_s, for all $s > t$. As in the usual Lucas model, a representative agent maximizes expected utility from current and future consumption.

Define the set of admissible consumption plans as $\mathcal{L}_\delta = \left\{C\colon \sum_{t=0}^{\infty}\delta^t\mathbb{E}[C_t] < \infty\right\}$, where the $0<\delta<1$ is the time-preference parameter. The Epstein-Zin utility of consumption at time t with horizon T is defined by the backward recursion

$$U_t^T(C) = \left\{(1-\delta)C_t^{\frac{1-\gamma}{\theta}} + \delta\left(\mathbb{E}_t[(U_{t+1}^T)^{1-\gamma}]\right)^{\frac{1}{\theta}}\right\}^{\frac{\theta}{1-\gamma}}, \qquad U_{T+1}^T = 0, \tag{5}$$

where C is the consumption process and $\theta = \frac{1-\gamma}{1-\rho}$. In addition to δ, the other preference parameters are $0 < \gamma \neq 1$ for risk aversion, and $\psi = \frac{1}{\rho}$ for intertemporal elasticity of substitution, with $0 < \rho \neq 1$.[7] The infinite-horizon Epstein-Zin utility is defined as the limit (which exists by Lemma B.1 below):

$$U_t(C) = \lim_{T\to\infty} U_t^T(C). \tag{6}$$

The agent chooses the consumption C_t and the number of shares ϕ_t to hold in the risky asset at time t, so that the budget equation governing the agent's wealth X_t is

$$X_t = \phi_{t-1}(S_t - S_{t-1}) + (1+r_{t-1})(X_{t-1} - \phi_{t-1}S_{t-1} - C_{t-1})$$

where S_t represents the (ex-dividend) price of the risky asset and r_t the safe rate, which are determined in equilibrium:

Definition 3.2. *An equilibrium is a pair (S_t, r_t) of price and rate processes such that*

(i) *Optimal consumption equals the dividend stream, i.e., $C_t = D_t,\ t \geq 0$;*

(ii) *Wealth equals the risky asset, i.e. $X_t = S_t, \phi_t = 1, t \geq 0$. Thus, the safe position is zero.*

[7] Recall that time-additive power utility with risk aversion γ recovers from the Epstein-Zin setting $\gamma = \rho$ and $\theta = 1$, and using the transformation $V_t = \frac{U_t^{1-\gamma}}{(1-\gamma)(1-\delta)}, \delta = \mathrm{e}^{-\beta}$, whence (5) becomes $V_t = \frac{C_t^{1-\gamma}}{1-\gamma} + \mathrm{e}^{-\beta}\,\mathbb{E}_t\left[V_{t+1}\right]$.

4 Main Results

The main result identifies the price-divide ratio and safe rate in equilibrium over time:

Theorem 4.1. *For $\beta > 0$ large enough, and additive power utility ($\gamma = \rho$ and $\theta = 1$) the price-dividend ratio and interest rate are respectively:*

$$\frac{S_t}{D_t} = \frac{1}{1 - \mathrm{e}^{((1-\gamma)\eta-\beta)}}\, {}_2F_1\left(1, (t+1)\hat{p}_t; t+1; \frac{1-\mathrm{e}^{(1-\gamma)s}}{1-\mathrm{e}^{-((1-\gamma)\eta-\beta)}}\right) - 1, \tag{7}$$

$$r_{t+1,t} = \frac{1}{\mathrm{e}^{-\beta-\eta\gamma}\,\mathbb{E}_t\left[\mathrm{e}^{-\gamma s X_{t+1}}\right]} - 1 = \frac{1}{\delta\,\mathrm{e}^{-\beta-\eta\gamma}\left(1-\hat{p}_t+\hat{p}_t\,\mathrm{e}^{-\gamma s}\right)} - 1, \tag{8}$$

where ${}_2F_1(a,b,c;d)$ is the (ordinary) hypergeometric function.

Epstein-Zin preferences do not lead to analogous closed-form solutions, but it is possible to find expansions around the long-term solution. To this end, define

$$S_t = \bar{c}_t^{\,1-\frac{1}{\psi}} D_t.$$

In other words, the price-dividend ratio is $\bar{c}_t^{\,1-\frac{1}{\psi}}$. Then

Theorem 4.2. *For $\beta > 0$ large enough,*

(i) *The price-dividend ratio and the interest rate are*

$$\frac{S_t}{D_t} = (\bar{c}_t(\hat{p}_t))^{1-\frac{1}{\psi}} = \left(c_\infty^{(t)}(\hat{p}_t)\right)^{1-\frac{1}{\psi}} - 1 + \sum_{i=1}^{\infty} \frac{\alpha_i(\hat{p}_t)}{t^i}, \tag{9}$$

$$r_{t+1,t} = \frac{1}{\delta\,\mathrm{e}^{\eta\left(-\frac{1}{\psi}\right)}\,\mathbb{E}_t\left[\bar{c}_{t+1}^{\frac{1}{\psi}-\gamma}\,\mathrm{e}^{-\gamma s X_{t+1}}\right]\left(\mathbb{E}_t\left[\bar{c}_{t+1}^{1-\gamma}\,\mathrm{e}^{(1-\gamma)sX_{t+1}}\right]\right)^{\frac{\gamma-\frac{1}{\psi}}{1-\gamma}}} - 1, \tag{10}$$

where

$$\left(c_\infty^{(t)}(\hat{p}_t)\right)^{1-\frac{1}{\psi}} = \frac{1}{1-\delta\,\mathrm{e}^{\eta(1-\frac{1}{\psi})}\left(\mathbb{E}_t\left[\mathrm{e}^{(1-\gamma)sX_{t+1}}\right]\right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}}} = \frac{1}{1-\delta\,\mathrm{e}^{\eta(1-\frac{1}{\psi})}\left((1-\hat{p}_t)+\hat{p}_t\,\mathrm{e}^{(1-\gamma)s}\right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}}}.$$

and the coefficients α_i are bounded and admit explicit expressions.

(ii) *In particular, for any $k \geq 0$ the error of the expansion (9) is:*

$$\left|(\bar{c}_t(\hat{p}_t))^{1-\frac{1}{\psi}} - \left(\left(c_\infty^{(t)}(\hat{p}_t)\right)^{1-\frac{1}{\psi}} - 1 + \sum_{i=1}^{k}\frac{\alpha_i(\hat{p}_t)}{t^i}\right)\right| = O\left(\frac{1}{t^{k+1}}\right). \tag{11}$$

Whereas the interest rate satisfies

$$\left|\frac{\left(\mathbb{E}_t\left[\left(c_\infty^{(t+1)}(\hat{p}_{t+1})\right)^{1-\gamma}\mathrm{e}^{(1-\gamma)sX_{t+1}}\right]\right)^{\frac{\frac{1}{\psi}-\gamma}{1-\gamma}}}{\delta\,\mathrm{e}^{\eta\left(-\frac{1}{\psi}\right)}\,\mathbb{E}_t\left[\left(c_\infty^{(t+1)}(\hat{p}_{t+1})\right)^{\frac{1}{\psi}-\gamma}\mathrm{e}^{-\gamma s X_{t+1}}\right]} - 1 - r_{t+1,t}\right| = O\left(\frac{1}{t}\right). \tag{12}$$

Similarly, an error of $O\left(\frac{1}{t^{k+1}}\right), k \geq 1$ is achieved through a higher-order approximation (9).

Quantity	Additive	Epstein-Zin A	Epstein-Zin B	Data
η	-0.0124	-0.0124	-0.0283	
$\beta(\delta)$	0.04 (0.96)	0.04 (0.96)	0.0273 (0.9731)	
γ	1.37	1.37	9.53	
p	0.0104	0.0104	0.4408	
s	3.118	3.118	0.133	
ψ	0.73	15	8.81	
Price Dividend ratio (1926)	26.1	24.6	29.9	22.5
Price Dividend ratio (2016)	24.5	28.4	45.9	38.7
Price Dividend ratio (long term)	24.1	30.1	64.2	
Average dividend growth	2%	2%	3%	1.2%
St. dev. of dividend growth	30%	30%	6%	11.1%

Table 1: Parameter calibrations with additive utility and Epstein-Zin preferences, with (A) recalibration of the elasticity of intertemporal substitution ψ, and (B) recalibration of all parameters. Estimates in the last column are from CRSP and Beeler et al. (2012).

Table 1 brings to life the above result through the calibration to realistic parameter values. In Table 1, the parameter combinations Additive and Epstein-Zin B are obtained by calibrating the model as to minimize the sum of squared errors of the empirical quantities in the last column, while the combination Epstein-Zin A is obtained by minimizing the sum of squared errors by varying only the EIS parameter ψ, while keeping the other ones equal to the Additive column. The left panel in Figure 2 is obtained from the closed form solution (7), whereas the right panel plots a numerical solution of the price-dividend ratio in (42), resulting from (5). For the left panel, parameters are as in the additive utility column of Table 1, whereas the right panel uses the parameters in the column Epstein-Zin A. Likewise, Figure 3 is obtained from the approximations in (9), combined with a numerical solution of the recursive equation for the price-dividend ratio, using the parameters in column Epstein-Zin B.

The central question is whether the model is able to reproduce the secular increase in the price-dividend ratio with realistic preference parameters, while also remaining consistent with the typical moments of aggregate dividend growth.

Both Table 1 and the left panel in Figure 2 show that additive utility does not lead to an *increase*, but to a slight *decrease* in the price-dividend ratio, as the representative agent responds to more uncertainty in assets' returns by bidding their prices up to hedge against potentially low growth. Thus, additive power utility generates a *learning discount* (parameter uncertainty increases prices) that recedes over time. An additional limit of the calibration with power utility is the extreme standard deviation of dividend growth.

A parsimonious attempt at keeping the same calibration parameters as additive utility, while optimizing the value of the elasticity of intertemporal substitution ψ, improves the results qualitatively, but not quantitatively, leading to a modest increase in the price-dividend ratio from 24.6 to 28.4, as shown from the calibration Epstein-Zin A.

A marked improvement takes place by reestimating all model parameters, as in the calibration Epstein-Zin B, which leads to higher risk aversion and lower elasticity of intertemporal substitution. Such calibration is both able to reproduce a realistic increase in the price-dividend ratio from 30 to 46 over a span of 90 years, while also generating a standard deviation of dividend growth more aligned to the data.

This calibration also suggests that, even after 90 years of observations, the learning premium

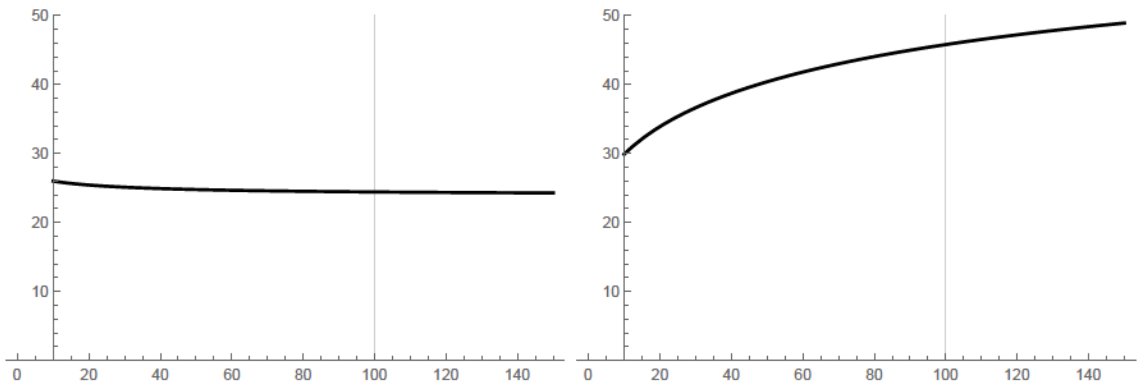

Figure 2: Price-dividend ratio over time, with time-additive, power utility (left) and with Epstein-Zin utility (right). Parameters are as in Table 1 (B for the right panel with Epstein-Zin).

has not vanished – parameter uncertainty still looms large. Indeed, the model implies that, if current parameter estimates were indeed perfectly accurate, by the time that their accuracy were resolved, the price-dividend ratio would have risen from 46 to 64, an increase of nearly 40%. Put differently, parameter uncertainty implies that current stock prices are nearly 30% lower that they would be if current parameter estimates could be trusted with absolute certainty.

Furthermore, the right panel in Figure 2 shows that such uncertainty is likely to be resolved very slowly: even if parameter estimates remained at their current levels, the price-dividend ratio would increase from 46 to 50 in the next fifty years, i.e., less than a quarter of its potential increase from 46 to 64.

Finally, as Epstein-Zin utility, in contrast to power utility, does not lead to closed-form solutions, Figure 3 displays the convergence of the approximation in Theorem 4.2 with the parameter values in the Epstein-Zin A row of Table 1. The convergence of the approximations in equation (9) to the numerical solution is significantly faster than the convergence of the price-dividend to its asymptotic value: under the current availability of accurate stock price data, which is longer than 90 years, the first-order approximation is within 10% of its numerical target. Yet, in the interest of precision, Table 1 reports figures obtained numerically.

5 Conclusion

This paper explores the extent to which the resolution of parameter uncertainty explains the secular increase in price-dividend ratios. Contrary to intuition, standard models of time-additive utility imply that stock prices are higher in the presence of uncertainty, and hence that price-dividend ratios counterfactually decline over time.

Epstein-Zin preferences are able to explain the empirical increase in price-dividend ratios with risk aversion near 10 and elasticity of intertemporal substitution near 9. Out of sample, the model implies that the current price-dividend ratio still reflects significant parameter uncertainty and that, even if current parameter estimates are close to their true value, in the future the price-dividend ratio may still increase considerably.

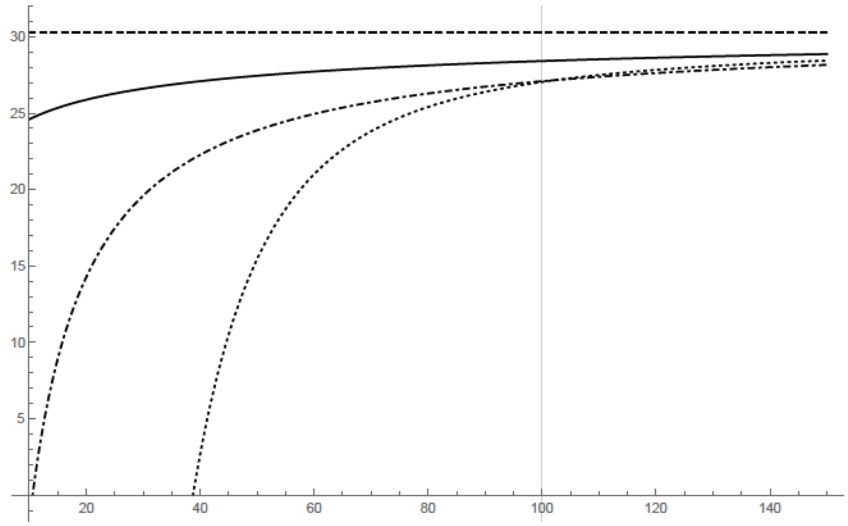

Figure 3: Left: Price-dividend ratio over time for Epstein-Zin utility: numerical solution (solid), zeroth order approximation (dashed), first order approximation (dash dotted), second order approximation (dotted) .

A Transitory vs. Permanent learning

This subsection demonstrates the difference between transitory and permanent learning by examining in detail the asset pricing implications of a model of transitory learning, and by contrasting them with the ones obtained in the prologue for permanent learning.

Consider the case of an unobservable dividend drift that follows an Ornstein-Uhlenbeck process with known coefficients. As before, the dividends themselves are still observable and can be used to estimate the current drift. Let the dividends again grow geometrically, i.e.,

$$dD_t = \mu_t D_t dt + \sigma_D D_t dW_t.$$

However, let now the growth rate μ_t follow a hidden Ornstein-Uhlenbeck process

$$d\mu_t = \kappa(\bar{\mu} - \mu_t)dt + \sigma_\mu dB_t,$$

which means that μ_t fluctuates around its long-term mean $\bar{\mu}$. Denoting by

$$\begin{aligned} R_t &= \int_0^t \frac{dD_s}{D_s} - \int_0^t \int_0^u \bar{\mu}\kappa\, \mathrm{e}^{-\kappa(u-s)}\, ds du = \int_0^t \frac{dD_s}{D_s} + \bar{\mu}t + \frac{\bar{\mu}}{\kappa}\left(\mathrm{e}^{-\kappa t} - 1\right) \\ \theta_t &= \mu_t - \bar{\mu}\kappa \int_0^t \mathrm{e}^{-\kappa(t-s)}\, ds, \end{aligned}$$

it follows that

$$\begin{aligned} dR_t &= \theta_t dt + \sigma_D dW_t, \\ d\theta_t &= -\kappa\theta_t dt + \sigma_\mu dB_t. \end{aligned}$$

Moreover, $\mathcal{F}_t^D = \sigma\left((D_u)_{0\le u\le t}\right) = \sigma\left((R_u)_{0\le u\le t}\right) = \mathcal{F}_t^R$. The Kalman-Bucy filter $\hat{\theta}_t = \mathbb{E}\left[\theta_t | \mathcal{F}_t^R\right]$

and its variance $\gamma(t) = \mathbb{E}[(\hat{\theta}_t - \theta_t)^2]$ satisfy

$$d\gamma(t) = \left(-2\kappa\gamma(t) - \frac{(\gamma(t))^2}{\sigma_D^2} + \sigma_\mu^2\right) dt, \tag{13}$$
$$d\hat{\theta}_t = -\kappa\hat{\theta}_t dt + \frac{\gamma(t)}{\sigma_D^2}\left(dR_t - \hat{\theta}_t dt\right).$$

Let $\gamma_\pm = -\kappa\sigma_D^2 \pm \sigma_D\sqrt{\kappa^2\sigma_D^2 + \sigma_\mu^2}$, be the two roots of the quadratic of the right hand side of (13). Assuming again that $\theta_0 \sim \mathbb{N}(\mu_0, \sigma_0^2)$, and setting $\hat{\theta}_0 = \mu_0, \gamma(0) = \sigma_0^2$, then the solution to the Kalman-Bucy filter is

$$\gamma(t) = \frac{\gamma_- - \gamma_+ \frac{\gamma_0^2 - \gamma_-}{\gamma_0^2 - \gamma_+} \mathrm{e}^{\frac{\gamma_+ - \gamma_-}{\sigma_D^2} t}}{1 - \frac{\gamma_0^2 - \gamma_-}{\gamma_0^2 - \gamma_+} \mathrm{e}^{\frac{\gamma_+ - \gamma_-}{\sigma_D^2} t}},$$
$$\hat{\theta}_t = \mathrm{e}^{-\int_0^t \left(\kappa + \frac{1}{\sigma_D^2}\gamma(s)\right) ds} \hat{\theta}_0 + \frac{1}{\sigma_D^2}\int_0^t \mathrm{e}^{-\int_s^t \left(\kappa + \frac{1}{\sigma_D^2}\gamma(u)\right) du} \gamma(s) dR_s.$$

and the Brownian Motion under $\mathcal{F}_t^R$ is $\widehat{W^D}$, defined as

$$d\widehat{W^D}_t = \frac{dR_t - \hat{\theta}_t dt}{\sigma_D} = \frac{\theta_t - \hat{\theta}_t}{\sigma_D} dt + dW_t,$$

so that

$$d\hat{\theta}_t = -\kappa\hat{\theta}_t dt + \frac{\gamma(t)}{\sigma_D} d\widehat{W^D}_t.$$

Thus

$$\frac{dD_t}{D_t} = dR_t + \int_0^t \bar{\mu}\kappa\, \mathrm{e}^{-\kappa(t-s)} ds dt = dR_t + \bar{\mu}\left(1 - \mathrm{e}^{-\kappa t}\right) dt \tag{14}$$
$$= \left(\hat{\theta}_t + \bar{\mu}\left(1 - \mathrm{e}^{-\kappa t}\right)\right) dt + \sigma_D d\widehat{W^D}_t.$$

Recall the price process in (3), with the state-price density M_t is proportional to the marginal utility of consumption $\mathrm{e}^{-\beta t} D_t^{-\gamma}$. Thus

$$D_t M_t = D_s M_s \, \mathrm{e}^{-\beta(t-s) + (1-\gamma)\left(\int_s^t \hat{\theta}_u du + \bar{\mu}(t-s) + \frac{\bar{\mu}}{\kappa}\left(\mathrm{e}^{-\kappa(t-s)} - 1\right) - \frac{\sigma_D^2}{2}(t-s) + \sigma_D\left(\widehat{W^D}_t - \widehat{W^D}_s\right)\right)}.$$

Note that, for $s \le t$, it holds that

$$\theta_t - \theta_s = -\kappa\int_s^t \hat{\theta}_u du + \int_s^t \frac{\gamma(u)}{\sigma_D} d\widehat{W^D}_u,$$

which, combined with $\hat{\theta}_t = \hat{\theta}_s \mathrm{e}^{-\kappa(t-s)} + \int_s^t \frac{\gamma(u)}{\sigma_D} \mathrm{e}^{-\kappa(t-u)} d\widehat{W^D}_u$, yields

$$\int_s^t \hat{\theta}_u du = -\frac{\theta_t - \theta_s}{\kappa} + \frac{1}{\kappa\sigma_D}\int_s^t \gamma(u) d\widehat{W^D}_u$$
$$= \frac{1 - \mathrm{e}^{-\kappa(t-s)}}{\kappa}\hat{\theta}_s - \frac{1}{\kappa\sigma_D}\int_s^t \gamma(u)(\mathrm{e}^{-\kappa(t-u)} - 1) d\widehat{W^D}_u.$$

Hence,

$$\int_s^t \hat{\theta}_u du + \sigma_D \left(\widehat{W^D}_t - \widehat{W^D}_s\right) = \frac{1 - \mathrm{e}^{-\kappa(t-s)}}{\kappa}\hat{\theta}_s - \frac{1}{\kappa\sigma_D}\int_s^t \left(\gamma(u)(\mathrm{e}^{-\kappa(t-u)} - 1) - \kappa\sigma_D^2\right) d\widehat{W^D}_u. \tag{15}$$

Therefore,

$$\int_s^t \hat{\theta}_u du + \sigma_D \left(\widehat{W^D}_t - \widehat{W^D}_s\right) \sim N\left(\frac{1 - \mathrm{e}^{-\kappa(t-s)}}{\kappa}\hat{\theta}_s, \frac{1}{(\kappa\sigma_D)^2}\int_s^t \left(\gamma(u)(1 - \mathrm{e}^{-\kappa(t-u)}) + \kappa\sigma_D^2\right)^2 du\right).$$

As in the long run $\gamma(u)$ converges to γ_+, it follows that for large s, t

$$\int_s^t \hat{\theta}_u du + \sigma_D \left(\widehat{W^D}_t - \widehat{W^D}_s\right) \sim N\left(\frac{1 - \mathrm{e}^{-\kappa(t-s)}}{\kappa}\hat{\theta}_s, H(t-s)\right),$$

where

$$H(\tau) = \frac{\gamma_+^2}{(\kappa\sigma_D)^2}\frac{\gamma_+^2\left(4\,\mathrm{e}^{-\kappa\tau} - \mathrm{e}^{-2\kappa\tau} + 2\kappa\tau - 3\right) - 4\gamma_+\kappa\sigma^2\left(\kappa\tau - \mathrm{e}^{-\kappa\tau} + 1\right) + 2\kappa^3\sigma^4\tau}{2k}.$$

Hence,

$$\begin{aligned}\mathbb{E}\left[M_t D_t | \mathcal{F}_s\right] = D_s M_s \exp&\left\{-\left(\beta - (1-\gamma)\left(\bar{\mu} - \frac{\sigma_D^2}{2}\right)\right)(t-s)\right\} \\ &\times \exp\left\{(1-\gamma)\left(\frac{\bar{\mu}}{\kappa}\left(\mathrm{e}^{-\kappa(t-s)} - 1\right) + \frac{1 - \mathrm{e}^{-\kappa(t-s)}}{\kappa}\hat{\theta}_s\right) + (1-\gamma)^2\frac{H(t-s)}{2}\right\}.\end{aligned}$$

This equality in turn implies that

$$\begin{aligned}S_t &= \frac{1}{M_t}\mathbb{E}\left[\int_t^\infty M_s D_s ds | \mathcal{F}_t\right] = \frac{1}{M_t}\int_t^\infty \mathbb{E}\left[M_s D_s | \mathcal{F}_t\right] ds \\ &= D_t \int_t^\infty \mathrm{e}^{-\left(\beta - (1-\gamma)\left(\bar{\mu} - \frac{\sigma_D^2}{2}\right)\right)(s-t) + (1-\gamma)\left(\frac{\bar{\mu}}{\kappa}\left(\mathrm{e}^{-\kappa(s-t)} - 1\right) + \frac{1 - \mathrm{e}^{-\kappa(s-t)}}{\kappa}\hat{\theta}_t\right) + (1-\gamma)^2 H(s-t)} ds < \infty.\end{aligned}$$

As $\kappa > 0$, the expression $H(\tau)$ grows at most linearly in τ. As a result, for $\beta > 0$ large enough, the above expression is finite. (There is no closed form solution, even in the stationary case $\gamma(u) = \gamma_+$ for all $u > 0$.) Thus, in contrast to the setting of *permanent* learning, described in the main text, this model of transitory learning gives rise to finite prices, at least for sufficiently large discount rates.

The same argument carries over to Epstein-Zin preferences. Denoting the aggregator by

$$\bar{f}(c, v) = \frac{\beta}{\rho}\frac{c^\rho - (\eta v)^{\frac{\rho}{\eta}}}{(\eta v)^{\frac{\rho}{\eta} - 1}}, \tag{16}$$

so that the indirect utility V_t satisfies

$$dV_t = -\bar{f}(C_t, V_t)dt + \bar{\sigma}_v(t)d\widehat{W^D}_t.$$

In order to find $\bar{\sigma}_v(t)$, recall that

$$dV_t = \left(-f(C_t, V_t) - \frac{1}{2}A(v)\sigma_v^2(t)\right)dt + \sigma_v(t)d\widehat{W^D}_t,$$

where

$$f(c,v) = \frac{\beta}{\rho}\frac{c^\rho - v^\rho}{v^{\rho-1}},\ A(v) = \frac{\eta-1}{v}.$$

As the only source of randomness of the utility comes from the consumption, and both the consumption and the utility processes are linear in consumption, $\sigma_v(t) = \sigma^D V_t$. The transformation to an equivalent normalized utility process is $\bar{U} = U \circ \phi$, where $\phi(v) = \int \mathrm{e}^{\int A(x)dx}\,dv$, which in this case, is $\phi(v) = \frac{v^\eta}{\eta}$. From Itô's formula, it follows that

$$\begin{aligned}
\bar{f}(c,\phi(v)) &= f(c,v)\phi'(v),\\
\bar{\sigma}_{\phi(v)}(t) &= \sigma_v(t)\phi'(v),\\
\bar{A}(\phi(v)) &= A(v)\phi'(v) - \phi''(v).
\end{aligned}$$

It then follows that $\bar{f}$ indeed equals (16), $\bar{A} = 0$, and

$$\bar{\sigma}_v(t) = \eta\sigma^D V_t. \tag{17}$$

As dividends coincide with consumption, i.e. $C_t = D_t$, the state-price deflator M_t is (Duffie and Epstein, 1992)

$$\begin{aligned}
M_t =& \exp\left\{\int_0^t \bar{f}_v(D_s, V_s)ds\right\}\bar{f}_c(D_t, V_t)\\
=& \beta \exp\left\{\frac{\beta}{\rho\eta^{\frac{\rho}{\eta}-1}}\left(1-\frac{\rho}{\eta}\right)\int_0^t \frac{D_s^\rho}{V_s^{\frac{\rho}{\eta}}}ds - \beta\frac{\eta}{\rho}t\right\}\frac{D_t^{\rho-1}}{(\eta V_t)^{\frac{\rho}{\eta}-1}}.
\end{aligned}$$

Using the fact that

$$d\left(V_t^{1-\frac{\rho}{\gamma}}\right) = -\left(1-\frac{\rho}{\gamma}\right)V_t^{1-\frac{\rho}{\gamma}}\left(\frac{\bar{f}(D_t,V_t)dt - \bar{\sigma}_v(t)d\widehat{W^D}_t}{V_t} + \frac{\rho}{2\gamma}\frac{\bar{\sigma}_v^2(t)}{V_t^2}dt\right),$$

it follows that

$$\begin{aligned}
d\left(\frac{D_t^{\rho-1}}{(\gamma V_t)^{\frac{\rho}{\gamma}-1}}\right) =& \frac{D_t^{\rho-1}}{(\gamma V_t)^{\frac{\rho}{\gamma}-1}}(\rho-1)\left(\left(\hat{\theta}_t - \bar{\mu}\left(1-\mathrm{e}^{-\kappa t}\right)\right)dt + \sigma^D d\widehat{W^D}_t + \frac{1}{2}(\rho-2)\left(\sigma^D\right)^2 dt\right)\\
&- \frac{D_t^{\rho-1}}{(\gamma V_t)^{\frac{\rho}{\gamma}-1}}\left(1-\frac{\rho}{\gamma}\right)\left(\frac{\bar{f}(D_t,V_t)dt - \bar{\sigma}_v(t)d\widehat{W^D}_t}{V_t} + \frac{\rho}{2\gamma}\frac{\bar{\sigma}_v^2(t)}{V_t^2}dt\right)\\
&+ (\rho-1)\frac{D_t^{\rho-1}}{(\gamma V_t)^{\frac{\rho}{\gamma}-1}}\left(1-\frac{\rho}{\gamma}\right)\frac{\sigma^D\bar{\sigma}_v(t)}{V_t}dt.
\end{aligned}$$

Therefore,

$$\begin{aligned}\frac{dM_t}{M_t} &= (\rho-1)\left(\left(\hat\theta_t - \bar\mu\left(1-\mathrm{e}^{-\kappa t}\right)\right)dt + \sigma^D d\widehat{W^D}_t + \frac{1}{2}(\rho-2)\left(\sigma^D\right)^2 dt\right)\\ &\quad - \left(1-\frac{\rho}{\gamma}\right)\left(\frac{\bar f(D_t,V_t)dt - \bar\sigma_v(t)d\widehat{W^D}_t}{V_t} + \frac{\rho}{2\gamma}\frac{\bar\sigma_v^2(t)}{V_t^2}dt\right)\\ &\quad + (\rho-1)\left(1-\frac{\rho}{\gamma}\right)\frac{\sigma^D\bar\sigma_v(t)}{V_t}dt + \left(\frac{\beta}{\rho\gamma^{\frac{\rho}{\gamma}-1}}\left(1-\frac{\rho}{\gamma}\right)\frac{D_t^\rho}{V_t^{\frac{\rho}{\gamma}}} - \beta\frac{\gamma}{\rho}\right)dt.\end{aligned}$$

and, substituting (16) and (17), yields

$$\begin{aligned}\frac{dM_t}{M_t} &= -r_t dt + (\rho-1)\sigma^D d\widehat{W^D}_t + \left(1-\frac{\rho}{\gamma}\right)\frac{\bar\sigma_v(t)}{V_t}d\widehat{W^D}_t\\ &= -r_t dt + (\gamma-1)\sigma^D d\widehat{W^D}_t,\end{aligned}$$

where

$$r_t = \beta + (1-\rho)\left(\hat\theta_t - \bar\mu\left(1-\mathrm{e}^{-\kappa t}\right)\right) + (2-\rho)(\gamma-1)\frac{(\sigma^D)^2}{2}.$$

Therefore

$$M_t = M_s \exp\left\{-\int_s^t r_u du - \frac{(\gamma-1)^2}{2}(\sigma^D)^2(t-s) + (\gamma-1)\sigma^D\left(\widehat{W^D}_t - \widehat{W^D}_s\right)\right\}. \qquad (18)$$

From (14),

$$D_t = D_s \exp\left\{\int_s^t \hat\theta_u du + \bar\mu(t-s) + \frac{\bar\mu}{\kappa}\left(\mathrm{e}^{-\kappa(t-s)} - 1\right) - \frac{(\sigma_D)^2}{2}(t-s) + \sigma_D\left(\widehat{W^D}_t - \widehat{W^D}_s\right)\right\}.$$

Together with (18) it now follows that

$$\begin{aligned}D_t M_t = D_s M_s &\exp\left\{-\beta(t-s) + \rho\left(\int_s^t \hat\theta_u du + \bar\mu(t-s) + \frac{\bar\mu}{\kappa}\left(\mathrm{e}^{-\kappa(t-s)} - 1\right)\right)\right\}\\ &\times \exp\left\{\left(\rho(\gamma-1)-\gamma^2\right)\frac{(\sigma^D)^2}{2}(t-s) + \gamma\sigma^D\left(\widehat{W^D}_t - \widehat{W^D}_s\right)\right\}.\end{aligned}$$

Thus, similarly to (15),

$$\int_s^t \hat\theta_u du + \frac{\gamma}{\rho}\sigma_D\left(\widehat{W^D}_t - \widehat{W^D}_s\right) \sim N\left(\frac{1-\mathrm{e}^{-\kappa(t-s)}}{\kappa}\hat\theta_s, H_1(s,t)\right),$$

where

$$H_1(s,t) = \frac{1}{\left(\kappa\sigma_D\right)^2}\int_s^t \left(\gamma(u)(1-\mathrm{e}^{-\kappa(t-u)}) + \kappa\frac{\gamma}{\rho}\sigma_D^2\right)^2 du.$$

Thus

$$\mathbb{E}\left[M_t D_t|\mathcal{F}_s\right] = D_s M_s \exp\left\{-\beta(t-s) + \rho\left(\bar{\mu}(t-s) + \frac{\bar{\mu}}{\kappa}\left(\mathrm{e}^{-\kappa(t-s)} - 1\right)\right)\right\}$$
$$\times \exp\left\{\rho\frac{1-\mathrm{e}^{-\kappa(t-s)}}{\kappa}\hat{\theta}_s + \left(\rho(\gamma-1)-\gamma^2\right)\frac{(\sigma^D)^2}{2}(t-s) + \rho^2\frac{H_1(s,t)}{2}\right\}.$$

We now calculate

$$S_t = \frac{1}{M_t}\mathbb{E}\left[\int_t^\infty M_s D_s ds|\mathcal{F}_t\right] = \frac{1}{M_t}\int_t^\infty \mathbb{E}\left[M_s D_s|\mathcal{F}_t\right] ds$$
$$= D_t \int_t^\infty \mathrm{e}^{-\beta(s-t)+\rho\left(\bar{\mu}(s-t)+\frac{\bar{\mu}}{\kappa}\left(\mathrm{e}^{-\kappa(s-t)}-1\right)\right)+\rho\frac{1-\mathrm{e}^{-\kappa(s-t)}}{\kappa}\hat{\theta}_t+\left(\rho(\gamma-1)-\gamma^2\right)\frac{(\sigma^D)^2}{2}(s-t)+\rho^2\frac{H_1(t,s)}{2}}\, ds.$$

Again, for the same reasons as in the additive utility case above, this expression is finite for a discount rate β large enough, which confirms the claim that prices remain finite even for Epstein-Zin preferences.

B Proofs

Proof of Lemma 3.1. For $n = 1$, it follows from the definition of X_1 that its distribution is $X_1 \sim BetaBin(1, 1, 1)$. For $n > 1$, we calculate the posterior distribution. Recall that

$$f_P(p|X_1, ..., X_{n-1}) \propto L(p) f_{p_0}(p) \propto p^{\sum_{i=1}^{n-1} X_i}(1-p)^{n-1-\sum_{i=1}^{n-1} X_i}, \tag{19}$$

where f_P is the pdf of P and L is the log-likelihood. Hence, $P|X_1, ..., X_{n-1} \sim Beta(\sum_{i=1}^{n-1} X_i + 1, n-1-\sum_{i=1}^{n-1} X_i + 1)$, and thus $X_n|X_1, .., X_{n-1} \sim BetaBin(1, \sum_{i=1}^{n-1} X_i + 1, n-1-\sum_{i=1}^{n-1} X_i + 1)$. Moreover, given the observations $X_1, ..., X_{n-1}$ and using (19), for $n \geq 1$ it holds that

$$\hat{p}_{n-1} = \mathbb{P}(X_n = 1|X_1, ..., X_{n-1}) = \int_0^1 \mathbb{P}(X_n = 1|p) f_P(p|X_1, ..., X_{n-1}) dp$$
$$= \int_0^1 p \frac{p^{\sum_{i=1}^{n-1} X_i}(1-p)^{n-1-\sum_{i=1}^{n-1} X_i}}{Beta(\sum_{i=1}^{n-1} X_i + 1, (n-1) - \sum_{i=1}^{n-1} X_i + 1)} dp = \frac{\sum_{i=1}^{n-1} X_i + 1}{n+1}.$$

□

We formulate this section for the general case of Epstein-Zin utility. The case of power utility corresponds to $\theta = 1$. First, we show the Epstein-Zin utility is well defined, i.e., the infinite-horizon limit in (6) exists. See also (Pennesi, 2018, Theorem 1) for a related result.

Lemma B.1. *Fix an admissible consumption* $C \in \mathcal{L}_\delta$. *Then*

$$U_t^N(C) \leq (1-\delta)^{\frac{1}{1-\rho}} \sum_{n=t}^{N-1} \delta^{n-t} \mathbb{E}_t[C_n]. \tag{20}$$

It follows that the limit in (6) is well defined (hence so is $U_t(C)$*). Moreover, such* $U_t(C)$ *is the unique solution to the recursive equation*

$$U_t(C) = \left\{(1-\delta) C_t^{\frac{1-\gamma}{\theta}} + \delta\left(\mathbb{E}_t[(U_{t+1})^{1-\gamma}]\right)^{\frac{1}{\theta}}\right\}^{\frac{\theta}{1-\gamma}} \tag{21}$$

with the asymptotic condition

$$\lim_{n\to\infty} U_t\left(C^{0,n}\right) = U_t(C), \tag{22}$$

where for any consumption streams $\tilde{C}, \hat{C}$, the modified process $\hat{C}^{\tilde{C},n}$ is defined as

$$\hat{C}_s^{\tilde{C},n} = \begin{cases} \hat{C}_s & : s \le n, \\ \tilde{C}_s & : s > n. \end{cases} \tag{23}$$

Proof of Lemma B.1. Fix any $N \ge t$. Then $U_N^N(C) = 0$. Similarly, if $N \ge 1$, then $U_{N-1}^N(C) = (1-\delta)^{\frac{1}{1-\rho}} \mathbb{E}_{N-1}[C_{N-1}]$.

By (backward) induction, assume that (20) is true for $t = k+1$, and show it for $t = k$. The induction assumption and Jensen's inequality imply that

$$\mathbb{E}_k\left[\left(U_{k+1}^N(C)\right)^{1-\gamma}\right]^{\frac{1}{1-\gamma}} \le (1-\delta)^{\frac{1}{1-\rho}} \sum_{n=k+1}^{N-1} \delta^{n-(k+1)} \mathbb{E}_k[C_n].$$

Then

$$\begin{aligned} U_k^N(C) &= \left\{(1-\delta)C_k^{1-\rho} + \delta\left(\mathbb{E}_k[(U_{k+1}^N)^{1-\gamma}]\right)^{\frac{1-\rho}{1-\gamma}}\right\}^{\frac{1}{1-\rho}} \\ &\le \left((1-\delta)C_k^{1-\rho} + \delta(1-\delta)\left(\sum_{n=k+1}^{N-1} \delta^{n-(k+1)}\mathbb{E}_k[C_n]\right)^{1-\rho}\right)^{\frac{1}{1-\rho}} \le (1-\delta)^{\frac{1}{1-\rho}} \sum_{n=k}^{N-1} \delta^{n-k}\mathbb{E}_k[C_n], \end{aligned}$$

where the first inequality follows from the induction step, and the second from Jensen's inequality, proving the induction step. It follows that the limit in (6) is well defined, as $\{U_t^N\}$ for fixed t is an increasing sequence in $N \ge t$. Hence, $U_t^N(C)$ is well defined for every N, and thus so is its limit $U_t(C)$ in (6).

Additionally, (21) now follows by continuity, after taking the limit $N \to \infty$ in $U_t^N(C) = \left\{(1-\delta)C_t^{\frac{1-\gamma}{\theta}} + \delta\left(\mathbb{E}_t[(U_{t+1}^N)^{1-\gamma}]\right)^{\frac{1}{\theta}}\right\}^{\frac{\theta}{1-\gamma}}$. Whereas the uniqueness of the solution follows from the uniqueness of U_t^N and the fact that $U_t(C^{0,N}) = U_t^N(C)$.

□

Next, set

$$m_{t+1,t} = \delta\left(\frac{D_{t+1}}{D_t}\right)^{\frac{1-\gamma}{\theta}-1} \left(\frac{U_{t+1}(D)}{\left(\mathbb{E}_t\left[U_{t+1}^{1-\gamma}(D)\right]\right)^{\frac{1}{1-\gamma}}}\right)^{-\frac{(1-\gamma)(1-\theta)}{\theta}}, \ m_{t,s} = \prod_{i=s+1}^{t} m_{i,i-1} \text{ for } t > s. \tag{24}$$

Define

$$\pi_{t,t} = \frac{\partial U_t}{\partial C_t}\Big|_{C=D} = \frac{\theta}{1-\gamma}(1-\delta)\frac{1-\gamma}{\theta}U_t(D)^{\frac{\theta}{1-\gamma}-1} D_t^{\frac{1-\gamma}{\theta}-1} = (1-\delta)U_t(D)^{\rho} D_t^{-\rho}, \tag{25}$$

$$\pi_{s,t} = m_{s,t}\pi_{t,t} \text{ for } s > t. \tag{26}$$

In view of (26), optimizing over the consumption at time $s \geq t$ without any constraint on initial wealth, leads to the problem

$$\max_{C\geq 0, C_u = D_u, u\neq s} \{U_t(C) - \mathbb{E}_t[\pi_{s,t}C_s]\} = U_t(D) - \mathbb{E}_t[\pi_{s,t}D_s]. \tag{27}$$

For convenience, denote P_t the cum-dividend price, defined as

$$P_t = S_t + D_t = \mathbb{E}_t\left[\sum_{n=t}^{\infty} m_{n,t}D_n\right]. \tag{28}$$

To complete the description of the market, define the price at time t of a bond maturing at $t+1$ as

$$B(t, t+1) = \mathbb{E}_t[m_{t+1,t}], \tag{29}$$

and the interest rate as

$$B(t, t+1) = \frac{1}{1 + r_{t+1,t}}. \tag{30}$$

Next, to define an equilibrium in this market it remains to define admissible consumption plans. Let X_t be total the wealth of the representative agent at time t (before any consumption takes place).

Definition B.2. *The wealth process X starting from time t_0 is admissible, if $X_t \geq 0$ for all times $t \geq t_0$. For a given consumption stream $C_t \geq 0,\ t = t_0, t_0 + 1, \ldots$ set value of the consumption stream starting from time $t \geq t_0$ as*

$$W_{t_0}(C) = \sum_{s=t_0}^{\infty} \mathbb{E}_{t_0}[m_{s,t_0}C_s].$$

Lemma B.3. *Let $t_0 \geq 0$, then for any admissible consumption $C_s,\ s \geq t_0$,*

$$\mathbb{E}_{t_0}[m_{T+1,t_0}X_{T+1}] = X_{t_0} - \sum_{t=t_0}^{T} \mathbb{E}_{t_0}[m_{t,t_0}C_t], \tag{31}$$

$$\text{and} \qquad X_{t_0} \geq \sum_{t=t_0}^{\infty} \mathbb{E}_{t_0}[m_{t,t_0}C_t]. \tag{32}$$

Moreover, if $X_{t_0} = P_{t_0}$ any admissible consumption C is dominated by D, in that $W_{t_0}(C) \leq W_{t_0}(D)$.

Proof. At any time, the agent can invest in two assets, the bond and the stock. Assume that at time at time t the portfolio is valued at X_t. The dividend is paid out first. Then the portfolio can be rebalanced, to include ϕ_t shares of stock and ψ_t cash. Thus $X_t = \phi_t(P_t - D_t) + \psi_t$, since the stock price P_t is cum-dividend, and whence $\psi_t = X_t - \phi_t(P_t - D_t)$. After which the consumption C_t happens. Then at the next period $t+1$, the portfolio is worth X_{t+1}, which is comprised of $\phi_t P_{t+1}$ wealth invested in stock and $(\psi_t - C_t)(1 + r_t)$ cash, i.e.,

$$X_{t+1} = \phi_t P_{t+1} + (\psi_t - C_t)(1 + r_t) = \phi_t P_{t+1} + (X_t - \phi_t(P_t - D_t) - C_t)(1 + r_t). \tag{33}$$

Because $P_t = D_t + \mathbb{E}_t[m_{t+1,t}P_{t+1}]$ for any t, from (29) it follows that

$$\begin{aligned}
\mathbb{E}_{t_0}[m_{t_0+1,t_0}X_{t_0+1}] &= \phi_{t_0}\mathbb{E}_{t_0}[m_{t_0+1,t_0}P_{t_0+1}] + (\psi_{t_0} - C_{t_0})(1 + r_{t_0+1,t_0})\mathbb{E}_{t_0}[m_{t_0+1,t_0}] \\
&= \phi_{t_0}(P_{t_0} - D_{t_0}) + (\psi_{t_0} - C_{t_0}) = X_{t_0} - C_{t_0}.
\end{aligned}$$

Repeating this argument, (31) follows.

By admissibility of X_{T+1} and the non-negativity of m it follows that $X_{t_0} \geq \sum_{t=t_0}^{T} \mathbb{E}_{t_0}\left[m_{t,t_0} C_t\right]$ for all $T \geq t_0$. Thus, (32) follows by letting $T \to \infty$:

$$X_{t_0} = P_{t_0} = \sum_{t=t_0}^{\infty} \mathbb{E}_{t_0}\left[m_{t,t_0} D_t\right] \geq \sum_{t=t_0}^{\infty} \mathbb{E}_{t_0}\left[m_{t,t_0} C_t\right]. \tag{34}$$

□

B.1 Additive Power Utility

Proof of Theorem 4.1. The closed form formula for the stock price P_n at time n is as follows

$$\begin{aligned}
D_n^{-\gamma} P_n &= \sum_{j=n}^{\infty} \mathbb{E}\left[\mathrm{e}^{-\beta(j-n)}\, D_j^{1-\gamma}\right] = D_n^{1-\gamma} \sum_{j=0}^{\infty} \mathrm{e}^{((1-\gamma)\eta-\beta)j}\, \mathbb{E}\left[\mathrm{e}^{(1-\gamma)sY_{j+n}^{(n)}}\right], \\
&= D_n^{1-\gamma} \sum_{j=0}^{\infty} \mathrm{e}^{((1-\gamma)\eta-\beta)j}\; {}_2F_1\left(-j, (n+1)p_n; n{+}1; 1-\mathrm{e}^{(1-\gamma)s}\right) \\
&= D_n^{1-\gamma} \sum_{j=0}^{\infty} \mathrm{e}^{((1-\gamma)\eta-\beta)j} \sum_{k=0}^{j} (-1)^k \binom{j}{k} \frac{((n+1)p_n)_k}{(n{+}1)_k} \left(1-\mathrm{e}^{(1-\gamma)s}\right)^k,
\end{aligned}$$

where

$$(q)_k = \begin{cases} 1 & : k = 0, \\ q(q+1)...(q+k-1) & : k > 0. \end{cases}$$

Now, changing the order of the summation

$$\begin{aligned}
P_n &= D_n \sum_{k=0}^{\infty} \frac{((n+1)p_n)_k}{(n+1)_k} (-1)^k \left(1-\mathrm{e}^{(1-\gamma)s}\right)^k \sum_{j=k}^{\infty} \mathrm{e}^{((1-\gamma)\eta-\beta)j} \binom{j}{k} \\
&= D_n \sum_{k=0}^{\infty} \frac{((n+1)p_n)_k}{(n+1)_k} (-1)^k \left(1-\mathrm{e}^{(1-\gamma)s}\right)^k \left(1-\mathrm{e}^{((1-\gamma)\eta-\beta)}\right)^{-k-1} \mathrm{e}^{((1-\gamma)\eta-\beta)k} \\
&= \frac{D_n}{1-\mathrm{e}^{((1-\gamma)\eta-\beta)}} \sum_{k=0}^{\infty} \frac{((n+1)p_n)_k}{(n+1)_k} (-1)^k \left(\frac{\mathrm{e}^{((1-\gamma)\eta-\beta)}\left(1-\mathrm{e}^{(1-\gamma)s}\right)}{1-\mathrm{e}^{((1-\gamma)\eta-\beta)}}\right)^k \\
&= \frac{D_n}{1-\mathrm{e}^{((1-\gamma)\eta-\beta)}} \sum_{k=0}^{\infty} \frac{k!((n+1)p_n)_k}{(n+1)_k} \frac{\left(-\frac{\left(1-\mathrm{e}^{(1-\gamma)s}\right)}{\mathrm{e}^{-((1-\gamma)\eta-\beta)}-1}\right)^k}{k!} \\
&= \frac{D_n}{1-\mathrm{e}^{((1-\gamma)\eta-\beta)}}\; {}_2F_1\left(1, (n+1)p_n; n+1; \frac{1-\mathrm{e}^{(1-\gamma)s}}{1-\mathrm{e}^{-((1-\gamma)\eta-\beta)}}\right),
\end{aligned}$$

where the second equality uses the identity $\sum_{j=k}^{\infty} q^j \binom{j}{k} = (1-q)^{-k-1} q^k$ with $q = \mathrm{e}^{((1-\gamma)\eta-\beta)}$, showing (7).

To show (8), recall the definition $B(t, t+1)$ – the price at time t of a zero coupon bound maturing at time $t+1$ in (29). As for power utility, (24) becomes $m_{t+1,t} = \mathrm{e}^{-\beta}\left(\frac{D_{t+1}}{D_t}\right)^{-\gamma}$, (8) readily follows by recalling the definition of the interest rate $r_{t+1,t}$ in (30).

Recall the with power utility the stochastic discount factor π in (25), (26) is $\pi_{t,t} = D_t^{-\gamma}$. Thus, by (24), (26) it follows that $\mathbb{E}_s[\pi_{t,s}C_t] = \mathbb{E}_{t_0}\left[\frac{1}{1-\gamma}\frac{\partial V_{t_0}(D)}{\partial D_s}C_t\right]$, $s \geq t \geq t_0$, for any $C_t \geq 0$ admissible consumption. Recall that from Lemma B.3 for any admissible consumption C with initial portfolio wealth P_{t_0}, $\sum_{t=t_0}^{\infty}\mathbb{E}_{t_0}[\pi_{t,t_0}D_t] \geq \sum_{t=t_0}^{\infty}\mathbb{E}_{t_0}[\pi_{t,t_0}C_t]$. For such consumption it follows that

$$
\begin{aligned}
\frac{1}{1-\gamma}V_{t_0}(C) &= \sum_{t=t_0}^{\infty}\mathbb{E}_{t_0}\left[\mathrm{e}^{-\beta(t-t_0)}\frac{C_t^{1-\gamma}}{1-\gamma}\right] \\
&\leq \sum_{t=0}^{\infty}\mathbb{E}_{t_0}\left[\mathrm{e}^{-\beta(t-t_0)}\frac{C_t^{1-\gamma}}{1-\gamma} + \pi_{t,t_0}(D_t - C_t)\right] \leq \sum_{t=0}^{\infty}\mathbb{E}_{t_0}\left[\mathrm{e}^{-\beta(t-t_0)}\frac{D_t^{1-\gamma}}{1-\gamma}\right] = \frac{V_{t_0}(D)}{1-\gamma},
\end{aligned}
$$

where the first inequality follows from

$$
\max_{C_t \geq 0}\mathbb{E}_{t_0}\left[\mathrm{e}^{-\beta(t-t_0)}\frac{C_t^{1-\gamma}}{1-\gamma} - \pi_{t,t_0}C_t\right] = \mathbb{E}_{t_0}\left[\mathrm{e}^{-\beta(t-t_0)}\frac{D_t^{1-\gamma}}{1-\gamma} - \pi_{t,t_0}D_t\right],
$$

which in turn follows from (27).

Assume for convenience that $t_0 = 0$. Note that, if $X_0 = P_0$, $\hat{C}_t = D_t$ and $\phi_t = 1$, then (33) implies by induction that $X_t = P_t$ for all $t \geq 0$. Now, consider the alternative strategy in which at time t the number of shares changes from 1 to $1+\varepsilon$ on some $\mathcal{F}_t$-measurable event $A \subset \{|P_t| < M, D_t > 1/M\}$, with $M > 0$. Note, that after the dividend is paid, the share price is $P_t - D_t$. Thus consumption correspondingly changes from D_t to $D_t - \varepsilon(P_t - D_t)$ and to $D_s(1+\varepsilon)$ for $s \geq t+1$. That is, define $\phi_s^{\varepsilon} = \phi_s + \varepsilon 1_{\{s\geq t\}\cap A}$ and $c_s^{\varepsilon} = D_s - \varepsilon P_t 1_{\{s=t\}\cap A} + \varepsilon D_s 1_{\{s\geq t+1\}\cap A}$, and note that this strategy continues to satisfy (33). (Note that ε may be either positive or negative.)

Setting $u(t, C_t) = \mathrm{e}^{-\beta t}\frac{C_t^{1-\gamma}}{1-\gamma}$, the change in expected utility from $(D, 1)$ to $(c^{\varepsilon}, \phi^{\varepsilon})$ is thus

$$
\Delta^{\varepsilon} = \mathbb{E}\left[1_A\left(u(t, D_t - \varepsilon(P_t - D_t)) - u(t, D_t) + \sum_{s=t+1}^{\infty}(u(s, D_s(1+\varepsilon)) - u(s, D_s))\right)\right] \leq 0 \quad (35)
$$

where the last inequality reflects the assumed optimality of the consumption stream D together with the treading strategy $\phi \equiv 1$. By concavity, note that for any $t, x, y > 0$:

$$
u_c(t, y)(y - x) \leq u(t, y) - u(t, x) \leq u_c(t, x)(y - x).
$$

Whence, on the event A, for $s > t$

$$
u_c(s, D_s(1+\varepsilon))\varepsilon D_s \leq u(s, D_s(1+\varepsilon)) - u(s, D_s) \leq u_c(s, D_s)\varepsilon D_s.
$$

Therefore, again on A,

$$
\begin{aligned}
&|u(s, D_s(1+\varepsilon)) - u(s, D_s)| \leq |\varepsilon|\, D_s \max(u_c(s, D_s), u_c(s, D_s(1+\varepsilon))) \\
&= |\varepsilon|\, D_s u_c(s, D_s) \leq |\varepsilon|\, D_s u_c(s, D_s),
\end{aligned} \quad (36)
$$

where for the first equality the fact that u is increasing and concave was used. Likewise,

$$
-\varepsilon P_t u_c(t, D_t - \varepsilon(P_t - D_t)) \leq u(t, D_t - \varepsilon(P_t - D_t)) - u(t, D_t) \leq -\varepsilon P_t u_c(t, D_t) \qquad \text{on } A.
$$

Hence on A, for $\epsilon > 0$ small enough

$$
|u(t, D_t - \varepsilon(P_t - D_t)) - u(t, D_t)| \leq |\varepsilon| P_t u_c(t, 1/M - \varepsilon(M - 1/M)) \leq |\varepsilon| P_t u_c(t, 1/(2M)). \quad (37)
$$

In view of (36) and (37), it follows that the respective incremental ratios are dominated by an integrable random variable, uniformly in ε. Thus, dividing Δ^ε in (35) by ε and passing to the limit as $\varepsilon \downarrow 0$, Lebesgue's dominated convergence theorem yields

$$\lim_{\varepsilon \downarrow 0} \frac{\Delta^\varepsilon}{\varepsilon} = \mathbb{E}\left[1_A\left(-u_c(t, D_t)(P_t - D_t) + \sum_{s=t+1}^{\infty} u_c(s, D_s)D_s\right)\right] \leq 0$$

Analogously, as $\varepsilon \uparrow 0$ it follows that $\lim_{\varepsilon \downarrow 0} \frac{\Delta^\varepsilon}{\varepsilon} \geq 0$, whence the limit must be zero. By the tower property of conditional expectation,

$$\mathbb{E}\left[1_A\left(-u_c(t, D_t)P_t + \mathbb{E}_t\left[\sum_{s=t}^{\infty} u_c(s, D_s)D_s\right]\right)\right] = 0.$$

As $M \uparrow \infty$, the event A spans any element of $\mathcal{F}_t$, which implies that

$$P_t = \mathbb{E}_t\left[\sum_{s=t}^{\infty} \frac{u_c(s, D_s)}{u_c(t, D_t)} D_s\right] \qquad \text{a.s..}$$

This completes the proof by recalling the definition of the SDF m in (24). □

We now adapt this proof to the Epstein-Zin recursive utility case.

B.2 Recursive Epstein-Zin Utility

The proof for the general recursive Epstein-Zin utility is more complicated, but the proof that the market is in equilibrium uses the same ideas as in the equivalent part of Theorem 4.1. The major difference is that there is no closed form solution to the price process, as opposed to the one found in Theorem 4.1. Hence, we proceed by finding a power expansion. First, it is more convenient to work with the following equilibrium price candidate P.

$$P_t = D_t + \mathbb{E}_t\left[m_{t+1,t}P_{t+1}\right]. \tag{38}$$

To establish the connection between utility U and price P, substitute (24) into (38) to get

$$\left(\mathbb{E}_t\left[U_{t+1}^{1-\gamma}(D)\right]\right)^{\frac{\theta-1}{\theta}} D_t^{\frac{1-\gamma}{\theta}-1} P_t = \left(\mathbb{E}_t\left[U_{t+1}^{1-\gamma}(D)\right]\right)^{\frac{\theta-1}{\theta}} D_t^{\frac{1-\gamma}{\theta}} + \delta \mathbb{E}_t\left[D_{t+1}^{\frac{1-\gamma}{\theta}-1} U_{t+1}^{\frac{(1-\gamma)(\theta-1)}{\theta}}(D)P_{t+1}\right].$$

Comparing this with (21) it follows that

$$U_t^{\frac{1-\gamma}{\theta}} = (1-\delta)D_t^{\frac{1-\gamma}{\theta}-1}P_t. \tag{39}$$

The proof that condition (22) holds is deferred to Lemma B.12. Next, let c_t defined by

$$P_t = c_t^{\frac{1-\gamma}{\theta}} D_t. \tag{40}$$

and attempt to find c_t. In other words $c_t^{\frac{1-\gamma}{\theta}}$ is the price dividend ratio. Then (39) becomes

$$U_t(D) = (1-\delta)^{\frac{\theta}{1-\gamma}} c_t D_t, \tag{41}$$

Substituting (41) into (21), it follows that

$$c_t^{\frac{1-\gamma}{\theta}} D_t^{\frac{1-\gamma}{\theta}} = D_t^{\frac{1-\gamma}{\theta}} + \delta \left(\mathbb{E}_t \left[c_{t+1}^{1-\gamma} D_{t+1}^{1-\gamma} \right] \right)^{\frac{1}{\theta}}.$$

and, using (4),

$$c_t^{\frac{1-\gamma}{\theta}} = 1 + \delta \, \mathrm{e}^{\eta \frac{1-\gamma}{\theta}} \left(\mathbb{E}_t \left[c_{t+1}^{1-\gamma} \, \mathrm{e}^{(1-\gamma)sX_{t+1}} \right] \right)^{\frac{1}{\theta}}. \tag{42}$$

Note that this is a backward recursion. If c_{t+1} is known and assuming $\hat{p}_t$ is also known, then c_t can be computed. Additionally, note that (42) can be solved if it is assumed that no more learning takes place, that is if $c_t = c_{t+1} = c$. In this case,

$$c^{\frac{1-\gamma}{\theta}} = 1 + \delta \, \mathrm{e}^{\eta \frac{1-\gamma}{\theta}} \, c^{\frac{1-\gamma}{\theta}} \left(\mathbb{E}_t \left[\mathrm{e}^{(1-\gamma)sX_{t+1}} \right] \right)^{\frac{1}{\theta}}.$$

It follows that

$$c^{\frac{1-\gamma}{\theta}} = \frac{1}{1 - \delta \, \mathrm{e}^{\eta \frac{1-\gamma}{\theta}} \left(\mathbb{E}_t \left[\mathrm{e}^{(1-\gamma)sX_{t+1}} \right] \right)^{\frac{1}{\theta}}} = \frac{1}{1 - \delta \, \mathrm{e}^{\eta \frac{1-\gamma}{\theta}} \left((1-\hat{p}_t) + \hat{p}_t \, \mathrm{e}^{(1-\gamma)s} \right)^{\frac{1}{\theta}}}.$$

Thus, define

$$c_\infty^{(n)}(p_n) = \left(\frac{1}{1 - \delta \, \mathrm{e}^{\eta \frac{1-\gamma}{\theta}} \left((1-p_n) + p_n \, \mathrm{e}^{(1-\gamma)s} \right)^{\frac{1}{\theta}}} \right)^{\frac{\theta}{1-\gamma}}.$$

Next, postulate that

$$(c_n(\hat{p}_n))^{\frac{1-\gamma}{\theta}} = \left(c_\infty^{(n)}(\hat{p}_n) \right)^{\frac{1-\gamma}{\theta}} + \sum_{i=1}^{\infty} \frac{\alpha_i(\hat{p}_n)}{n^i}, \tag{43}$$

and seek the coefficients α_i by subsisting into (42). Henceforth, the argument $\hat{p}_n$ of $c_n, c_\infty^{(n)}, \alpha_n$ is dropped for convenience. The coefficients in this expansion are solved explicitly by inserting (43) into (42). For example, the first one equals

$$\alpha_1(p) = \frac{(p(\mathrm{e}^{(1-\gamma)s} - 1) + 1)^{\frac{1}{\theta}} \, \mathrm{e}^{2\eta \frac{1-\gamma}{\theta}} \, \delta^2 (p-1) p (\mathrm{e}^{(1-\gamma)s} - 1)^2 (p(\mathrm{e}^{(1-\gamma)s} - 1) + 1)^{\frac{1}{\theta} - 2}}{\theta \left(\mathrm{e}^{\eta \frac{1-\gamma}{\theta}} \, \delta (p(\mathrm{e}^{(1-\gamma)s} - 1) + 1)^{\frac{1}{\theta}} - 1 \right)^3}.$$

and explicit formulas for higher-order coefficients follow similarly. The next auxiliary lemmas helps to verify the expansion (43).

Lemma B.4. *There exists $\nu_0 > 0$, such that*

$$\nu_0^{-(n-m)} D_m \le D_n \le \nu_0^{n-m} D_m, \text{ for any } n \ge m \ge 0. \tag{44}$$

Moreover, fix the starting point $n_0 \ge 0$, and assume that

$$0 < \delta < \delta_1 \triangleq 1 \wedge \max\{\mathrm{e}^{-(\eta+s)\frac{1-\gamma}{\theta}}, \mathrm{e}^{-\eta \frac{1-\gamma}{\theta}}\}.$$

Then for $n \ge n_0$,

$$c_{\min} \triangleq 1 \le c_n \le \left(\frac{1}{1 - \delta \, \mathrm{e}^{(1-\rho)(\eta + (s)^+)}} \right)^{\frac{1}{1-\rho}} \triangleq c_{\max}. \tag{45}$$

So that

$$(1-\delta)^{\frac{1}{1-\rho}} c_{\min} D_{n_0} \le U_{n_0}(D) \le (1-\delta)^{\frac{1}{1-\rho}} c_{\max} D_{n_0}. \tag{46}$$

Proof. Set $D^*_{n+n_0} \triangleq D_{n_0} \max\{e^{(n+n_0)\eta}, e^{(n+n_0)(\eta+s)}\} = D_{n_0} e^{(n+n_0)(\eta+(s)^+)}$. Then $0 < D_{n+n_0} \leq D^*_{n+n_0}$. Similarly, $D_{n_0} e^{(n+n_0)(\eta-(s)^-)} \leq D_{n+n_0}$. It then follows that (44) is satisfied with $\nu_0 = e^{|\eta|+|s|}$. To show (45), using the fact that U is increasing in consumption, it immediately follows from the definition of c_{n_0} in (41) that $c_{n_0} \geq 1$, whence $U^*_{n_0} = U_{n_0}(D^*) \geq U_{n_0}(D)$. Thus, (21) becomes

$$U^*_{n_0} = \left\{(1-\delta)(D^*_{n_0})^{\frac{1-\gamma}{\theta}} + \delta(U^*_{n_0+1})^{\frac{1-\gamma}{\theta}}\right\}^{\frac{\theta}{1-\gamma}},$$

where we used the identity $\mathbb{E}_t[(U^*_{n_0+1})^{1-\gamma}] = (U^*_{n_0+1})^{1-\gamma}$ because the consumption D^*_t is deterministic for $t \geq n_0$. Recalling that $\theta = \frac{1-\gamma}{1-\rho}$, it follows that for $V^*_{n_0} = \frac{(U^*_{n_0})^{1-\rho}}{(1-\rho)(1-\delta)}$

$$V^*_{n_0} = (D^*_{n_0})^{1-\rho} + \delta V^*_{n_0+1},$$

which is the power utility case, with risk aversion ρ. Hence,

$$V^*_{n_0} = \sum_{n=n_0}^{\infty} \delta^{n-n_0}(D^*_n)^{1-\rho} = \sum_{n=0}^{\infty}(D^*_{n_0})^{1-\rho}\delta^n e^{n(1-\rho)(\eta+(s)^+)} = \frac{(D^*_{n_0})^{1-\rho}}{1-\delta\, e^{(1-\rho)(\eta+(s)^+)}},$$

which implies (45) by recalling that $c_{n_0} = \frac{U_{n_0}}{(1-\delta)^{\frac{\theta}{1-\gamma}} D_{n_0}} \leq c^*_{n_0} = \frac{U^*_{n_0}}{(1-\delta)^{\frac{\theta}{1-\gamma}} D_{n_0}} = \frac{((1-\delta)V^*_{n_0})^{\frac{1}{1-\rho}}}{(1-\delta)^{\frac{1}{1-\rho}} D_{n_0}} = \left(\frac{1}{1-\delta\, e^{(1-\rho)(\eta+(s)^+)}}\right)^{\frac{1}{1-\rho}} = c_{\max}$. This also shows (46), as $(1-\delta)^{\frac{1}{1-\rho}} c_{\min} D_{n_0} \leq U_{n_0}(D)$.

□

Similarly, any admissible consumption stream admits the following bounds.

Lemma B.5. *Let $n_0 \geq 0$ be the initial time. Then there exists a constant $K_0 > 0$, independent of n_0, such that $U_n(C) \leq K_0 X_n$, for any $n \geq n_0$ and for any admissible consumption process C.*

Proof. Using ν_0 from Lemma B.4 and recalling (40), it follows that $K_1^{-1}\nu_0^{-(n-n_0)} D_{n_0} \leq c_{max}^{1-\rho} \wedge c_{min}^{1-\rho} D_n \leq P_n \leq c_{max}^{1-\rho} \vee c_{min}^{1-\rho} D_n \leq K_1 \nu_0^{n-n_0} D_{n_0}$, for some constant $K_1 > 0$ and $n \geq n_0$. Hence, it also follows that $\frac{P_n}{P_{n-1}} \leq K_1^2 \nu_0$. Using the bounds on U from Lemma B.4, for another constant $K_2 > 0$ it follows that $m_{n+1,n} \geq \frac{1}{K_2 \nu_0^{-\rho}}$, which implies that the same bound holds for $1 + r_{n+1,n} \leq K_2 \nu_0^{-\rho}$. Thus for $\nu_1 = K_2 \nu_0^{-\rho} \vee K_1^2 \nu_0$, it follows that $C_n \leq \nu_1^{n-n_0} X_{n_0}$, $n \geq n_0$. A similar calculation as in Lemma B.4 yields the upper bound $U_{n_0}(C) \leq K_0 X_{n_0}$, for some constant $K > 0$.

□

Lemma B.6. *Set*

$$Err \triangleq \max_{p\in[0,1]} \frac{\delta^2(p-1)pe^{2\eta(1-\gamma)}\left(e^{(1-\gamma)s}-1\right)^2\left(p\left(e^{(1-\gamma)s}-1\right)+1\right)^{\frac{2}{\theta}-2}}{\theta\left(\delta e^{\eta(1-\gamma)}\left(p\left(e^{(1-\gamma)s}-1\right)+1\right)^{1/\theta}-1\right)^2} + 1,$$

$$B_1 \triangleq 1 + c_{\max}^{\frac{1-\gamma}{\theta}} e^{\left(\frac{1-\gamma}{\theta}(\eta+s)\right)^+} + Err, \tag{47}$$

$$B_2 \triangleq 1 + c_{\max}^{1-\gamma} e^{\left(\frac{1-\gamma}{\theta}s\right)^+}, \tag{48}$$

and assume that $0 < \bar{\delta} < 1$, where

$$\bar{\delta} = \delta \max\left\{|\theta|\,(1 \vee |B_1|^{\theta-1})\, e^{\left(\eta\frac{1-\gamma}{\theta}\right)^+}, (1 \vee |B_1|^{\theta-1})(1 \vee |B_2|^{-1-\frac{1}{\theta}})\, e^{\eta\frac{1-\gamma}{\theta}(1-\gamma)(s)^+}\right\}. \tag{49}$$

Moreover, let the assumptions of Lemma B.4 hold. Then $\left|\left(c_\infty^{(n)}(p_n)\right)^{\frac{1-\gamma}{\theta}} - (c_n(p_n))^{\frac{1-\gamma}{\theta}}\right| = O\left(\frac{1}{n}\right).$

Proof. First, note that $c_\infty^{(n)}$ almost satisfy (42), more specifically, for $n > 0$ big enough

$$
\begin{aligned}
&(c_\infty^{(n)}(\hat{p}_n))^{\frac{1-\gamma}{\theta}} - 1 - \delta\, \mathrm{e}^{\eta\frac{1-\gamma}{\theta}} \left(\mathbb{E}_t\left[(c_\infty^{(n+1)}(\hat{p}_{n+1}))^{1-\gamma}\, \mathrm{e}^{(1-\gamma)sX_{n+1}}\right]\right)^{\frac{1}{\theta}} \\
&= \frac{1}{n}\frac{\delta^2(\hat{p}_n-1)p_n e^{2\eta(1-\gamma)}\left(e^{(1-\gamma)s}-1\right)^2\left(\hat{p}_n\left(e^{(1-\gamma)s}-1\right)+1\right)^{\frac{2}{\theta}-2}}{\theta\left(\delta e^{\eta(1-\gamma)}\left(\hat{p}_n\left(e^{(1-\gamma)s}-1\right)+1\right)^{1/\theta}-1\right)^2} + O\left(\frac{1}{n^2}\right) \le \frac{\mathrm{Err}}{n}.
\end{aligned}
$$

Fix n and $N > n$. The idea is to express the difference between $(c_n(\hat{p}_n))^{1-\gamma}$ and $(c_\infty^{(n)}(\hat{p}_n))^{1-\gamma}$ using the difference at time $n+1$, and then recursively repeat the process until time N. Observe that

$$
\begin{aligned}
&\left|(c_n(\hat{p}_n))^{1-\gamma} - (c_\infty^{(n)}(\hat{p}_n))^{1-\gamma}\right| \\
&\le \Bigg|\left(1+\delta\,\mathrm{e}^{\eta\frac{1-\gamma}{\theta}}\left(\mathbb{E}_n\left[(c_{n+1}(\hat{p}_{n+1}))^{1-\gamma}\,\mathrm{e}^{(1-\gamma)sX_{n+1}}\right]\right)^{\frac{1}{\theta}}\right)^\theta \\
&- \left(1+\delta\,\mathrm{e}^{\eta\frac{1-\gamma}{\theta}}\left(\mathbb{E}_n\left[(c_\infty^{(n+1)}(\hat{p}_{n+1}))^{1-\gamma}\,\mathrm{e}^{(1-\gamma)sX_{n+1}}\right]\right)^{\frac{1}{\theta}} + \frac{\mathrm{Err}}{n}\right)^\theta\Bigg| \\
&\le |\theta|\,|\zeta_n|^{\theta-1}\,\delta\,\mathrm{e}^{\eta\frac{1-\gamma}{\theta}} \\
&\quad\times\left|\left(\mathbb{E}_n\left[(c_{n+1}(\hat{p}_{n+1}))^{1-\gamma}\,\mathrm{e}^{(1-\gamma)sX_{n+1}}\right]\right)^{\frac{1}{\theta}} - \left(\mathbb{E}_n\left[(c_\infty^{(n+1)}(\hat{p}_{n+1}))^{1-\gamma}\,\mathrm{e}^{(1-\gamma)sX_{n+1}}\right]\right)^{\frac{1}{\theta}} + \frac{\mathrm{Err}}{n}\right| \\
&\le |\theta|\,|\zeta_n|^{\theta-1}\,\delta\,\mathrm{e}^{\eta\frac{1-\gamma}{\theta}}\left|\frac{\left|\hat{\zeta}_n\right|^{-1-\frac{1}{\theta}}}{|\theta|}\mathbb{E}_n\left[\left((c_{n+1}(\hat{p}_{n+1}))^{1-\gamma} - (c_\infty^{(n+1)}(\hat{p}_{n+1}))^{1-\gamma}\right)\mathrm{e}^{(1-\gamma)sX_{n+1}}\right] + \frac{\mathrm{Err}}{n}\right| \\
&\le |\theta|\,|\zeta_n|^{\theta-1}\,\delta\,\mathrm{e}^{\eta\frac{1-\gamma}{\theta}}\frac{\mathrm{Err}}{n} \\
&\quad + |\zeta_n|^{\theta-1}\left|\hat{\zeta}_n\right|^{-1-\frac{1}{\theta}}\delta\,\mathrm{e}^{\eta\frac{1-\gamma}{\theta}(1-\gamma)(s)^+}\,\mathbb{E}_n\left[\left|(c_{n+1}(\hat{p}_{n+1}))^{1-\gamma} - (c_\infty^{(n+1)}(\hat{p}_{n+1}))^{1-\gamma}\right|\right],
\end{aligned}
$$

here ζ_n, and $\hat{\zeta}_n$ are unknown points in the Taylor remainder. Note that both ζ_n and $\hat{\zeta}_n$ are uniformly bounded, independently of n. Indeed, the point ζ_n is located somewhere between $1+\delta\,\mathrm{e}^{\eta\frac{1-\gamma}{\theta}}\left(\mathbb{E}_n\left[(c_{n+1}(\hat{p}_{n+1}))^{1-\gamma}\,\mathrm{e}^{(1-\gamma)sX_{n+1}}\right]\right)^{\frac{1}{\theta}}$ and $\delta\,\mathrm{e}^{\eta\frac{1-\gamma}{\theta}}\left(\mathbb{E}_n\left[(c_\infty^{(n+1)}(\hat{p}_{n+1}))^{1-\gamma}\,\mathrm{e}^{(1-\gamma)sX_{n+1}}\right]\right)^{\frac{1}{\theta}}+1+\frac{\mathrm{Err}}{n}$. Both of these quantities are bounded between 1 and B_1 from (47). Similarly, the point $\hat{\zeta}_n$ is located between $\mathbb{E}_n\left[(c_{n+1}(\hat{p}_{n+1}))^{1-\gamma}\,\mathrm{e}^{(1-\gamma)sX_{n+1}}\right]$ and $\mathbb{E}_n\left[(c_\infty^{(n+1)}(\hat{p}_{n+1}))^{1-\gamma}\,\mathrm{e}^{(1-\gamma)sX_{n+1}}\right]$, which are bounded by $\mathrm{e}^{(1-\gamma)s^-}$ and B_2 from (48). Recalling the definition of $\bar{\delta}$ in (49), the previous chain of inequalities continues as

$$
\begin{aligned}
\left|(c_n(\hat{p}_n))^{1-\gamma} - (c_\infty^{(n)}(\hat{p}_n))^{1-\gamma}\right| &\le \bar{\delta}\frac{\mathrm{Err}}{n} + \bar{\delta}\mathbb{E}_n\left[\left|(c_{n+1}(\hat{p}_{n+1}))^{1-\gamma} - (c_\infty^{(n+1)}(\hat{p}_{n+1}))^{1-\gamma}\right|\right] \\
&\le \bar{\delta}\frac{\mathrm{Err}}{n} + \bar{\delta}^2\frac{\mathrm{Err}}{n+1} + \bar{\delta}^2\mathbb{E}_n\left[\left|(c_{n+2}(\hat{p}_{n+2}))^{1-\gamma} - (c_\infty^{(n+2)}(\hat{p}_{n+2}))^{1-\gamma}\right|\right] \\
&\le (\bar{\delta}+\bar{\delta}^2)\frac{\mathrm{Err}}{n} + \bar{\delta}^2\mathbb{E}_n\left[\left|(c_{n+2}(\hat{p}_{n+2}))^{1-\gamma} - (c_\infty^{(n+2)}(\hat{p}_{n+2}))^{1-\gamma}\right|\right].
\end{aligned}
$$

Which implies that

$$\left|(c_n(\hat{p}_n))^{1-\gamma} - (c_\infty^{(n)}(\hat{p}_n))^{1-\gamma}\right| \le \left(\sum_{k=1}^{\infty} \bar{\delta}^k\right) \frac{\mathrm{Err}}{n} + \bar{\delta}^{N-n} \mathbb{E}_n \left[\left|(c_N(\hat{p}_N))^{1-\gamma} - (c_\infty^{(N)}(\hat{p}_N))^{1-\gamma}\right|\right]$$
$$= \frac{1}{1-\bar{\delta}} \frac{\mathrm{Err}}{n} + \bar{\delta}^{N-n} \mathbb{E}_n \left[\left|(c_N(\hat{p}_N))^{1-\gamma} - (c_\infty^{(N)}(\hat{p}_N))^{1-\gamma}\right|\right].$$

Letting $N \to \infty$ the claim of the lemma now follows as both c_N and $c_\infty^{(N)}$ are bounded. □

This lemma can be generalized to higher orders. (The corresponding proof is omitted.)

Lemma B.7. *For any $k \ge 1$, there exists $\delta > 0$ small enough, such that*

$$\left|\left(c_\infty^{(n)}(\hat{p}_n)\right)^{\frac{1-\gamma}{\theta}} + \sum_{i=1}^{k} \frac{\alpha_i(\hat{p}_n)}{n} - (c_n(\hat{p}_n))^{\frac{1-\gamma}{\theta}}\right| = O\left(\frac{1}{n^{k+1}}\right).$$

Lemma B.8. *The interest rate $r_{t,t+1}$ with Epstein-Zin recursive utility is as in* (10).

Proof. Using (41), and (4) the SDF from (24) becomes

$$m_{t+1,t} = \delta \left(\frac{D_{t+1}}{D_t}\right)^{\frac{1-\gamma}{\theta}-1} \left(\frac{c_{t+1} D_{t+1}}{\left(\mathbb{E}_t\left[(c_{t+1}D_{t+1})^{1-\gamma}\right]\right)^{\frac{1}{1-\gamma}}}\right)^{\frac{(1-\gamma)(\theta-1)}{\theta}}$$
$$= \delta \left(\frac{D_t\, \mathrm{e}^{\eta + sX_{t+1}}}{D_t}\right)^{\frac{1-\gamma}{\theta}-1} \left(\frac{c_{t+1} D_t\, \mathrm{e}^{\eta+sX_{t+1}}}{\left(\mathbb{E}_t\left[(c_{t+1}D_t\, \mathrm{e}^{\eta+sX_{t+1}})^{1-\gamma}\right]\right)^{\frac{1}{1-\gamma}}}\right)^{\frac{(1-\gamma)(\theta-1)}{\theta}}$$
$$= \delta\, \mathrm{e}^{\eta\left(\frac{1-\gamma}{\theta}-1\right)} c_{t+1}^{\frac{(1-\gamma)(\theta-1)}{\theta}} \left(\mathbb{E}_t\left[c_{t+1}^{1-\gamma}\, \mathrm{e}^{(1-\gamma)sX_{t+1}}\right]\right)^{\frac{1-\theta}{\theta}} \mathrm{e}^{-\gamma s X_{t+1}}. \tag{50}$$

Recall the definition of bond price $B(t, t+1)$ in (29). It follows from (50) that

$$B(t,t+1) = \delta\, \mathrm{e}^{\eta\left(\frac{1-\gamma}{\theta}-1\right)} \mathbb{E}_t\left[c_{t+1}^{\frac{(1-\gamma)(\theta-1)}{\theta}} \mathrm{e}^{-\gamma s X_{t+1}}\right] \left(\mathbb{E}_t\left[c_{t+1}^{1-\gamma}\, \mathrm{e}^{(1-\gamma)sX_{t+1}}\right]\right)^{\frac{1-\theta}{\theta}}.$$

The desired result (10) follows readily now from the definition of the interest rate $r_{t+1,t}$ in (30). □

Corollary B.9. *For $\delta > 0$ small enough,* (50) *implies that*

$$\left|\frac{\left(\mathbb{E}_n\left[\left(c_\infty^{(n+1)}(\hat{p}_{n+1})\right)^{1-\gamma} \mathrm{e}^{(1-\gamma)sX_{n+1}}\right]\right)^{\frac{\rho-\gamma}{1-\gamma}}}{\delta\, \mathrm{e}^{-\eta\rho}\, \mathbb{E}_n\left[\left(c_\infty^{(n+1)}(\hat{p}_{n+1})\right)^{\rho-\gamma} \mathrm{e}^{-\gamma s X_{n+1}}\right]} - 1 - r_{n+1,n}\right| = O\left(\frac{1}{n}\right)$$

And an error of $O\left(\frac{1}{n^{k+1}}\right), k \ge 1$ can be achieved if higher order approximation of $\left(c_\infty^{(n)}(\hat{p}_n)\right)^{\frac{1-\gamma}{\theta}} + \sum_{i=1}^{k} \frac{\alpha_i(\hat{p}_n)}{n}$ is used to approximate $(c_n(\hat{p}_n))^{\frac{1-\gamma}{\theta}}$.

Proof. The proof follows from the combination of Lemmas B.4, B.6, B.7, B.8. □

Corollary B.10. *For $\delta > 0$ small enough, note that* $\lim_{t\to\infty} \mathbb{E}_{t_0}[m_{t,t_0} P_t] = 0$.

Proof. Recall that $\theta = \frac{1-\gamma}{1-\rho}$. Then, under the assumption that $\delta > 0$ small enough,

$$0 < m_{t+1,t} \frac{D_{t+1}}{D_t} \le \delta_2. \tag{51}$$

for some $\delta_2 < 1$. This can be seen by considering different cases. For example, when $\gamma > 1$ and $\rho < \gamma$, so that $1-\gamma, \frac{1-\theta}{\theta} = \frac{\gamma-\rho}{1-\gamma} < 0$, using the definition of m in (24) and Lemma B.4 it follows that

$$\begin{aligned} 0 < m_{t+1,t} \frac{D_{t+1}}{D_t} &\le \delta \, \mathrm{e}^{-\eta\rho} \, \mathbb{E}_t \left[\left(\frac{c_{\max}}{c_{\min}} \right)^{1-\gamma} \mathrm{e}^{(1-\gamma)sX_{t+1}} \right]^{\frac{1-\theta}{\theta}} \mathrm{e}^{-\gamma s X_{t+1}} \\ &\le \left(\frac{c_{\max}}{c_{\min}} \right)^{\gamma-\rho} \delta \, \mathrm{e}^{-\eta\rho} \, \mathrm{e}^{-\gamma s^-} \, \mathrm{e}^{(\gamma-\rho)s^+} = \delta \left(\frac{c_{\max}}{c_{\min}} \right)^{\gamma-\rho} \mathrm{e}^{-\eta\rho+\gamma s-\rho s^+} . \end{aligned}$$

Thus (51) holds for $\delta > 0$ small enough. Thus, for any $t_0 \ge 0$,

$$\lim_{t\to\infty} m_{t,t_0} D_t = D_{t_0} \lim_{t\to\infty} \prod_{n=t_0+1}^{t} m_{n,n-1} \frac{D_n}{D_{n-1}} \le D_{t_0} \lim_{t\to\infty} \delta_2^{t-t_0} = 0,$$

whence $\lim_{t\to\infty} \mathbb{E}_{t_0}[m_{t,t_0} P_t] = 0$. □

The next corollary is presented for completeness only. It shows that the two price candidates (38) and (28) in the power utility and Epstein-Zin utility coincide.

Corollary B.11. *The price P in* (38) *equals* (28).

Proof. Recall that $m_{t_0,t_0} = 1$. The equality between (38) and (28) follows from Corollary B.10. □

So far we have been using the recursion (21). We are now ready to show that the asymptotic condition (22) holds.

Lemma B.12. *Let $U_t(D)$ be as in* (41). *Then for $\delta > 0$ small enough,*

$$\lim_{N\to\infty} \left| U_t(D) - U_t\left(D^{0,N}\right) \right| = 0.$$

Proof. Observe, that the equivalent of (46) also holds for $U_t\left(D^{0,N}\right)$, for $N \ge t+1$. Namely,

$$(1-\delta)^{\frac{1}{1-\rho}} c_{\min} D_t \le U_t\left(D^{0,N}\right) \le (1-\delta)^{\frac{1}{1-\rho}} c_{\max} D_t.$$

Thus, similarly to (51) and using the same δ_2 we can bound $0 < m^N_{t+1,t} \frac{D_{t+1}}{D_t} \le \delta_2$, where

$$m^N_{t+1,t} = \delta \left(\frac{D_{t+1}}{D_t} \right)^{\frac{1-\gamma}{\theta}-1} \left(\frac{U_{t+1}\left(D^{0,N}\right)}{\left(\mathbb{E}_t \left[\left(U_{t+1}\left(D^{0,N}\right) \right)^{1-\gamma} \right] \right)^{\frac{1}{1-\gamma}}} \right)^{-\frac{(1-\gamma)(1-\theta)}{\theta}} .$$

Then from Lemma B.4 it follows that

$$\left|U_t(D) - U_t\left(D^{0,N}\right)\right| \le \mathbb{E}_t\left[\pi_{t,t}\prod_{n=t}^{N} m^N_{n+1,n}\left|U_N(D) - (1-\delta)^{\frac{1}{1-\rho}}D_N\right|\right]$$
$$\le \mathbb{E}_t\left[\pi_{t,t}\prod_{n=t}^{N} m^N_{n+1,n}U_N(D)\right] \le \pi_{t,t}(1-\delta)^{\frac{1}{1-\rho}}c_{\max}\delta_2^{N-t+1}D_t,$$

which in turn converges to zero as $N \to \infty$. □

We are now ready for the equilibrium proof for Epstein-Zin utility.

Proof of Theorem 4.2. First, note that we have already proved parts of Theorem 4.2. Specifically, Lemmas B.6 and B.7 show the validity of (9) and (11), and (10) and of (12) follow from Lemma B.8 and Corollary B.9 respectively . The next two steps similar to the ones in the proof of Theorem 4.1 is to show that the consumption D maximizes the utility U subject to the budget constraint and then use this result to show the market is in equilibrium.

Let $\epsilon > 0$ and let $t \ge 0$ be the initial time. Assume the initial wealth is $X_t = P_t$, so that the consumption stream D is admissible (otherwise, it suffices to scale it). Fix a consumption process C, also admissible for this initial wealth. The first goal is to show that $U_t(C) \ge U_t(D)$. Without loss of generality assume that $\sum_{s=t}^{\infty}\mathbb{E}_t\left[m_{s,t}C_s\right] = X_t$. Indeed, $\sum_{s=t}^{\infty}\mathbb{E}_t\left[m_{s,t}C_s\right] \le X_t$ by Lemma B.3. Thus if the inequality is strict we may increase the consumption, and thereby increase the utility.

The goal now is to show that $U_t(C) \le U_t(D)$. From (34), there exists $n \ge t$ such that $\sum_{s=n+1}^{\infty}\mathbb{E}_t\left[m_{s,t}C_s\right] \le \epsilon$, $\sum_{s=n+1}^{\infty}\mathbb{E}_t\left[m_{s,t}D_s\right] \le \epsilon$, and hence

$$\sum_{s=n+1}^{\infty}\mathbb{E}_t\left[\pi_{s,t}(C_s - D_s)\right] \le 2\pi_{t,t}\epsilon.$$

Recall the definition (23), which defines the modified consumption process $D^{C,n} = \begin{cases} D_s & : s \le n, \\ C_s & : s > n. \end{cases}$
It then follows from Lemma B.3 that

$$\mathbb{E}_t\left[\pi_{n+1,t}X_{n+1}(D^{C,n})\right] = \mathbb{E}_t\left[\pi_{n+1,t}X_{n+1}(C)\right] = X_t - \sum_{s=t}^{n}\mathbb{E}_t\left[\pi_{s,t}C_s\right] = \sum_{s=n+1}^{\infty}\mathbb{E}_t\left[\pi_{s,t}C_s\right] \le \pi_{t,t}\epsilon.$$

We next show that $U_t(D^{C,n}) \le U_t(D) + K_0\epsilon$, where $K_0 > 0$ is the constant from Lemma B.5. Clearly, we only need to consider the case, when $U_t(D^{C,n}) \ge U_t(D)$. Then from the concavity of U and

$$U_t(D^{C,n}) - U_t(D) \le \mathbb{E}_t\left[\frac{\partial U_t(D)}{\partial U_{n+1}}\left|U_{n+1}(D^{C,n}) - U_{n+1}(D)\right|\right] \le \mathbb{E}_t\left[\frac{\partial U_t(D)}{\partial U_{n+1}}U_{n+1}(D^{C,n})\right]$$
$$= \mathbb{E}_t\left[m_{n+1,t}U_{n+1}(D^{C,n})\right] \le \mathbb{E}_t\left[m_{n+1,t}K_0X_{n+1}\right] \le K_0\epsilon,$$

where the third inequality is from Lemma B.5. Then

$$U_t(C) \le U_t(C) - \sum_{s=t}^{n}\mathbb{E}_t\left[\pi_{s,t}\left(C_s - D_s\right)\right] - \sum_{s=n+1}^{\infty}\mathbb{E}_t\left[\pi_{s,t}\left(C_s - D_s\right)\right]$$
$$\le U_t(C) - \sum_{s=t}^{n}\mathbb{E}_t\left[\pi_{s,t}\left(C_s - D_s\right)\right] + 2\epsilon\pi_{t,t}$$
$$\le U_t(D^{C,n}) + 2\epsilon\pi_{t,t} \le U_t(D) + \left(2\pi_{t,t} + K_0\right)\epsilon.$$

where the first inequality holds by (34), and the third from (27). Letting $\epsilon \to 0$ it follows that $U_t(C) \le U_t(D)$ and thus D maximizes the utility of consumption from a given initial wealth.

We now proceed in a similar fashion to the proof of Theorem 4.1. Consider the alternative strategy in which at time t the number of shares changes from 1 to $1+\varepsilon$ on some $\mathcal{F}_t$-measurable event $A \subset \{|P_t|, < M, D_t > 1/M\}$, with $M > 0$, while at the next time step $t+1$ the extra shares now worth εP_{t+1} are consumed in addition to D_{t+1}, and for times $s \ge t+2$ the consumption remains the same as before D_s. That is, define $\phi_s^\varepsilon = \phi_s + \varepsilon 1_{\{s=t\}} \cap A$ and $C_s^\varepsilon = D_s - \varepsilon P_s 1_{\{s=t\}\cap A} + \varepsilon P_s 1_{\{s=t+1\}\cap A}$, and note that this strategy continues to satisfy (33). (Note that ε may be either positive or negative.) The change in expected utility from $(D, 1)$ to $(c^\varepsilon, \phi^\varepsilon)$ is thus

$$\begin{aligned}\Delta^\varepsilon = \mathbb{E}\Bigg[1_A \Bigg\{&(1-\delta)(D_t - \varepsilon(P_t - D_t))^{\frac{1-\gamma}{\theta}} \\ &+\delta\left(\mathbb{E}\left[\left\{(1-\delta)(D_{t+1}+\varepsilon P_{t+1})^{\frac{1-\gamma}{\theta}} + \delta\left(\mathbb{E}_{t+1}[U_{t+2}(D)^{1-\gamma}]\right)^{\frac{1}{\theta}}\right\}^\theta\right]\right)^{\frac{1}{\theta}}\Bigg\}^{\frac{\theta}{1-\gamma}}\Bigg] \\ -\mathbb{E}\Bigg[1_A\Bigg\{&(1-\delta)D_t^{\frac{1-\gamma}{\theta}} + \delta\left(\mathbb{E}_t\left[\left\{(1-\delta)D_{t+1}^{\frac{1-\gamma}{\theta}} + \delta\left(\mathbb{E}_{t+1}[U_{t+2}(D)^{1-\gamma}]\right)^{\frac{1}{\theta}}\right\}^\theta\right]\right)^{\frac{1}{\theta}}\Bigg\}^{\frac{\theta}{1-\gamma}}\Bigg] \le 0\end{aligned} \tag{52}$$

where the last inequality reflects the assumed optimality of $(D, 1)$. For any increasing, concave function $u(x, y)$, it holds that

$$u_x(x_2,y_2)(x_2-x_1)+u_y(x_2,y_2)(y_2-y_1) \le u(x_2,y_2)-u(x_1,y_1) \le u_x(x_1,y_1)(x_2-x_1)+u_y(x_1,y_1)(y_2-y_1)$$

whence, on the event A,

$$\begin{aligned}&\frac{\partial U_t}{\partial C_t}(D)\varepsilon(P_t - D_t) + \varepsilon\mathbb{E}_t\left[\frac{\partial U_t}{\partial C_{t+1}}(D)P_{t+1}\right] \le U_t(D) - U_t(C^\varepsilon) \\ &\le \frac{\partial U_t}{\partial C_t}(C^\varepsilon)\varepsilon(P_t - D_t) + \varepsilon\mathbb{E}_t\left[\frac{\partial U_t}{\partial C_{t+1}}(C^\varepsilon)P_{t+1}\right].\end{aligned} \tag{53}$$

Note that from Lemma B.5 on A, we also have that $P_{t+1} \le M_1 \triangleq \frac{M}{\nu_0}$, and $D_{t+1} \ge \frac{1}{M_1}$. Set

$$C_s^M = \begin{cases} 1/(2M) & : s = t, \\ 1/(2M_1) & : s = t+1, \\ 0 & : s \ge t+2 \end{cases}, \text{ for } s \ge t.$$

Assuming $0 < \varepsilon < 1/(2M^2)$, we have that $D_s, C_s^\varepsilon \ge C_s^M$ for all $s \ge t$. Thus, from (53)

$$\begin{aligned}|U_t(D) - U_t(C^\varepsilon)| &\le \left|\frac{\partial U_t}{\partial C_t}(C^M)\varepsilon(P_t - D_t)\right| + \left|\varepsilon\frac{\partial U_t}{\partial C_{t+1}}(C^M)\mathbb{E}_t[P_{t+1}]\right| \\ &\le \left|\frac{\partial U_t}{\partial C_t}(C^M)\right|\varepsilon M + M_1\varepsilon\left|\frac{\partial U_t}{\partial C_{t+1}}(C^M)\right|.\end{aligned} \tag{54}$$

In view of (54), it follows that the respective incremental ratios are dominated by an integrable random variable, uniformly in ε. Thus, dividing Δ^ε in (52) by ε and passing to the limit as $\varepsilon \downarrow 0$, Lebesgue's dominated convergence theorem yields

$$\lim_{\varepsilon\downarrow 0}\frac{\Delta^\varepsilon}{\varepsilon} = \mathbb{E}\left[1_A\left(-\frac{\partial U_t(D)}{\partial C_t}(P_t - D_t) + \frac{\partial U_t(D)}{\partial C_{t+1}}P_{t+1}\right)\right] \le 0$$

Analogously, as $\varepsilon \uparrow 0$ it follows that $\lim_{\varepsilon \downarrow 0} \frac{\Delta^\varepsilon}{\varepsilon} \geq 0$, whence the limit must be zero. By the tower property of conditional expectation,

$$\mathbb{E}\left[1_A\left(-\frac{\partial U_t(D)}{\partial C_t}(P_t - D_t) + \frac{\partial U_t(D)}{\partial C_{t+1}} P_{t+1}\right)\right] = 0.$$

As $M \uparrow \infty$, the event A spans any element of $\mathcal{F}_t$, and recalling the definition $m_{t+1,t}$ in (24), we get that that

$$P_t = D_t + E_t\left[m_{t+1,t} P_{t+1}\right] \qquad \text{a.s..}$$

□

References

Andrew Ang, Geert Bekaert, and Min Wei. Do macro variables, asset markets, or surveys forecast inflation better? *Journal of monetary Economics*, 54(4):1163–1212, 2007.

Malcolm Baker and Jeffrey Wurgler. The equity share in new issues and aggregate stock returns. *the Journal of Finance*, 55(5):2219–2257, 2000.

Jason Beeler, John Y Campbell, et al. The long-run risks model and aggregate asset prices: An empirical assessment. *Critical Finance Review*, 1(1):141–182, 2012.

Van Binsbergen, H Jules, and Ralph SJ Koijen. Predictive regressions: A present-value approach. *The Journal of Finance*, 65(4):1439–1471, 2010.

William Breen, Lawrence R Glosten, and Ravi Jagannathan. Economic significance of predictable variations in stock index returns. *The Journal of Finance*, 44(5):1177–1189, 1989.

Frode Brevik and Stefano d'Addona. Information quality and stock returns revisited. *Journal of Financial and Quantitative Analysis*, 45(6):1419–1446, 2010.

John Y Campbell and Robert J Shiller. The dividend-price ratio and expectations of future dividends and discount factors. *The Review of Financial Studies*, 1(3):195–228, 1988.

John Y Campbell and Samuel B Thompson. Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies*, 21(4):1509–1531, 2007.

John Y Campbell and Tuomo Vuolteenaho. Inflation illusion and stock prices. Technical report, National bureau of economic research, 2004.

Long Chen, Zhi Da, and Xinlei Zhao. What drives stock price movements? *The Review of Financial Studies*, 26(4):841–876, 2013.

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

John H Cochrane. The dog that did not bark: A defense of return predictability. *The Review of Financial Studies*, 21(4):1533–1575, 2007.

Timothy Cogley and Thomas J Sargent. Diverse beliefs, survival and the market price of risk. *The Economic Journal*, 119(536):354–376, 2009.

Pierre Collin-Dufresne, Michael Johannes, and Lars A Lochstoer. Parameter learning in general equilibrium: The asset pricing implications. *The American Economic Review*, 106(3):664–698, 2016.

John C Cox, Stephen A Ross, and Mark Rubinstein. Option pricing: A simplified approach. *Journal of financial Economics*, 7(3):229–263, 1979.

Mariano M Croce, Martin Lettau, and Sydney C Ludvigson. Investor information, long-run risk, and the term structure of equity. *The Review of Financial Studies*, 28(3):706–742, 2014.

Zhi Da, Ravi Jagannathan, and Jianfeng Shen. Growth expectations, dividend yields, and future stock returns. Technical report, National Bureau of Economic Research, 2014.

Darrell Duffie and Larry G Epstein. Asset pricing with stochastic differential utility. *The Review of Financial Studies*, 5(3):411–436, 1992.

Larry G Epstein and Stanley E Zin. Substitution, risk aversion and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica*, 57(4):937–969, 1989.

Eugene F Fama and Kenneth R French. Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, 33(1):3–56, 1993.

Hans-Otto Georgii. *Stochastics: introduction to probability and statistics.* Walter de Gruyter, 2013.

Lawrence R Glosten, Ravi Jagannathan, and David E Runkle. On the relation between the expected value and the volatility of the nominal excess return on stocks. *The journal of finance*, 48(5):1779–1801, 1993.

Amit Goyal and Ivo Welch. Predicting the equity premium with dividend ratios. *Management Science*, 49 (5):639–654, 2003.

Lars Peter Hansen. Beliefs, doubts and learning: Valuing economic risk. Technical report, National Bureau of Economic Research, 2007.

Ravi Jagannathan and Binying Liu. Dividend dynamics, learning, and expected stock index returns. Technical report, National Bureau of Economic Research, 2015.

Michael Johannes, Lars A Lochstoer, and Yiqun Mou. Learning about consumption dynamics. *The Journal of Finance*, 71(2):551–600, 2016.

Bryan Kelly and Seth Pruitt. Market expectations in the cross-section of present values. *The Journal of Finance*, 68(5):1721–1756, 2013.

David M Kreps. Anticipated utility and dynamic choice. *Econometric Society Monographs*, 29:242–274, 1998.

Owen Lamont. Earnings and expected returns. *The journal of Finance*, 53(5):1563–1587, 1998.

Martin Lettau and Sydney Ludvigson. Consumption, aggregate wealth, and expected stock returns. *the Journal of Finance*, 56(3):815–849, 2001.

Martin Lettau and Sydney C Ludvigson. Expected returns and expected dividend growth. *Journal of Financial Economics*, 76(3):583–626, 2005.

Martin Lettau and Stijn Van Nieuwerburgh. Reconciling the return predictability evidence: The review of financial studies: Reconciling the return predictability evidence. *The Review of Financial Studies*, 21(4): 1607–1652, 2007.

Yan Li, David T Ng, and Bhaskaran Swaminathan. Predicting market returns using aggregate implied cost of capital. *Journal of Financial Economics*, 110(2):419–436, 2013.

Robert S Liptser and Albert N Shiryaev. *Statistics of random Processes: I. general Theory*, volume 5. Springer Science & Business Media, 2013.

Robert E Lucas and Thomas J Sargent. *Rational expectations and econometric practice*, volume 2. U of Minnesota Press, 1981.

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

Ian Martin. The lucas orchard. *Econometrica*, 81(1):55–111, 2013.

Franco Modigliani. The monetarist controversy; or, should we forsake stabilization policies? *Economic Review*, (Spr suppl):27–46, 1977.

Ľuboš Pástor and Robert F Stambaugh. Are stocks really less volatile in the long run? *The Journal of Finance*, 67(2):431–478, 2012.

Daniele Pennesi. Asset prices in an ambiguous economy. *Mathematics and Financial Economics*, 12(1): 55–73, 2018.

Monika Piazzesi and Martin Schneider. Interest rate risk in credit markets. *The American Economic Review*, 100(2):579–584, 2010.

Christopher Polk, Samuel Thompson, and Tuomo Vuolteenaho. Cross-sectional forecasts of the equity premium. *Journal of Financial Economics*, 81(1):101–141, 2006.

Christian Robert. *The Bayesian choice: from decision-theoretic foundations to computational implementation*. Springer Science & Business Media, 2007.

Jules Van Binsbergen, Wouter Hueskes, Ralph Koijen, and Evert Vrugt. Equity yields. *Journal of Financial Economics*, 110(3):503–519, 2013.

Pietro Veronesi. How does information quality affect stock returns? *The Journal of Finance*, 55(2):807–837, 2000.

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