WP 2004-7

Latent Utility Shocks in a Structural Empirical Asset Pricing Model

by

Bent Jesper Christensen & Peter Raahauge

INSTITUT FOR FINANSIERING, Handelshøjskolen i København
Solbjerg Plads 3, 2000 Frederiksberg C
tlf.: 38 15 36 15 fax: 38 15 36 00

DEPARTMENT OF FINANCE, Copenhagen Business School
Solbjerg Plads 3, DK - 2000 Frederiksberg C, Denmark
Phone (+45)38153615, Fax (+45)38153600
www.cbs.dk/departments/finance

ISBN 87-90705-83-1
ISSN 0903-0352

LATENT UTILITY SHOCKS IN A STRUCTURAL EMPIRICAL ASSET PRICING MODEL*

Bent Jesper Christensen
School of Economics and Management
University of Aarhus

Peter Raahauge
Department of Finance
Copenhagen Business School

December 14, 2004

We consider a random utility extension of the fundamental Lucas (1978) equilibrium asset pricing model. The resulting structural model leads naturally to a likelihood function. We estimate the model using U.S. asset market data from 1871 to 2000, using both dividends and earnings as state variables. We find that current dividends do not forecast future utility shocks, whereas current utility shocks do forecast future dividends. The estimated structural model produces a sequence of predicted utility shocks which provide better forecasts of future long-horizon stock market returns than the classical dividend-price ratio.

KEYWORDS: Random utility, asset pricing, maximum likelihood, structural model, return predictability

1. INTRODUCTION

The consumption-based intertemporal asset pricing model has been the work horse in theoretical as well as empirical asset pricing for more than two decades. The seminal theoretical contributions are due to Lucas (1978) and Breeden (1979). The basic model is aimed at explaining the behavior of stock returns over time. Numerous extensions to other markets cover interest rates, bond returns, foreign exchange, and derivative pricing. The basic premise is that a representative investor maximizes expected discounted time-separable utility over an infinite horizon by choice of consumption and asset holdings.

Unfortunately, empirical evidence shows that the model suffers from several shortcomings in terms of explaining actual observations (see Campbell (2000) for a recent critical review and discussion). One of the striking empirical findings about asset market data is that long term stock returns are predicted by the dividend-price ratio (Fama and French (1988)). There is no compelling reason within the intertemporal asset pricing model that precisely this transformation of current dividend and price (namely, the raw ratio) should predict particularly well, although some justification may be given using the loglinear approximation of Campbell and Shiller (1988a). Empirical evidence has given

*We are grateful to Serena Atim Ngoro-Menyang for useful discussions and collaboration at the early stages of this project.

rise to a number of additional puzzles, e.g., the excess volatility puzzle, that returns vary too much relative to subsequent changes in fundamentals (Shiller (1981)), and the equity premium puzzle, that expected stock returns are higher than model predictions at reasonable parameter values (Mehra and Prescott (1985)). In response, several modifications of the basic model have been introduced, drawing, e.g., on habit persistence (Constantinides (1990)), recursive (or other forms of non-separable) utility in the sense of Kreps and Porteus (1978) and Epstein and Zin (1989), or market imperfections (Aiyagari and Gertler (1991)) to explain the observed patterns—see Kocherlakota (1996) for a comprehensive review. At this point, no single modification of the basic model stands out as the preferred specification. Some of the alternative theories are nearly indistinguishable, based on available empirical data. Furthermore, none of the modifications of the model so far has provided any further justification for using the dividend-price ratio as a return forecast, or for choosing some other transformation of dividend and price.

In this paper, we introduce a random utility shock into the fundamental consumption-based intertemporal asset pricing model. Our approach leads naturally to an empirically tractable structural likelihood function. Random utility models are wide-spread and powerful tools in many other disciplines, particularly in microeconometric studies, based on the early work of McFadden (1973) on the static random utility model of discrete choice and the generalization of Rust (1987) to the dynamic optimizing case. Surprisingly, the asset pricing literature is nearly void of applications of this approach. A notable exception is Hansen and Singleton (1983), who also consider a random utility extension of the Lucas (1978) model and carry out likelihood analysis. However, in their paper, they impose an assumption on the joint distribution of consumption, asset returns, and the utility shock, and derive parametric asset pricing restrictions. As noted by Gallant (1987, pp. 437-38), this approach is less satisfactory than full, structural maximum likelihood, i.e., imposing distributional assumptions on consumption and the utility shock, and deriving the resulting endogenous distribution of asset prices, consistent with the Lucas model. We adopt the later approach in this paper.

The likelihood approach ensures asymptotic efficiency in the inference stage and allows easy derivation of powerful tests of distributional misspecification. In our empirical work, the tests fail to reject some of our estimated specifications. Thus, the resulting efficiently estimated structural model is consistent with the data, and although it is based on premises (in particular, a suitable utility shock process) that are merely alternatives to other empirical measures (habit persistence, market imperfections, etc.), we argue that it provides a convenient means of capturing and describing the important intertemporal features of the data.

Our specification, which we label the RUAP (random utility asset pricing) model, generates a stochastic process for the marginal rate of intertemporal substitution (MRIS) which translates into a model predicted asset pricing function matching the observed data on prices and fundamentals perfectly. This model based MRIS is a function of dividend and price which specializes to the raw dividend-price ratio under restrictive conditions that are rejected by the data. Without these restrictions it is shown to forecast future stock returns better than the dividend-price ratio. Thus, our approach provides a theoretical foundation for the derivation of the predictor, and leads to improved empirical forecasting performance.

MRIS in our model takes the form $y_{t+1}d_t/(y_td_{t+1})$, where d_t is the dividend and y_t the random utility shock. The process y_t may in turn be identified from data on asset prices p_t and dividends as

$$y_t = \frac{d_t(b_1 \log(d_t) + k)}{p_t - d_t b_2},$$

where b_1, b_2, and k are known functions of model parameters. It follows that the dividend-price ratio only emerges as a special case when $b_1 = b_2 = 0$. We show that this in turn requires, among others, a serially uncorrelated random utility shock y_t. Our empirical results support serial dependence in y_t, and it is the corresponding generalization (with non-zero b's) of the dividend-price ratio that produces a superior forecast of future long horizon returns. Notably, this improvement is obtained without using the long horizon returns in the model estimation stage, i.e., out-of-sample returns are brought in as a new data element for the forecast comparison.

In sum, our estimation technique produces a predicted utility shock sequence y_t characterizing the features of the data not captured in the basic model and improving forecasting performance. Further structural model developments can now be aimed at explaining our predicted shock sequence. This way, we provide a unifying tool for organizing further asset market research.

The paper is organized as follows. In Section 2, we introduce the random utility asset pricing (RUAP) model. The associated likelihood function is derived in Section 3. Section 4 introduces the data used in our empirical study. Following the literature, we try both dividends and earnings as state variables in the asset pricing model. Section 5 presents the results of estimation and hypothesis testing, and Section 6 concludes.

2. The Random Utility Asset Pricing Model

We consider a random utility extension of the Lucas (1978) asset pricing model. The infinitely lived representative agent is assumed to maximize expected discounted utility

$$E_0 \sum_{t=0}^{\infty} \beta^t u(c_t, y_t),$$

where c_t is consumption, y_t is the random utility shock, β is the discount factor, and u is the period utility function. The optimization is subject to the budget constraint

$$c_t = s_t d_t - p_t(s_{t+1} - s_t),$$

where s_t is current asset holding, s_{t+1} is next period asset holding, p_t is the asset price, and d_t the current dividend. Thus, s_t is determined in the previous period, and the choice variable in period t is s_{t+1}, which in turn determines current consumption via the budget constraint. This is of course based on non-satiation, and we adopt the particular specification

$$u(c_t, y_t) = y_t \log(c_t).$$

That is, preferences are logarithmic, but there is a random utility shift y_t each period. The state variables in period t are s_t, d_t and y_t. Here, s_t is last period's control, and we assume a bivariate vector-autoregressive (VAR) model for the dividends and utility shocks,

$$\begin{bmatrix}\log(d_t)-\mu_d\\ y_t-\mu_y\end{bmatrix}=\begin{bmatrix}a_{dd} & a_{dy}\\ a_{yd} & a_{yy}\end{bmatrix}\begin{bmatrix}\log(d_{t-1})-\mu_d\\ y_{t-1}-\mu_y\end{bmatrix}+\begin{bmatrix}\epsilon_t^d\\ \epsilon_t^y\end{bmatrix}, \tag{1}$$

or, upon defining $\omega_t=(\log(d_t),y_t)'$,

$$\omega_t-\mu=A(\omega_{t-1}-\mu)+\epsilon_t, \qquad \epsilon_t\sim i.i.d., \qquad E(\epsilon_t)=0,$$

where $\mu=(\mu_d,\mu_y)'$ is the vector of unconditional means, A is the matrix of autoregressive coefficients a_{dd}, a_{dy}, a_{yd}, a_{yy}, and $\epsilon_t=(\epsilon_t^d,\epsilon_t^y)'$ is the vector of current innovations to log-dividends and utility shocks. Note that each of the two state variables $\log(d_t)$ and y_t may depend on both current and lagged values of the other. Thus, while a_{dd} and a_{yy} indicate the dependence of log-dividends respectively utility shocks on their own past, a_{dy} allows the past utility shock to condition the conditional mean of current log-dividends, and a_{yd} governs the similar effect of past log-dividends on the current utility shock. A contemporaneous relation between $\log(d_t)$ and y_t results if ϵ_t^d and ϵ_t^y are dependent. We assume that the process ω_t is stationary, i.e., that the eigenvalues of A are less than unity in magnitude. This random utility asset pricing model is henceforth labelled the RUAP model.

Total asset supply is normalized to unity. In equilibrium, $s_t=1$, so $c_t=d_t$. Inserting this condition in the first order conditions for the representative agent's problem leads to the stochastic Euler equation

$$\begin{aligned}p_t&=\beta E_t\left[\frac{u_c'(c_{t+1},y_{t+1})}{u_c'(c_t,y_t)}(d_{t+1}+p_{t+1})\right]\\&=\beta E_t\left[\frac{y_{t+1}d_t}{y_t d_{t+1}}(d_{t+1}+p_{t+1})\right],\end{aligned}$$

which is similar to the Euler equation in the Lucas model, but in addition involves the random utility shock. From this, we get the equilibrium asset pricing function, as stated in the following theorem.

THEOREM 1 *The unique finite equilibrium asset pricing function is given by*

$$p_t=d_t\left(\frac{b_1\log(d_t)+k}{y_t}+b_2\right), \tag{2}$$

where

$$b=[\,b_1\;b_2\,]=[\,0\;1\,]\beta A(I-\beta A)^{-1},$$

$$\begin{aligned}k&=[\,0\;1\,]\Big(\beta I(I-\beta I)^{-1}-\beta A(I-\beta A)^{-1}\Big)\mu\\&=\frac{\beta}{1-\beta}\mu_y-b\mu.\end{aligned}$$

PROOF: From the Euler equation,

$$
\begin{aligned}
p_t &= \beta\frac{d_t}{y_t}E_t[y_{t+1}] + \beta\frac{d_t}{y_t}E_t\left[\frac{y_{t+1}}{d_{t+1}}p_{t+1}\right] \\
&= \beta\frac{d_t}{y_t}E_t[y_{t+1}] + \beta^2\frac{d_t}{y_t}E_t\left[\frac{y_{t+1}}{d_{t+1}}\frac{y_{t+2}d_{t+1}}{y_{t+1}d_{t+2}}(d_{t+2}+p_{t+2})\right] \\
&= \beta\frac{d_t}{y_t}E_t[y_{t+1}] + \beta^2\frac{d_t}{y_t}E_t[y_{t+2}] + \beta^3\frac{d_t}{y_t}E_t\left[\frac{y_{t+2}}{d_{t+2}}p_{t+2}\right].
\end{aligned}
$$

Recursive substitution for future asset prices produces

$$
p_t = \frac{d_t}{y_t}\sum_{i=1}^{\infty}\beta^i E_t[y_{t+i}],
$$

using that the eigenvalues of A and hence of βA are less than unity in magnitude, so that $\sum_i \beta^i A^i$ is summable. In particular,

$$
E_t[\omega_{t+i}] = A^i(\omega_t - \mu) + \mu,
$$

so that

$$
E_t\left[\sum_{i=1}^{\infty}\beta^i\omega_{t+i}\right] = \beta A(I-\beta A)^{-1}(\omega_t-\mu) + \beta\mu(1-\beta)^{-1}.
$$

Therefore, writing $y_{t+i} = [\,0\ 1\,]\omega_{t+i}$, we have

$$
\begin{aligned}
p_t &= \frac{d_t}{y_t}E_t\left[\sum_{i=1}^{\infty}\beta^i[\,0\ 1\,]\omega_{t+i}\right] \\
&= \frac{d_t}{y_t}[\,0\ 1\,](\beta A(I-\beta A)^{-1}(\omega_t-\mu) + \beta(1-\beta)^{-1}\mu) \\
&= \frac{d_t}{y_t}(b\omega_t + k) \\
&= \frac{d_t}{y_t}(b_1\log(d_t) + b_2 y_t + k),
\end{aligned}
$$

with b and k as defined in the Theorem. *Q.E.D.*

The Theorem yields a closed form asset pricing function $p_t = p(d_t, y_t)$, thus facilitating analysis. The functional form depends on the structural parameters β, A and μ through the three scalars b_1, b_2, and k. The matrix expressions for these are given in the Theorem, and they may be spelled out explicitly in terms of the individual scalar entries in A as

$$b = \left[\frac{\beta a_{yd}}{1 - \beta \mathrm{tr}(A) + \beta^2 |A|} \quad \frac{\beta \left(a_{yy} - \beta (a_{dd} a_{yy} - a_{dy} a_{yd}) \right)}{1 - \beta \mathrm{tr}(A) + \beta^2 |A|} \right]$$
$$(3) \qquad = \left[\frac{\beta a_{yd}}{1 - \beta \mathrm{tr}(A) + \beta^2 |A|} \quad \frac{\beta \left(a_{yy} - \beta |A| \right)}{1 - \beta \mathrm{tr}(A) + \beta^2 |A|} \right],$$

$$k = \frac{\beta}{1-\beta} \mu_y - \frac{\beta \left(a_{yd} \mu_d + (a_{yy} - \beta |A|) \mu_y \right)}{1 - \beta \mathrm{tr}(A) + \beta^2 |A|},$$

where tr($\cdot$) denotes the matrix trace and $| \cdot |$ the determinant. We note also the alternative forms of the numerator

$$\begin{aligned} 1 - \beta \mathrm{tr}(A) + \beta^2 |A| &= 1 - \beta (a_{dd} + a_{yy}) + \beta^2 (a_{dd} a_{yy} - a_{dy} a_{yd}) \\ &= (1 - \beta a_{dd})(1 - \beta a_{yy}) - \beta^2 a_{dy} a_{yd}. \end{aligned}$$

Importantly, the distributional form of the VAR shocks ϵ_t does not matter for the functional form of the asset pricing function. Of course, the asset pricing function is time invariant, i.e., we may analyze the function $p(d, y)$ in general, dropping explicit time subscripts. We collect several results in a corollary.

COROLLARY 1 *For any realization $y > 0$ of the random utility shock, the asset pricing function $d \to p(d, y)$ satisfies*

$$\lim_{d \to 0} p(d, y) = 0.$$

For sufficiently large $d > 0$, the asset pricing function is strictly increasing and strictly convex, with elasticity strictly in excess of unity, provided that $b_1, b_2 > 0$.

PROOF: For non-zero y and b_1, the limit result follows from l'Hospital's rule applied to the ratio $\log d / d^{-1}$. Differentiation of numerator and denominator yields $-d^{-1}/d^{-2} = -d$, which is 0 at $d = 0$. Differentiation of (2) produces

$$p'_d(d, y) = \frac{b_1 (\log(d) + 1) + k}{y} + b_2,$$

and the second assertion follows by taking $d > \exp\left(-(1 + k/b_1)\right)$, provided $b_1, b_2 > 0$. Further differentiation yields

$$p''_{dd}(d, y) = \frac{b_1}{dy} > 0,$$

showing strict convexity. Elasticity in excess of unity, i.e., $p'_d \; d/p > 1$, is equivalent to $p'_d > p/d$ and so follows from strict convexity and $p(0, y) = 0$. *Q.E.D.*

Non-negative entries in A and $A \neq 0$ are sufficient conditions for $b_1, b_2 > 0$. Non-negative entries in A imply that $\beta^i A^i$ and hence it's sum has non-negative entries, and b_1 and b_2 are the lower left and right corners of this matrix. It becomes an empirical question whether A has non-negative entries, or, more generally, whether $b_1, b_2 > 0$. We now turn to an empirical investigation of this model.

3. The Likelihood Function

Consider a dataset of the form $\{d_t, p_t\}_{t=0}^T$. We wish to draw inference on the parameters of the RUAP model from the conditional log likelihood function for $\{d_t, p_t\}_{t=1}^T$, given the initial observations (d_0, p_0). Here, we adopt the additional assumption that $\epsilon_t \sim iiN(0, \Sigma)$, the bivariate normal distribution, with Σ a positive definite variance-covariance matrix. The normality assumption is not rejected by the data in some of our specifications, but it is clear from below how to derive the likelihood function from alternative distributional assumptions, given that the asset pricing function from Theorem 1 does not depend on the distribution of ϵ_t. Of course, with normality, the random utility shock y_t could in principle take on the value zero, or even negative values, but for σ_{yy} (lower right corner element of Σ) sufficiently small relative to the mean utility shock μ_y, the likelihood of such an occurrence is diminutive. Ultimately, in practice, we are most interested in the estimated version of the corresponding empirical model. Here, μ_y may be normalized to unity without loss of generality, since the expected utility analysis is invariant to affine transformations of u. Given $\mu_y = 1$, the likelihood function will keep the estimate of σ_{yy} away from any value that implies an appreciable probability of zero or lower values for y_t. Furthermore, the mean μ_d of log-dividends vanishes in the empirical model, since in the implementation a common exponential detrending procedure is applied to dividends and asset prices. Thus, the parameters to be estimated are β and the elements of A and Σ, a total of eight.

The conditional RUAP model log likelihood function for the sample $\{d_t, p_t\}_{t=1}^T$ given (d_0, p_0) takes the form

$$\begin{aligned} \ell(d,p,\theta) &= \log(T) - \log(2\sqrt{\pi^2|\Sigma|}) \\ &\quad + \sum_{t=1}^{T} \log\left|\frac{y_t^2}{d_t^2(b_1\log(d_t)+k)}\right| - \sum_{t=1}^{T} \tfrac{1}{2}\epsilon_t^\top \Sigma^{-1}\epsilon_t, \end{aligned} \tag{4}$$

where the term $|y_t^2/(b_1\log(d_t)+k)|$ is the absolute value of the inverse determinant of the Jacobian of the map from the error terms $\epsilon_t = (\epsilon_t^d, \epsilon_t^y)^\top$ in (1) to the data (d_t, p_t). Note that the full dependence on data and parameters is somewhat implicit in (4). The details are as follows. From Theorem 1, the asset pricing function is readily inverted with respect to y_t, allowing identification of the random utility shock in terms of data and parameters as

$$y_t = \frac{d_t(b_1\log(d_t)+k)}{p_t - d_t b_2}. \tag{5}$$

Here, b_1, b_2 and k are explicit functions of the structural parameters in β and A, given in the Theorem. Thus, y_t is directly computable and may be inserted in the Jacobian term in (4). Also, with y_t given, along with data and parameters, ϵ_t may be solved recursively from the vector-autoregression (1), for $t \geq 1$, using the observed $\log(d_0)$ along with y_0 identified from d_0, p_0 and parameters to start the iterations. This completes the computation of the log likelihood function.

The expression for the Jacobian above may be understood as follows. From (1), the map from ϵ_t to $(\log(d_t), y_t)$ has unit Jacobian. The map from $(\log(d_t), y_t)$ to (d_t, y_t) has

Jacobian d_t. Using Theorem 1, the Jacobian matrix of the map from (d_t, y_t) to (d_t, p_t) is given by

$$\mathcal{J}_t = \begin{bmatrix} 1 & 0 \\ \frac{b_1(1+\log(d_t))+k}{y_t} + b_2 & -\frac{d_t(b_1 \log(d_t)+k)}{y_t^2} \end{bmatrix},$$

so the relevant absolute determinant is $||\mathcal{J}_t|| = |\partial p_t/\partial y_t| = |d_t(b_1 \log(d_t) + k)/y_t^2|$. Multiplication and inversion of the Jacobians produces the term $|y_t^2/(d_t^2(b_1 \log(d_t) + k))|$ used in (4).

Intuitively, if data had been available directly on ϵ_t, but not on (d_t, p_t), then the parameters that could be identified would be those in the distribution for ϵ_t, namely Σ, only. If the available data had been on (d_t, y_t), instead, the vector-autoregression could have been analyzed, too, and both Σ and A would be identified. In reality, we have data on (d_t, p_t), so potentially the parameters in the map from (d_t, y_t) to (d_t, p_t) can be identified, in addition. Clearly, both β and A enter this map, but identification of the remaining structural parameter β requires in addition that this enters in a suitably non-redundant way in the asset pricing function, or, equivalently, in (5). In fact, from the expressions for b_1, b_2 and k, β is a strong determining factor for these, and hence for the asset pricing function, as might be expected, and all eight parameters are identified.

The maximum likelihood estimator (MLE) $\hat{\theta}$ of $\theta = (\beta, A, \Sigma)$ is determined by maximizing the log likelihood function iteratively using a numerical scheme similar to Newton-Raphson. The actual implementation uses the BFGS (Broyden-Fletcher-Goldfarb-Shanno) algorithm (see Coleman and Li (1996)), which applies rank-two updates to the Hessian and line search in each iteration and is more efficient than the Davidon-Fletcher-Powell algorithm. The BFGS algorithm allows using nonlinear constraints to avoid that the iterations diverge, but the final optima were interior in all cases, and asymptotic standard errors calculated off the squareroots of minus the inverse Hessian $-H^{-1}$ at $\hat{\theta}$. Under the distributional assumptions, the MLE is consistent, asymptotically normal and efficient, $\sqrt{T}(\hat{\theta} - \theta) \overset{\sim}{\to} N(0, i^{-1})$, where i is the information matrix and i^{-1} the Cramer-Rao lower bound. The indicated procedure yields consistent estimates of the asymptotic standard errors.

Since the asset pricing function from Theorem 1 does not depend on normality of the error terms ϵ_t, the log likelihood functions corresponding to alternative distributional assumptions would involve the same Jacobian and the same ϵ_t's as in (4). This point is important. In particular, the only parameters in the Jacobian and in ϵ_t are those from the structural model, namely, β and A. The parameters in Σ do not enter here. Changing distributional assumption therefore amounts to using some other bivariate log density $\log f(\epsilon_t, \widetilde{\Sigma})$ instead of the Gaussian $-\log(2\sqrt{\pi^2|\Sigma|}) - \frac{1}{2}\epsilon_t^\top \Sigma^{-1}\epsilon_t$. The revised parameter set $\widetilde{\Sigma}$ could include the variance-covariance matrix Σ, as well as other non-structural (e.g. shape) parameters. In our empirical work below, we find that the normality assumption is adequate for ϵ_t^y for the full data period, whereas normality of ϵ_t^d applies in subperiods (skewness presents no problem, but the variance is changing, leading to overall excess kurtosis). We focus on the structural model issues and retain the Gaussian likelihood function throughout. For robustness against departures from distributional assumptions, we compared standard errors to those based on the squareroots of the sandwich-type (quasi-maximum likelihood) estimator $H^{-1}BH^{-1}$, where B is the sum of outer products of the individual score contributions, and the differences were negligible.

4. DATA

We use the data on stock prices, dividends, and earnings provided by Robert J. Shiller.[1] The data span the period from 1871 to 2000. The stock prices are January levels of the S&P Composite Stock Price Index. The model has been cast in terms of dividends, following Lucas (1978), but we do consider earnings data as a possible alternative to dividend data, following, e.g., LeRoy and Porter (1981) and Campbell and Shiller (1988a). The dividends and earnings series are based on annual data from the S&P Statistical Service Security Price Index Record back to 1926, and annual data from Cowles and Associates (1939) before that. All series are deflated using the CPI. For comparison with the literature as well as to achieve robustness with respect to possible different behavior during the apparent stock market bubble period starting in 1995, we also consider a pre-bubble subperiod, ending in 1994.

The deflated series are transformed to deviations around a common long-run exponential trend $\tau_t = e^{\alpha_0 + \alpha_1 t}$ by measuring their values as relative to this. The coefficients are determined by regression using raw log-dividends, $\log(d_t^{raw}) = \alpha_0 + \alpha_1 t + u_t$, and the series actually used in the analysis are $d_t = d_t^{raw}/\tau_t$ for dividends and $p_t = p_t^{raw}/\tau_t$ for prices. The transformation implies that the zero-mean condition $E(\log(d)) = 0$ on dividends is without loss of generality. The trend coefficients are reestimated for each period under consideration, and also reestimated when using earnings in place of dividends. The α_1-estimate for the full period is 1.0245% for dividends and 1.4388% for earnings. Figure 1 shows the resulting full period series for p_t and d_t in standarized form (zero mean, unit variance).

5. ESTIMATION RESULTS

In the following, we consider the results of structural estimation and model selection using annual data for the full 1871-2000 period, as well as for the different subperiods. In addition, we consider the forecasting properties of the utility shock sequence y_t identified from the estimated model, and we compare with results using earnings in place of dividends.

5.1. RESULTS FOR THE FULL PERIOD 1871-2000

Table 1 shows the estimation results for the full period from 1871, using dividends (as opposed to earnings) for d_t in the model. The first line shows the results from estimation of the full model, with eight parameters. The discount factor β is estimated at .955, significantly below unity at conventional levels (asymptotic standard errors in parentheses). The parameters a_{dd} and a_{yy} are estimated at just below .8, and quite precisely, indicating strong persistence in both state variables, but no unit root problem. The parameter a_{dy} is significantly negative, judged from the estimated standard error. This suggests that

[1] See www.econ.yale.edu

current, unobserved utility shocks impact future consumption (dividends). In particular, a higher marginal utility function now implies lower future consumption. This is consistent with rational, maximizing behavior, since the marginal utility process is mean reverting ($a_{yy} < 1$), so that current increases are viewed as temporary and thus yield an incentive to consume now rather than later. Of course, in equilibrium, the dividends supplied must be consumed, and the estimation procedure generates fitted state variables rationalizing the observed dividends as consumption choices. The literature is rich on examples of observed variables that do not predict future dividends, which are typically taken as driven by an exogenous process, as in the original Lucas (1978) model. Our results identify a variable y_t which does predict future dividends. This causality is unidirectional: From Table 1, the parameter a_{yd} is insignificantly different from zero, indicating that lagged dividends do not condition the distribution of utility shocks.

Of the three parameters in Σ, the variance-covariance matrix of the disturbances, the two individual variances σ_{dd} and σ_{yy} of d and y are significant. In particular, the utility shock is, indeed, non-degenerate and random, consistent with the need for this random utility generalization of the basic asset pricing model. Shocks to marginal utility and consumption are contemporaneously positively correlated, as might be expected, but the relevant parameter, σ_{dy}, is only borderline significant, based on the associated standard error (asymptotic t-statistic of 1.97).

Each of the following lines in Table 1 shows the results of estimation under a particular parametric restriction. In turn, each of the parameters a_{dy}, a_{yd} and σ_{dy}, and all combinations hereof, are restricted to zero. The last column shows the value of the maximized log likelihood function, for each of the specifications. Likelihood ratio tests confirm the conclusions based on the asymptotic standard errors, that marginal utility causes consumption ($a_{dy} < 0$), but not the reverse ($a_{yd} = 0$ is not rejected), and that the contemporaneous conditional correlation of the two is barely significant ($\sigma_{dy} \approx 0$). In view of the weak correlation, it seems reasonable to consider also the joint test of $H_0 : a_{yd} = 0, \sigma_{dy} = 0$ against the full model in the first line. The estimates under these restrictions appear in line 7 of Table 1. The test fails to reject at conventional levels ($-2 \log Q = 4.85$, with a p-value of 8.8% in the χ^2-distribution on two degrees of freedom). None of the other parameters are particularly sensitive to the restrictions imposed. The final model is that in the third line of the Table, with seven parameters, or that in the seventh line, with six parameters, and the implications of the two are similar. Both models imply that $a_{yd} = 0$, that is, current dividends do not predict future utility shocks. In the model in line seven, $\sigma_{dy} = 0$ in addition, implying also lack of contemporaneous conditional correlation between dividends and utility shocks. In either model, the restriction $a_{dy} = 0$ is strongly rejected, i.e., dividends are indeed predicted by previous utility shocks. Thus, the relation between marginal utility and dividends is predominantly dynamic in nature, and negative ($a_{dy} < 0$, $\sigma_{dy} \approx 0$).

Figure 1 shows the time series of predicted values of random utility shocks y_t from the structural model, using parameter estimates from the unrestricted model (first line of Table 1), together with p_t and d_t. Evidently, y_t is countercyclical, corresponding to the negative estimate of a_{dy} in the model. Low marginal utility of current consumption spurs investment, leading to increases in asset prices and future dividends.

Figure 2 shows the standardized series for the two stochastic shocks ϵ_t^y and ϵ_t^d in the model. The figure shows that although there are signs of positive contemporaneous correlation, the effect is not particularly strong. This confirms the estimation results, where σ_{dy} is found to be at best borderline significant. Figure 2 furthermore reveals a data break around 1959, when the procedure for calculation of the CPI deflator changed. After the break, both error terms appear smoother and less volatile. This change in volatility could lead to apparent excess kurtosis in normality tests based on the marginal distribution. This is investigated in Figure 3. The figure shows the histograms and fitted normal distributions for ϵ_t^y (panel (a)) and ϵ_t^d (panel (b)). There is no problem with the normality assumption for ϵ_t^y. The joint Jarque-Bera test for lack of skewness and excess kurtosis (also shown in the figure) takes the value 4.07, for a p-value of 13% in the asymptotic χ_2^2-distribution. The separate tests for no skewness and kurtosis also fail to reject (p-values of 5.3% and 58%, respectively). From panel (b), for ϵ_t^d, skewness is not a major problem (p-value of 4.5%), but the series does exhibit excess kurtosis, thus confirming the impression from Figure 2. We reestimated the model for the separate subperiods before and after the 1959 break in CPI calculation, and found no departures from normality within subperiods. Figure 4 shows the distributions and (insignificant) test statistics for the most recent subperiod, after 1959. From Section 2, the asset pricing function does not depend on normality, so the expressions for the calculated ϵ_t-series and the Jacobian in the likelihood function are invariant to distributional assumptions (only their numerical values would change with the parameters β and A), and we do not pursue the possibility of modifying the distributional assumptions any further. Instead, we focus on results on the structural model, using the Gaussian conditional log likelihood function, and inference and conclusions are similar whether using standard errors based on the negative inverse Hessian or the sandwich-style QML formula (Section 3).

5.2. The pre-bubble period 1871-1994

The finding $a_{dy} < 0$ and the implied predictive power of y_t with respect to future dividends and prices is a strong result on the structure of asset pricing in this market. It is clearly important to assess the robustness of the result. For one, it may be speculated that the findings are driven by the apparent bubble effect that characterized stock markets since the mid-nineties and until the end of our sample period. To examine this issue, we reestimate the model after discarding all observations from 1995 and later.

The results appear in Table 2 , which is laid out as Table 1. The parameter estimates and standard errors are very similar to those in Table 1. The estimates of a_{dd} and a_{yy} are slightly lower than in Table 1, as expected, since the bubble has been removed, but the difference is very slight. The point estimate of a_{yd} is now negative, but again highly insignificant. As in Table 1, the final model is that in the third or seventh line, where $a_{yd} = 0$, and possibly $\sigma_{dy} = 0$ (seventh line). Specifically, the test of the model in the seventh line against that in the first now takes the value 4.052, with a p-value of 13.2%. In the model in line seven, the parameter a_{dy} remains negative and strongly significant. Thus, the main findings, that y_t predicts d_t, but not the reverse, and that the contemporaneous conditional correlation between the two is weak, are robust to removing the bubble period from the data.

5.3. Out-of-sample long-horizon return regressions

The observation that our model produces a utility shock sequence y_t that predicts dividends and prices suggests that it may be used in a more general return forecasting context. In the literature, there has been a strong interest in predictions of multi-period returns from the dividend-price ratio or dividend yield, see, e.g., Fama and French (1988), henceforth FF. In particular, this variable has been found to have more predictive power with respect to long-term stock returns than other candidates, including interest rate related variables. Following FF, the relevant regression specification takes the form

$$r_{t+1} + \cdots + r_{t+K} = \alpha_K + \beta_K z_t + \epsilon_t, \tag{6}$$

where r_{t+j} is the log stock return j periods in the future, and z_t is the current variable used to predict the K-period stock return. Given our results, it is natural to inquire how well our estimated y_t series fares in comparison with the dividend-price ratio d_t/p_t in place of z_t in the regression. In particular, for $b_1 = b_2 = 0$, it follows from (5) that our model based y_t reduces to the dividend-price ratio d_t/p_t. Our empirical results are consistent with $a_{yd} = 0$, and hence $b_1 = 0$, using the expressions for the b-coefficients in (3). However, given $a_{yd} = 0$, the condition $b_2 = 0$ requires that either $a_{yy} = 0$ or $\beta a_{dd} = 1$. Clearly, both may be rejected, based on our empirical results in Tables 1 and 2. Thus, utility shocks are significantly positively autocorrelated, and both the discount factor β and the autocorrelation parameter of log-dividends (in deviations from the exponential trend) a_{dd} fall short of unity. We conclude that y_t differs statistically from d_t/p_t.[2] It remains an empirical matter whether it also forecasts long-horizon stock returns better.

Simply sticking the y_t-series considered e.g. in Figure 1 into the above specification (6) in place of z_t would produce a regression of the in-sample variety, in that the structural parameter estimates used in calculating the y_t-variables are based on past as well as future information. This would in part give the y_t-forecast the benefit of hindsight. Therefore, we adopt an out-of-sample approach, instead, for our forecasting analysis. Thus, for each t, starting at $t = 20$ (the year 1890), we now carry out a full structural estimation using only data through t, and use the resulting parameter estimates when calculating y_t. These consecutive estimations therefore use only information available to investors at t in constructing y_t. This provides for a fair comparison with the alternative forecast d_t/p_t, which is also part of the same information set. In particular, the question is whether the estimated structural asset pricing model may be used to produce a better forecast of subsequent returns than that entailed in the dividend-price ratio at t, using only information available at t.

We implement the regression for the full period, with the forecast horizon K ranging from 1 to 4 years. The results are shown in Table 3. Panel A shows results using y_t based on the model from the first line in Table 1, with eight parameters. Panel B shows results using y_t from the model in line seven of Table 1, with six parameters ($a_{yd} = \sigma_{dy} = 0$ imposed). The first two columns show the univariate regression results, where y_t plays the role of z_t in the regression (6). In both panels, the estimated regression coefficients

[2] In relation to Corollary 1, the parameter estimates in the first line of Table 1 give rise to $b_1 = .27$, $b_2 = 3.09$, and $k = 18.13$, and those in the seventh line yield $b_1 = 0$, $b_2 = 3.12$, and $k = 18.10$. For the unrestricted parameters, the conditions $b_1, b_2 > 0$ from the Corollary are satisfied, even though the entries of A are not all non-negative, and since $b_2 > 0$ for both parameter sets, y_t does not reduce to d_t/p_t.

increase with the forecast horizon, from about .12 for the one-year forecast, $K = 1$, to about .47 for the four-year forecast, $K = 4$. Similarly, the associated t-statistics (in parentheses) increase from 1.9 ($K = 1$) to 2.4 ($K = 4$), and adjusted R^2 increases from 3% to 11%. All t-statistics are based on Newey and West (1987) (NW) heteroskedasticity and autocorrelation consistent standard errors. In particular, for K-period returns, there is an overlap of length $K - 1$ among consecutive left hand side variables, and we use the NW estimator with lag length $K - 1$.

The next two columns of Table 3 show the results when the dividend-price ratio $z_t = d_t/p_t$ enters the regression (6) instead of y_t from our model. These results are common across the first two panels, as our model plays no role for these univariate regressions. As expected, the regression coefficients are different from the coefficients on y_t. Consistent with FF, the coefficients increase with the forecast horizon. The t-statistics and R^2 also increase with K, from 1.6 to 2.2, and from 2% to 9%, similarly to the $z_t = y_t$ case. The common pattern of increase in coefficient with forecast horizon and, particularly, the very similar t-statistics and R^2 suggest that y_t has as much predictive power for long-horizon stock returns as the dividend-price ratio.

We also consider the encompassing specification

$$(7) \qquad r_{t+1} + \cdots + r_{t+K} = \alpha_K + \gamma_K y_t + \delta_K \frac{d_t}{p_t} + \epsilon_t.$$

The final three columns of Table 3 report the results. Quite strikingly, the coefficient γ_K on y_t is positive and significant and the coefficient δ_K on the dividend-price ratio d_t/p_t is negative throughout. This indicates that y_t is actually a better forecast of long-horizon returns than the dividend-price ratio.

We also carry out the analysis using the log dividend-price ratio $\log(d_t/p_t)$ instead of the raw ratio d_t/p_t. This may provide a useful comparison and robustness check, for at least two reasons. Firstly, the log-transformed series may be expected to have better statistical properties than the untransformed series. Secondly, in Campbell and Shiller (1988a)'s present value framework using loglinear approximation, $\log(d_t/p_t)$ comes out as a cointegration residual suitable e.g. as explanatory variable in forecasting regressions. In contrast, the similar residual in the present value model without loglinear approximation would be $d_t/(1 + R) - p_t$, but this requires that R is a constant conditional expected return. As the log dividend-price ratio emerges from a model not assuming constant conditional expected returns and is free of unknown parameters (such as R), it provides a useful benchmark forecast of long-horizon returns, to compare y_t from our model against. To be sure, Campbell and Shiller (1988a) worked with data that were not detrended, but as our detrending procedure consists of dividing each series by a common deflator τ_t, this cancels when forming the ratio d_t/p_t, and the logarithm of this.

The results of the forecast comparison with the log dividend-price ratio are shown in the last two panels of Table 3. The first two columns are unchanged from the panels above, as y_t from the RUAP model is unchanged. Thus, Panel C corresponds to Panel A, using the same estimated model from Table 1, and Panel D corresponds to Panel B. The next two columns in each panel show the results for the new variable $\log(d_t/p_t)$. Using this as the explanatory variable z_t in the univariate regression (6) produces coefficients that increase with K, from .06 to .34, and are relatively close to the similar coefficients from univariate regression on y_t from our model. Recalling the functional dependence of y_t

on p_t and both d_t and $\log d_t$, it is clear that y_t can carry similar information to that in the log dividend-price ratio. In terms of inference, we find that t-statistics and R^2 for the univariate regressions on $\log(d_t/p_t)$ again are similar to those from the regressions on y_t and d_t/p_t, although in this case in fact slightly lower. The last columns show the results of the encompassing regression analysis, substituting $\log(d_t/p_t)$ for d_t/p_t in (7). Here, the coefficient γ_K on y_t is again significantly positive and increasing in the forecast horizon K. On the other hand, the coefficient δ_K on $\log(d_t/p_t)$ gets a perverse negative sign throughout.

The results suggest that y_t is a useful forecast of long-horizon stock returns, subsuming the information content of the dividend-price ratio (dividend yield), whether or not the log-transform is applied to the latter.

5.4. Pre-bubble forecast regressions

To test the robustness of the results from the forecast regressions, we again consider the pre-bubble period 1871-1994, as in Table 2 above. The results appear in Table 4, which is laid out as Table 3. The time period used for the forecast regressions in FF was 1927 to 1986. Ending in 1986 of course avoids the bubble issue, too. While we include more data both before and after the FF period, eliminating the bubble-period still allows us to compare results more directly with FF than when the bubble-period is included. Continuing until 1994 instead of 1986 in addition facilitates comparison with results in Campbell, Lo, and MacKinlay (1997, p. 269) (CLM), who use a 1927-1994 period. In fact, from Table 4, parameter estimates, t-statistics and R^2 are now quite close to those reported in FF, and especially to those in CLM, for the comparable regressions with $z_t = d_t/p_t$, respectively $z_t = \log(d_t/p_t)$. We find coefficients on d_t/p_t increasing with K from 3.1 to 10.8, with t-statistics from 2.8 to 3.5, and R^2 from 5% to 16%. FF report estimates for $K = 1$ to 4 ranging from 3.4 to 14.4, with t-statistics and R^2 growing from 1.7 to 3.3 and from 3% to 29%. The source of differences in results is the remaining difference in periods, after eliminating the bubble-period. FF report results for both nominal and real returns, and for the dividend yield defined as either d_t/p_t or d_t/p_{t-1}. As our data are deflated by the CPI, our results are most comparable to their results for real returns. As we use the Shiller data, where the stock price is measured at the beginning of the year or traded cum-dividend, our results are most comparable to their results for d_t/p_{t-1}. Overall, the FF results for the two versions of the dividend yield are very similar in any case, and close to ours, and we get results essentially identical to theirs for d_t/p_{t-1} when restricting attention to the 1927-1986 period (not reported).

The alternative univariate regression on y_t from our model (first columns of Table 4, Panels A and B) produces t-statistics and R^2 that are slightly higher than those obtained using $z_t = d_t/p_t$ (following two columns). The multivariate regression for the full eight-parameter model (last three columns of Table 4, Panel A) shows that, again, y_t is a better forecast than the dividend-price ratio. The coefficient on y_t is positive and significant, and that on the dividend-price ratio negative. The results from the reduced six-parameter model (Panel B) are similar.

We get results even closer to those in CLM in the last two panels, C and D, as they used data through 1994, and applied the log-transform to the dividend-price ratio. Thus, for the univariate regression with $z_t = \log(d_t/p_t)$, we get coefficients increasing from .15 to .51, with t-statistics from 2.8 to 3.4, and R^2 from 5% to 16%. CLM report coefficients from .20 to .65, with t-statistics from 2.3 find to 4.6 and R^2 from 7% to 26%. As before, t-statistics and R^2 are higher in the univariate regressions on y_t (first columns of each panel) than on $\log(d_t/p_t)$. Again, the coefficient δ_K on $\log(d_t/p_t)$ turns negative in the multivariate regression, whereas the coefficient γ_K on y_t is positive, and for large K also statistically significant.

The comparisons with both FF and CLM show that our data are not unusual. We get similar results for comparable univariate regressions. Although our multivariate results may be plagued by multicollinearity to some extent, the predictive power of y_t from our model appears from both the univariate and multivariate out-of-sample forecasting results to be stronger than that of d_t/p_t or $\log(d_t/p_t)$ for long-horizon return forecasts.

5.5. EARNINGS DATA

In part of the literature on intertemporal asset pricing, earnings are used in place of dividends (see, e.g., LeRoy and Porter (1981) and Campbell and Shiller (1988b)). It is of interest to investigate which of our results carry over to the earnings case. Thus, we reestimate our model for the full period, as in Table 1, with dividends replaced by earnings throughout.

The results based on earnings appear in Table 5. Clearly, there are some differences in inferences, compared to Table 1. In the full model, first line of Table 5, both a_{dy} and a_{yd} are negative and insignificant, whereas σ_{dy} is strongly significant. It is as though the predictive effect of y_t with respect to dividends in Table 1 ($a_{dy} < 0$) has been replaced by a contemporaneous conditional correlation between y_t and earnings, consistent with the notion that dividends adjust sluggishly, relative to earnings. The final model is now that in the second or third line, setting either a_{dy} or a_{yd} to zero, which in either case then leaves the opposite parameter negative and significant. Thus, the data reveal some information on a link between y_t and earnings: One of the two must forecast the other.

Since the results are less clearcut than in the dividend case, we look to the long-horizon forecasting regressions for additional information. Table 6 is laid out as Table 3. Earnings replace dividends everywhere in the regressions. As in the dividend case, coefficient estimates, t-statistics and R^2 all increase with the forecast horizon K in the univariate regressions. This is so in all models considered, whether y_t from our model or the earnings-price ratio is used as explanatory variable, and with or without the log-transform. In Panel A (full eight-parameter model), the t-statistics for the coefficients on y_t range from 2.5 to 2.8, and the associated adjusted R^2-ratios increase from 5% ($K = 1$) to 13% ($K = 4$). The corresponding t-ratios and R^2 are quite similar in the regressions using the earnings-price ratio instead of y_t (next two columns of the table), when using the log earnings-price ratio (Panel D) and when using y_t from the restricted models, with $a_{dy} = 0$ in Panels B and E, and $a_{yd} = 0$ in Panels C and F. Turning to the multivariate regressions (last three columns of the table), the picture is somewhat as in the dividend case, albeit a bit more mixed. In Panels A and B, the coefficients γ_K on y_t and δ_K on d_t/p_t are both insignificant for $K = 1$. The coefficient γ_K on y_t and it's t-statistic are positive and increasing in K, reaching $t = 1.8$ and 2.0 for $K = 4$ in the two panels, whereas the earnings-price ratio gets a perverse negative coefficient in the multivariate regressions for $K > 1$. These results are essentially the same when using the log earnings-price ratio (Panels D and E), and they suggest that y_t is a better forecast of returns than the earnings-price ratio. The results from Panels C and F, imposing $a_{yd} = 0$, are different, and suggest that the earnings-price ratio is a better forecast than y_t.

All in all, the conclusion is that y_t from our model predicts long-term stock returns out-of-sample at least as well as the dividend-price ratio, the earnings-price ratio and their logarithms, and actually better in several of our specifications. Specifically, when using dividends rather than earnings, the results uniformly support y_t as the best forecast. When using earnings data, instead, the results depend on the precise version of the RUAP model adopted (whether earnings depend on lagged utility shocks or the reverse), but the results for the full, unrestricted model (mutual dependence) again suggest that y_t is the best forecast, subsuming the information content of the earnings-price ratio and it's logarithm.

5.6. Out-of-Sample Forecast Errors

The forecasts considered in the long-horizon regressions are out-of-sample forecasts in the sense that both y_t and d_t/p_t are based on data available to investors at time t. In particular, the parameters that enter y_t are estimated from prevailing data at time t. On the other hand, the regression coefficients are estimated using all data. Thus, the overall forecast, including regression coefficients, namely, $\alpha_K + \beta_K z_t$ in (6), where z_t may be y_t or d_t/p_t, is not a true out-of-sample forecast. This issue is addressed, e.g., in FF and in Goyal and Welch (2003). In these papers, forecast errors are calculated by reestimating the regression coefficients α_K and β_K every period, using only prevailing data at each point in time. In order to investigate whether y_t continues to perform as well as, or better than, the dividend-price ratio, even in the sense of producing smaller forecast errors when forecasting regressions use only prevailing data at each point in time, we replicate the procedure from these papers.

Our y_t series starts in 1891, since the first 20 years are set aside for the first structural estimation of the RUAP model. This is reestimated in each of the following years, using prevailing data, thus allowing construction of the y_t series, $t = 1891, ..., 2000$, exactly as in the previous subsections 5.3-5.5. Next, the long-horizon regressions are implemented, but using only data on y_t and d_t/p_t from $t = 1891, ..., T$, starting with $T = 1911$, and running new regressions for each $T = 1911, ..., 2000-K$, where K is the forecast horizon. Each regression is used to calculate a single true out-of-sample forecast error, namely, the forecast error obtained in period $T + 1$ if using coefficients estimated based on prevailing data at T.

The resulting distribution of out-of-sample forecast errors is summarized in Table 7. What is reported is the mean squared forecast error R^2, following FF. Thus, $R^2 = 1 - \text{MSFE}/s^2$, where MSFE is the mean squared forecast error across the periods used, and s^2 is the variance of the long-term return being forecast. Note that this R^2 is not the usual coefficient of determination for any single regression, but is calculated from a sequence of forecast errors, each based on a separate regression using prevailing data. The first column of the table indicates the forecast horizon K. Following Goyal and Welch (2003), we show in the second column the R^2 obtained when using the prevailing mean of the long-term return as a forecast of the future return. The third column shows the R^2 that obtains when using our RUAP model forecast y_t (full, unrestricted model). The last two columns show the results when using the dividend-price ratio or its logarithm. Panel A shows the full period results. Panel B shows the results using the pre-bubble period ending in 1994. Panel C shows the results using earnings in place of dividends throughout. For sufficiently long forecast horizons, $K = 3$ and 4, the highest R^2 is in all cases obtained using the forecast based on our y_t. For $K = 1$ and using dividends, the prevailing mean is actually the best forecast in the full period. The same is true for $K = 2$, but here, this forecast becomes the worst when removing the bubble period (Panel B). In general, the results for $K = 1$ and 2 are quite mixed.

Although the superiority of y_t for longer horizons is uniform, based on the raw R^2 values, Diebold and Mariano (1995) tests for differences in performance between the prevailing mean and the alternative forecasts are all insignificant, taking values in the range from $-.76$ to 1.05. This result is similar to that in Goyal and Welch (2003), who examined the performance of the dividend-price ratio relative to the prevailing mean. The lack of significance is not surprising in view of recent results of Brennan and Xia (2004). They show that even if there is true return predictability, available regression techniques might often not recover this for realistic parameter values. Nonetheless, they show that use of the forecasting relation may well be beneficial in a portfolio planning context, even when it appears statistically insignificant. In spite of the insignificant Diebold-Mariano tests, it is thus of interest to investigate further the relative performance of the alternative forecasts. In fact, the R^2 results in Table 7 summarize the average performance over the period, but does not show the evolution over time of the relative performance of the alternative forecasts. Following Goyal and Welch (2003), the dynamics may be captured by graphically depicting the relative performance of each measure, vis-a-vis the prevailing mean. This is done in Figure 5.

What is shown in the Figure 5 is the cumulated net squared out-of-sample forecast error through the period indicated on the first axis. Precisely, this is $\sum_{t=1911}^{T}(SE_t(\text{P}){-}SE_t(\text{M})$, where $SE_t(\text{P})$ is the squared forecast error in period t when using the prevailing mean as forecast, $SE_t(\text{M})$ is the squared forecast error when using the forecast based on the model in question (RUAP or dividend-price ratio), and T indicates the period on the first axis. For the model to do better than the prevailing mean through T, the line must be above the zero level. For long-term forecasts, $K = 4$, the figure shows that all candidate forecasts generally do better than the prevailing mean, and that the line corresponding to the RUAP-based forecast lies above those for the dividend-price ratio and its logarithm. Thus, the latent state variable y_t produced by our model outperforms the alternative measures dynamically, and not only in terms of the overall R^2. This reinforces the impression that the RUAP-based forecast is in fact the best of those considered here.

Figure 5 also shows results for short-term forecasts, $K = 1$. This was the case considered in Goyal and Welch (2003), so our figure is comparable to theirs as far as the forecast based on d_t/p_t goes, and for the period common to the studies, i.e., from $T = 1946$ and on. In fact, our d_t/p_t measure based on the Shiller data is most comparable to the CRSP based measure they label dividend yield, and our figure confirms the dip in performance relative to the prevailing mean which they report for the period from the mid-fifties to the early seventies. Consistent with the R^2 values from Table 7, the figure also shows that the superiority of the RUAP forecast does not extend to the short-term forecasts, $K = 1$.

Overall, the analysis of true out-of-sample forecast errors confirms that the sequence of latent variables y_t backed out from the RUAP model provides better forecasts of long horizon stock returns than the dividend-price ratio and its logarithm.

6. CONCLUSION

In this paper, we have introduced a structural asset pricing model, the RUAP model, with random utility shocks that simultaneously serve to capture movements in the marginal rate of intertemporal substitution (MRIS) and allow estimation by full-information maximum likelihood. Rather than imposing an outside assumption on the joint distribution of dividends, prices, and utility shocks, we impose assumptions on dividends and utility shocks and derive the resulting endogenous distribution of asset prices. Our empirical application shows that the model fits the data reasonably well. The estimated version of the model produces a time series of predicted latent utility or MRIS shocks y_t, for which the associated fitted residual series conforms with the distributional assumption. Furthermore, the resulting series y_t is shown to be a more informative forecast of long-horizon asset returns than the classical dividend-price ratio, and variations thereof.

Why would the particular function of prices and dividends given by y_t from the RUAP model forecast future returns better than transformations of the dividend-price ratio? It is possible that by inverting a structural model with respect to the relevant unobservable, we have identified an important portion of the information in current observables. Thus, y_t may be viewed as an approximately sufficient statistic for the history of dividends and asset prices through t. The question remains whether a more complicated structural model, e.g., allowing for different shocks to MRIS and the rate of relative risk aversion (unity in our model), would allow backing out an even more informative latent variable than y_t. We leave this question for future research.

TABLE 1

STRUCTURAL PARAMETER ESTIMATES FOR RUAP MODEL, 1871-2000

β	a_{dd}	a_{yy}	a_{dy}	a_{yd}	σ_{dd}	σ_{yy}	σ_{dy}	ℓ
0.955 (0.003)	0.778 (0.062)	0.797 (0.039)	−0.089 (0.033)	0.018 (0.089)	0.111 (0.007)	0.181 (0.018)	0.205 (0.104)	-247.710
0.956 (0.003)	0.801 (0.053)	0.797 (0.035)	0	−0.055 (0.087)	0.114 (0.007)	0.184 (0.020)	0.130 (0.106)	-251.523
0.955 (0.003)	0.772 (0.052)	0.793 (0.034)	−0.088 (0.033)	0	0.111 (0.007)	0.181 (0.018)	0.193 (0.086)	-247.730
0.955 (0.002)	0.758 (0.053)	0.782 (0.036)	−0.074 (0.033)	−0.074 (0.076)	0.111 (0.007)	0.182 (0.017)	0	-249.650
0.956 (0.003)	0.811 (0.052)	0.804 (0.033)	0	0	0.114 (0.007)	0.186 (0.021)	0.169 (0.087)	-251.715
0.956 (0.003)	0.802 (0.052)	0.787 (0.035)	0	−0.111 (0.071)	0.114 (0.007)	0.181 (0.019)	0	-252.279
0.955 (0.002)	0.763 (0.053)	0.793 (0.034)	−0.082 (0.032)	0	0.111 (0.007)	0.181 (0.017)	0	-250.136
0.956 (0.003)	0.801 (0.052)	0.792 (0.034)	0	0	0.114 (0.007)	0.184 (0.019)	0	-253.518

NOTE: Asymptotic standard errors in parentheses.

TABLE 2

STRUCTURAL PARAMETER ESTIMATES FOR RUAP MODEL, 1871-1994

β	a_{dd}	a_{yy}	a_{dy}	a_{yd}	σ_{dd}	σ_{yy}	σ_{dy}	ℓ
0.954 (0.002)	0.762 (0.059)	0.763 (0.046)	−0.108 (0.039)	−0.013 (0.083)	0.113 (0.007)	0.171 (0.015)	0.173 (0.100)	-226.081
0.954 (0.003)	0.799 (0.054)	0.772 (0.043)	0	−0.060 (0.085)	0.116 (0.007)	0.172 (0.017)	0.123 (0.106)	-230.140
0.954 (0.002)	0.767 (0.053)	0.766 (0.042)	−0.109 (0.038)	0	0.113 (0.007)	0.171 (0.015)	0.181 (0.088)	-226.093
0.954 (0.002)	0.751 (0.054)	0.752 (0.044)	−0.097 (0.038)	−0.078 (0.075)	0.113 (0.007)	0.171 (0.015)	0	-227.550
0.954 (0.003)	0.810 (0.053)	0.781 (0.040)	0	0	0.116 (0.007)	0.175 (0.018)	0.166 (0.090)	-230.387
0.954 (0.002)	0.802 (0.053)	0.759 (0.044)	0	−0.109 (0.071)	0.116 (0.007)	0.170 (0.016)	0	-230.820
0.954 (0.002)	0.757 (0.054)	0.766 (0.042)	−0.104 (0.038)	0	0.113 (0.007)	0.171 (0.015)	0	-228.107
0.954 (0.003)	0.800 (0.054)	0.766 (0.042)	0	0	0.116 (0.007)	0.173 (0.017)	0	-232.020

NOTE: Asymptotic standard errors in parentheses.

TABLE 3

LONG-HORIZON RETURN REGRESSIONS, 1871-2000

	$r = \beta\, y + e$		$r = \beta\ \ d/p + e$		$r = \gamma\, y + \delta\ \ d/p + e$		
K	β_K	R^2	β_K	R^2	γ_K	δ_K	R^2
			Panel A: Full Model				
1	0.119	0.028	1.743	0.017	1.513	−27.899	0.083
	(1.898)		(1.579)		(3.377)	(−2.960)	
2	0.243	0.054	3.725	0.037	2.631	−47.893	0.133
	(2.115)		(1.829)		(3.308)	(−2.854)	
3	0.327	0.071	5.324	0.056	3.240	−58.572	0.160
	(2.082)		(1.878)		(2.986)	(−2.559)	
4	0.471	0.109	7.849	0.089	4.045	−72.025	0.210
	(2.442)		(2.235)		(2.919)	(−2.472)	
			Panel B: Restricted Model, $a_{yd} = \sigma_{dy} = 0$				
1	0.111	0.025	1.743	0.017	0.952	−17.024	0.047
	(1.792)		(1.579)		(2.181)	(−1.833)	
2	0.234	0.051	3.725	0.037	1.989	−35.615	0.099
	(2.062)		(1.829)		(2.725)	(−2.258)	
3	0.317	0.069	5.324	0.056	2.593	−46.324	0.129
	(2.038)		(1.878)		(2.566)	(−2.130)	
4	0.456	0.105	7.849	0.089	3.180	−55.597	0.170
	(2.380)		(2.235)		(2.331)	(−1.910)	

	$r = \beta\, y + e$		$r = \beta \log(d/p) + e$		$r = \gamma\, y + \delta \log(d/p) + e$		
K	β_K	R^2	β_K	R^2	γ_K	δ_K	R^2
			Panel C: Full Model				
1	0.119	0.028	0.064	0.012	0.703	−0.515	0.073
	(1.898)		(1.207)		(3.932)	(−3.153)	
2	0.243	0.054	0.148	0.028	1.379	−1.030	0.123
	(2.115)		(1.410)		(3.354)	(−2.526)	
3	0.327	0.071	0.221	0.045	1.828	−1.387	0.152
	(2.082)		(1.543)		(3.039)	(−2.290)	
4	0.471	0.109	0.338	0.076	2.327	−1.737	0.197
	(2.442)		(1.905)		(2.763)	(−2.076)	
			Panel D: Restricted Model, $a_{yd} = \sigma_{dy} = 0$				
1	0.111	0.025	0.064	0.012	0.552	−0.394	0.053
	(1.792)		(1.207)		(3.160)	(−2.388)	
2	0.234	0.051	0.148	0.028	1.153	−0.845	0.102
	(2.062)		(1.410)		(3.036)	(−2.153)	
3	0.317	0.069	0.221	0.045	1.556	−1.162	0.130
	(2.038)		(1.543)		(2.827)	(−2.024)	
4	0.456	0.105	0.338	0.076	1.953	−1.421	0.169
	(2.380)		(1.905)		(2.566)	(−1.814)	

NOTE: Asymptotic t-statistics based on Newey and West (1987) standard errors in parentheses.

TABLE 4

LONG-HORIZON RETURN REGRESSIONS, 1871-1994

	$r = \beta\, y + e$		$r = \beta\ \ d/p + e$		$r = \gamma\, y + \delta\ \ d/p + e$		
K	β_K	R^2	β_K	R^2	γ_K	δ_K	R^2
				Panel A: Full Model			
1	0.195	0.065	3.061	0.046	1.285	−22.169	0.096
	(3.194)		(2.753)		(2.599)	(−2.104)	
2	0.365	0.111	5.892	0.083	2.253	−38.443	0.157
	(3.598)		(3.157)		(2.678)	(−2.153)	
3	0.467	0.141	7.946	0.116	2.675	−45.040	0.188
	(3.425)		(3.129)		(2.406)	(−1.913)	
4	0.624	0.188	10.802	0.161	3.359	−55.866	0.243
	(3.736)		(3.498)		(2.354)	(−1.857)	
				Panel B: Restricted Model, $a_{yd} = \sigma_{dy} = 0$			
1	0.185	0.060	3.061	0.046	0.716	−10.923	0.068
	(3.056)		(2.753)		(1.502)	(−1.062)	
2	0.353	0.107	5.892	0.083	1.591	−25.512	0.128
	(3.509)		(3.157)		(2.092)	(−1.544)	
3	0.454	0.136	7.946	0.116	2.014	−32.201	0.162
	(3.341)		(3.129)		(2.017)	(−1.494)	
4	0.607	0.182	10.802	0.161	2.493	−39.019	0.211
	(3.626)		(3.498)		(1.864)	(−1.364)	

	$r = \beta\, y + e$		$r = \beta \log{(d/p)} + e$		$r = \gamma\, y + \delta \log{(d/p)} + e$		
K	β_K	R^2	β_K	R^2	γ_K	δ_K	R^2
				Panel C: Full Model			
1	0.195	0.065	0.153	0.051	0.380	−0.177	0.068
	(3.194)		(2.845)		(1.450)	(−0.651)	
2	0.365	0.111	0.286	0.087	0.864	−0.481	0.121
	(3.598)		(3.080)		(1.737)	(−0.909)	
3	0.467	0.141	0.377	0.116	1.240	−0.745	0.159
	(3.425)		(3.084)		(2.095)	(−1.201)	
4	0.624	0.188	0.507	0.156	1.696	−1.037	0.215
	(3.736)		(3.408)		(2.138)	(−1.286)	
				Panel D: Restricted Model, $a_{yd} = \sigma_{dy} = 0$			
1	0.185	0.060	0.153	0.051	0.203	−0.017	0.060
	(3.056)		(2.845)		(0.830)	(−0.068)	
2	0.353	0.107	0.286	0.087	0.641	−0.281	0.110
	(3.509)		(3.080)		(1.432)	(−0.577)	
3	0.454	0.136	0.377	0.116	0.999	−0.533	0.146
	(3.341)		(3.084)		(1.893)	(−0.940)	
4	0.607	0.182	0.507	0.156	1.372	−0.750	0.197
	(3.626)		(3.408)		(1.976)	(−1.037)	

NOTE: Asymptotic t-statistics based on Newey and West (1987) standard errors in parentheses.

TABLE 5

STRUCTURAL PARAMETER ESTIMATES USING EARNINGS DATA, 1871-2000

β	a_{dd}	a_{yy}	a_{dy}	a_{yd}	σ_{dd}	σ_{yy}	σ_{dy}	ℓ
0.927 (0.004)	0.755 (0.071)	0.681 (0.048)	-0.077 (0.067)	-0.134 (0.098)	0.209 (0.013)	0.239 (0.022)	0.529 (0.084)	-281.142
0.928 (0.004)	0.734 (0.061)	0.726 (0.031)	0	-0.202 (0.066)	0.210 (0.013)	0.235 (0.022)	0.479 (0.073)	-281.870
0.927 (0.004)	0.851 (0.052)	0.651 (0.041)	-0.146 (0.051)	0	0.213 (0.014)	0.244 (0.022)	0.610 (0.058)	-281.851
0.945 (0.015)	0.973 (0.089)	0.995 (0.052)	0.562 (0.699)	0.012 (0.015)	0.224 (0.017)	0.027 (0.038)	0	-284.770
0.927 (0.004)	0.807 (0.052)	0.700 (0.032)	0	0	0.211 (0.013)	0.251 (0.023)	0.561 (0.062)	-286.540
0.929 (0.004)	0.900 (0.054)	0.741 (0.032)	0	-0.297 (0.062)	0.216 (0.014)	0.224 (0.024)	0	-292.591
0.927 (0.004)	0.748 (0.062)	0.670 (0.041)	-0.051 (0.050)	0	0.209 (0.013)	0.245 (0.023)	0	-309.355
0.927 (0.004)	0.732 (0.061)	0.670 (0.041)	0	0	0.210 (0.013)	0.245 (0.023)	0	-309.882

NOTE: Asymptotic standard errors in parentheses.

TABLE 6

LONG-HORIZON RETURN REGRESSIONS USING EARNINGS DATA, 1871-2000

	$r = \beta\, y + e$		$r = \beta\ d/p + e$		$r = \gamma\, y + \delta\ d/p + e$		
K	β_K	R^2	β_K	R^2	γ_K	δ_K	R^2
			Panel A: Full Model				
1	0.106 (2.659)	0.050	1.726 (3.188)	0.065	0.038 (0.289)	1.004 (0.541)	0.051
2	0.197 (2.825)	0.080	2.652 (2.418)	0.073	0.405 (1.633)	−3.043 (−0.832)	0.087
3	0.248 (2.508)	0.094	3.324 (2.146)	0.085	0.466 (1.700)	−3.185 (−0.776)	0.100
4	0.329 (2.780)	0.125	4.330 (2.384)	0.110	0.554 (1.819)	−3.293 (−0.738)	0.129
			Panel B: Restricted Model, $a_{dy} = 0$				
1	0.106 (2.729)	0.051	1.726 (3.188)	0.065	0.052 (0.389)	0.805 (0.415)	0.052
2	0.195 (2.821)	0.080	2.652 (2.418)	0.073	0.446 (1.640)	−3.702 (−0.914)	0.089
3	0.247 (2.493)	0.095	3.324 (2.146)	0.085	0.535 (1.780)	−4.220 (−0.949)	0.104
4	0.330 (2.791)	0.128	4.330 (2.384)	0.110	0.677 (2.040)	−5.090 (−1.072)	0.138
			Panel C: Restricted Model, $a_{yd} = 0$				
1	0.088 (2.116)	0.035	1.726 (3.188)	0.065	−0.513 (−2.346)	8.757 (2.855)	0.088
2	0.132 (1.461)	0.037	2.652 (2.418)	0.073	−1.133 (−3.135)	18.437 (3.701)	0.150
3	0.165 (1.271)	0.043	3.324 (2.146)	0.085	−1.584 (−3.685)	25.471 (4.334)	0.203
4	0.228 (1.461)	0.062	4.330 (2.384)	0.110	−1.931 (−3.538)	31.416 (4.233)	0.247

NOTE: Asymptotic t-statistics based on Newey and West (1987) standard errors in parentheses.

TABLE 6

(*Continued*)

	$r = \beta\, y + e$		$r = \beta \log(d/p) + e$		$r = \gamma\, y + \delta \log(d/p) + e$		
K	β_K	R^2	β_K	R^2	γ_K	δ_K	R^2
			Panel D: Full Model				
1	0.106 (2.659)	0.050	0.150 (3.085)	0.074	−0.039 (−0.294)	0.177 (1.029)	0.059
2	0.197 (2.825)	0.080	0.230 (2.515)	0.081	0.211 (0.863)	−0.018 (−0.055)	0.080
3	0.248 (2.508)	0.094	0.279 (2.244)	0.088	0.344 (1.459)	−0.119 (−0.370)	0.096
4	0.329 (2.780)	0.125	0.358 (2.470)	0.110	0.466 (1.898)	−0.169 (−0.532)	0.127
			Panel E: Restricted Model, $a_{dy} = 0$				
1	0.106 (2.729)	0.051	0.150 (3.085)	0.074	−0.031 (−0.238)	0.168 (0.972)	0.059
2	0.195 (2.821)	0.080	0.230 (2.515)	0.081	0.215 (0.877)	−0.025 (−0.075)	0.080
3	0.247 (2.493)	0.095	0.279 (2.244)	0.088	0.366 (1.506)	−0.147 (−0.448)	0.098
4	0.330 (2.791)	0.128	0.358 (2.470)	0.110	0.523 (2.011)	−0.239 (−0.716)	0.133
			Panel F: Restricted Model, $a_{yd} = 0$				
1	0.088 (2.116)	0.035	0.150 (3.085)	0.074	−0.309 (−1.937)	0.485 (2.378)	0.089
2	0.132 (1.461)	0.037	0.230 (2.515)	0.081	−0.636 (−2.166)	0.949 (2.742)	0.131
3	0.165 (1.271)	0.043	0.279 (2.244)	0.088	−0.720 (−1.877)	1.093 (2.646)	0.135
4	0.228 (1.461)	0.062	0.358 (2.470)	0.110	−0.796 (−1.752)	1.265 (2.623)	0.156

NOTE: Asymptotic t-statistics based on Newey and West (1987) standard errors in parentheses.

TABLE 7

MEAN SQUARED ERROR R^2 FOR OUT-OF-SAMPLE FORCASTS

K	Prevailing Mean	y	d/p	$\log(d/p)$
		Panel A: 1871-2000		
1	0.106	0.063	0.073	0.066
2	0.189	0.171	0.172	0.160
3	0.276	0.283	0.275	0.264
4	0.332	0.359	0.346	0.335
		Panel B: 1871-1994		
1	0.078	0.077	0.082	0.089
2	0.150	0.194	0.191	0.194
3	0.236	0.310	0.297	0.296
4	0.291	0.390	0.375	0.372
		Panel C: Earnings Data, 1871-2000		
1	0.222	0.205	0.205	0.227
2	0.348	0.367	0.340	0.356
3	0.461	0.491	0.483	0.483
4	0.528	0.568	0.561	0.561

NOTE: Reported statistics are mean squared out-of-sample forecast error R^2, given by $R^2 = 1 - MSFE/s^2$, where $MSFE$ is the mean squared forecast error and s^2 is the variance of the long-term return being forecast.

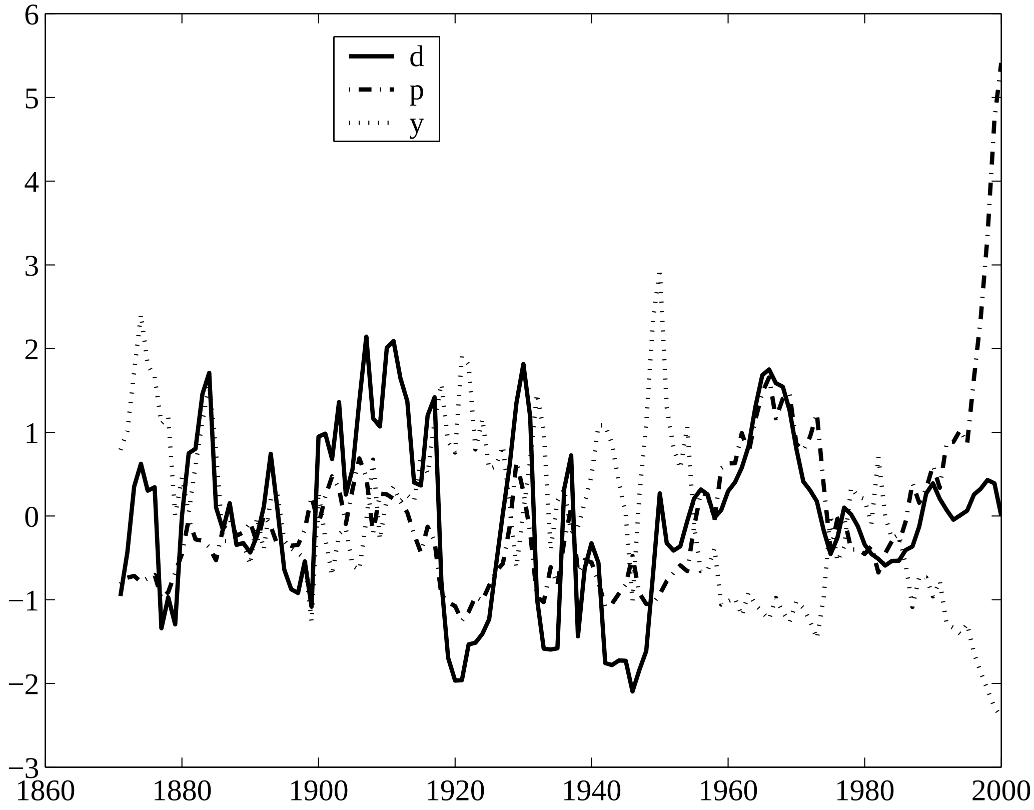


FIGURE 1: Time series for d_t, p_t and y_t
Note: Full sample, no restrictions on parameters in y_t. Standardized variables.

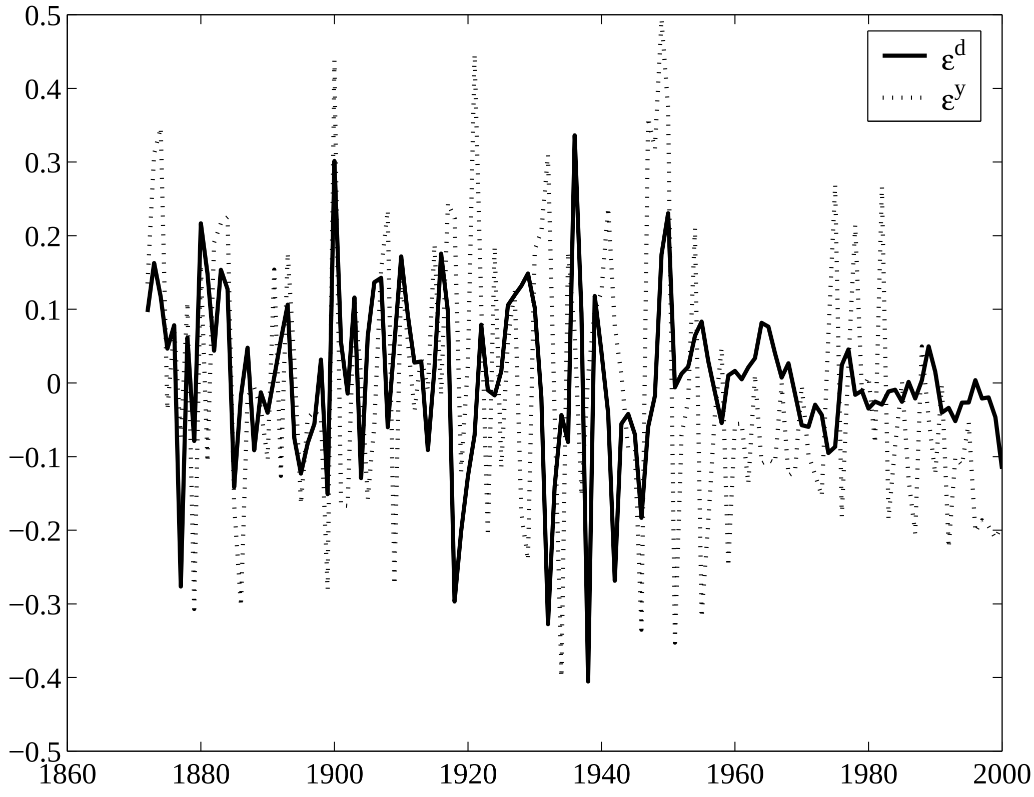
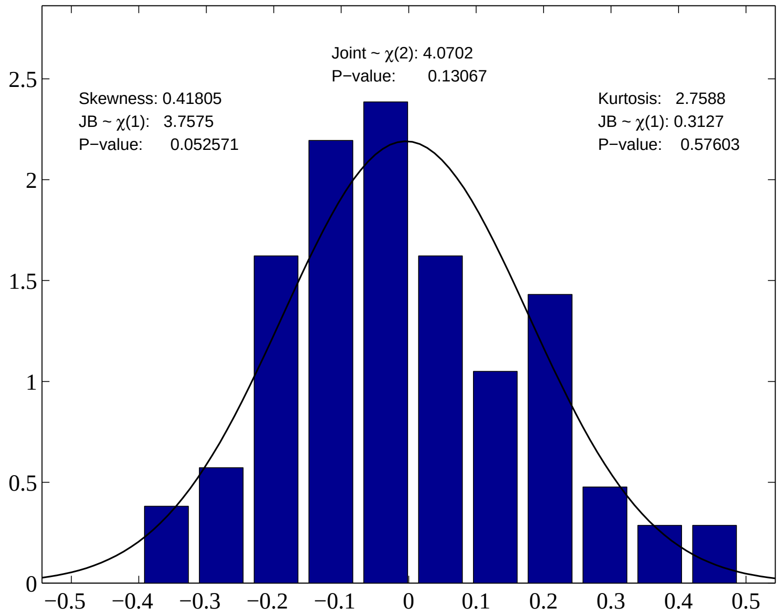
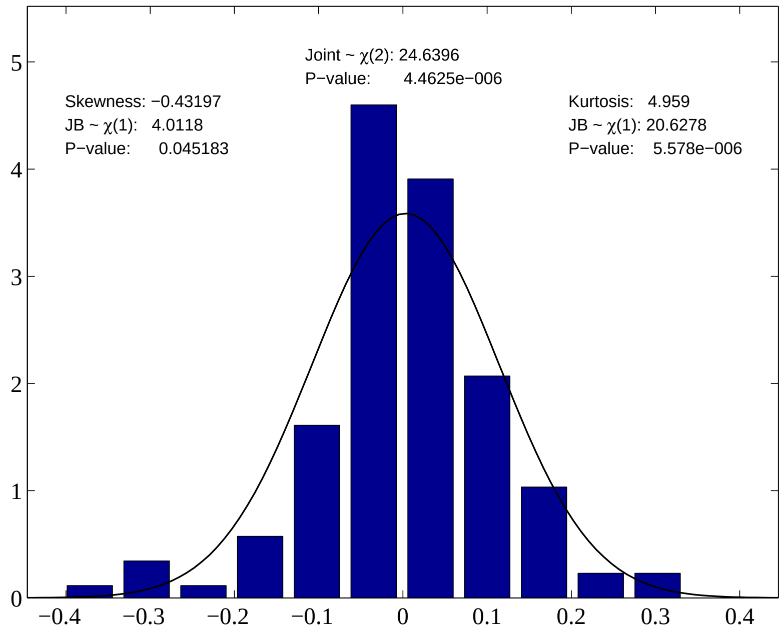


FIGURE 2: The error terms
Note: Full sample, no restrictions on parameters

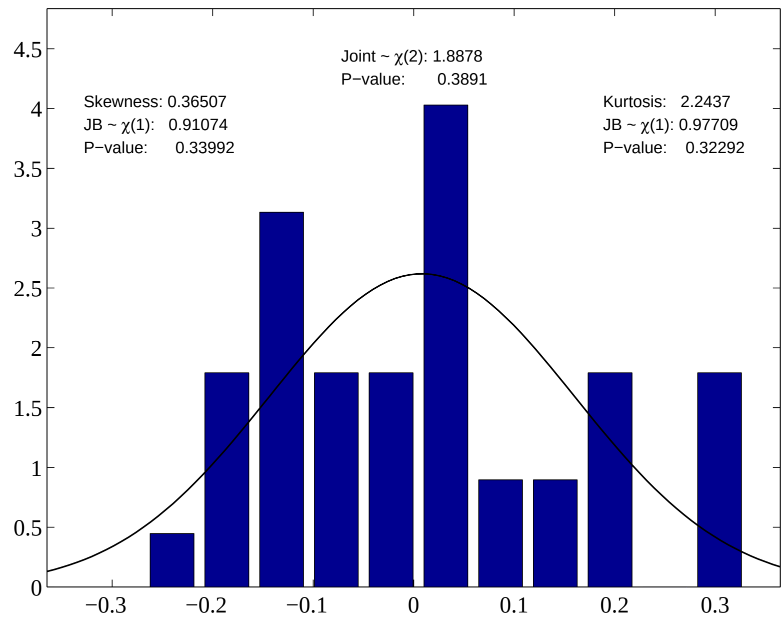


(a) Distribution of ϵ_t^y

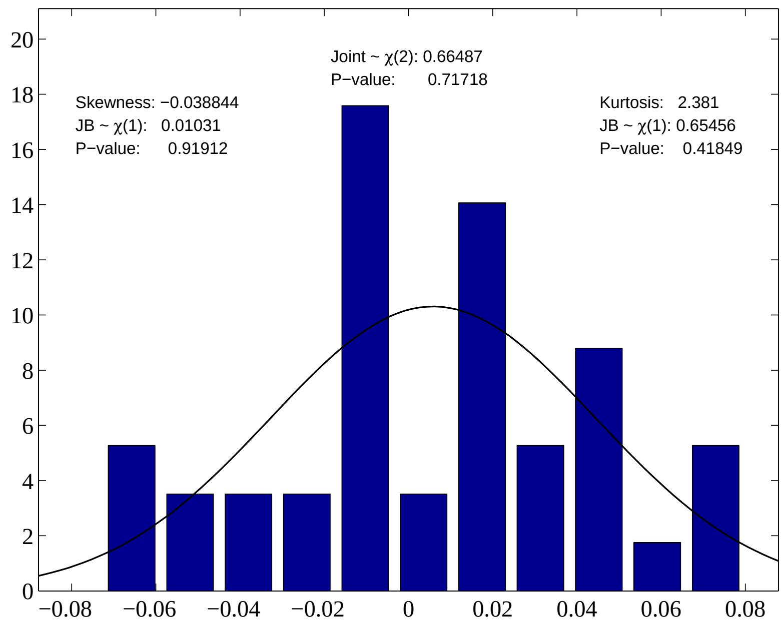


(b) Distribution of ϵ_t^d

FIGURE 3: Distribution of error terms ϵ_t
Error terms ϵ_t defined as in (1). Full sample. No restrictions on parameters. Values of Jarque-Bera tests indicated.



(a) Distribution of ϵ_t^y



(b) Distribution of ϵ_t^d

FIGURE 4: Distribution of error terms ϵ_t

Error terms ϵ_t defined as in (1). Period: 1959-2000. No restrictions on parameters. Values of Jarque-Bera tests indicated.

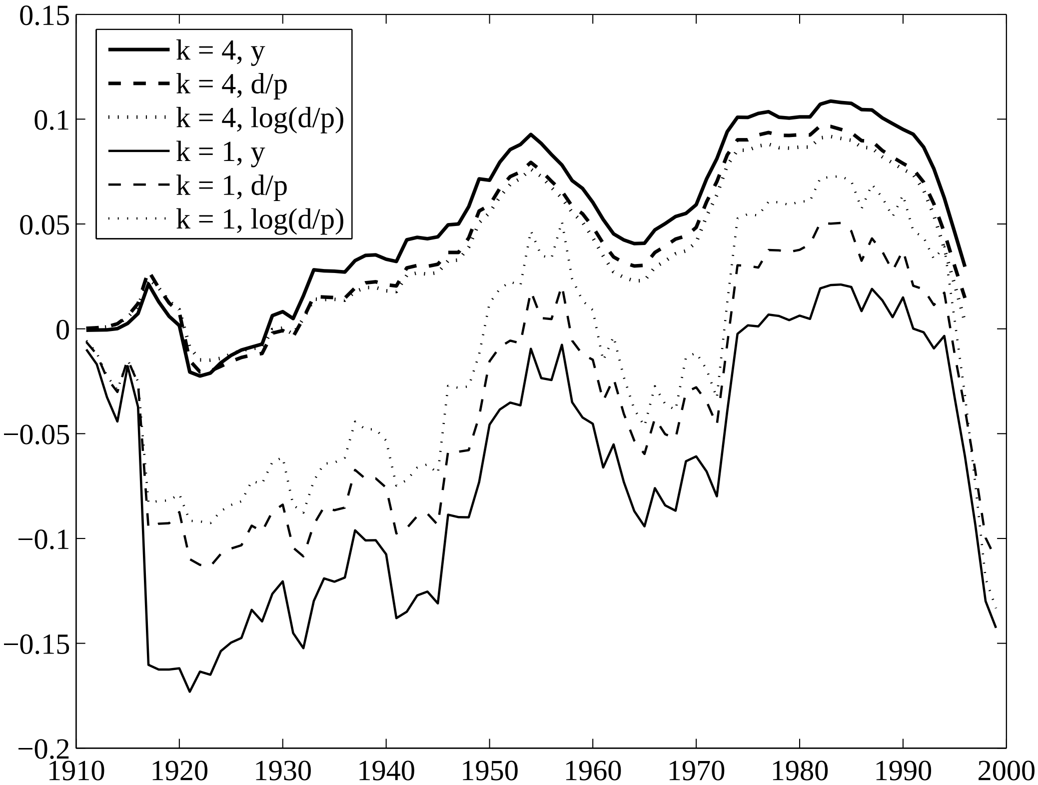


FIGURE 5: Cumulative relative out-of-sample performance.

Note: The figure plots the cumulated net squared out-of-sample forecast error, given by $\sum_{t=1911}^{T}(SE_t(\text{P}) - SE_t(\text{M}))$, where $SE_t(\text{P})$ is the squared forecast error in period t when using the prevailing mean as forecast, and $SE_t(\text{M})$ is the squared forecast error when using the model in question.

REFERENCES

AIYAGARI, S. R. AND M. GERTLER (1991): "Asset Returns with Transaction Costs and Uninsured Individual Risk," *Journal of Monetary Economics*, 27(3):311–331.

BREEDEN, D. T. (1979): "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities," *Journal of Financial Economics*, 7(3):265–296.

BRENNAN, M. J. AND Y. XIA (2004): "Persistence, Predictability, and Portfolio Planning," Manuscript, Finance Department, Wharton School, University of Pennsylvania.

CAMPBELL, J. Y. (2000): "Asset Pricing at the Millennium," *Journal of Finance*, 55(4):1515–1567.

CAMPBELL, J. Y., A. W. LO, AND A. C. MACKINLAY (1997): *The Econometrics of Financial Markets,* Princeton University Press, Princeton, New Jersey.

CAMPBELL, J. Y. AND R. J. SHILLER (1988a): "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1:195–227.

——— (1988b): "Stock Prices, Earnings, and Expected Dividends," *Journal of Finance*, 43:661–676.

COLEMAN, T. AND Y. LI (1996): "An Interior, Trust Region Approach for Nonlinear Minimization Subject to Bounds," *SIAM Journal on Optimization*, 6:418–445.

CONSTANTINIDES, G. M. (1990): "Habit Formation: A Resolution of the Equity Premium Puzzle," *Journal of Political Economy*, 98(3):519–543.

COWLES, ALFRED, III AND ASSOCIATES (1939): *Common Stock Indexes*, volume 3 of *Cowles Commission Monograph*. Principia Press, Bloomington, Indiana, 2ed edition.

DIEBOLD, F. X. AND R. S. MARIANO (1995): "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, 13(3):253–263.

EPSTEIN, L. G. AND S. E. ZIN (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57(4):937–969.

FAMA, E. F. AND K. R. FRENCH (1988): "Dividend Yields and Expected Stock Returns," *Journal of Financial Economics*, 22(1):3–27.

GALLANT, A. R. (1987): *Nonlinear Statistical Models,* John Wiley & Sons.

GOYAL, A. AND I. WELCH (2003): "Predicting the Equity Premium with Dividend Ratios," *Management Science*, 49(5):639–654.

HANSEN, L. P. AND K. J. SINGLETON (1983): "Stochastic Consumption, Risk Aversion, and the Temporal Behaviour of Asset Returns," *Journal of Political Economy*, 91(2):249–265.

KOCHERLAKOTA, N. R. (1996): "The Equity Premium: It's Still a Puzzle," *Journal of Economic Literature*, 34(1):42–71.

KREPS, D. M. AND E. L. PORTEUS (1978): "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 46(1):185–200.

LEROY, S. F. AND R. PORTER (1981): "The Present Value Relation: Tests Based on Variance Bounds," *Econometrica*, 49:555–577.

LUCAS, JR., R. E. (1978): "Asset Prices in an Exchange Economy," *Econometrica*, 46(6):1429–1445.

MCFADDEN, D. (1973): "Conditional Logit Analysis of Qualitative Choice Behavior," in Zarembka, P., editor, *Frontiers of Econometrics*. New York: Academic Press.

MEHRA, R. AND E. C. PRESCOTT (1985): "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15:145–161.

NEWEY, W. K. AND K. D. WEST (1987): "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55(3):703–708.

RUST, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55(5):999–1033.

SHILLER, R. J. (1981): "Do Stock Pries Move Too Much to be Justified by Subsequent Changes in Dividends," *American Economic Review*, 71(3):421–436.

INSTITUT FOR FINANSIERINGs working papers siden 2000:

WP 2000-1	Bjarne Astrup Jensen and Carsten Sørensen: Paying for minimum interest rate guarantees: Who should compensate who?
WP 2000-2	Jan Jakobsen and Ole Sørensen: Decomposing and testing Long-run Returns with an application to initial public offerings in Denmark.
WP 2000-3	Jan Jakobsen and Torben Voetmann: Volatility-Adjusted Performance An Alternative Approach to Interpret Long-Run Returns.
WP 2000-4	Jan Jakobsen and Torben Voetmann: Post-Acquisition Performance in the Short and Long-Run Evidence from the Copenhagen Stock Exchange 1993-1997.
WP 2000-5	Ken L. Bechmann and Johannes Raaballe: A Regulation of Bids for Dual Class Shares. Implication: Two Shares – One Price.
WP 2000-6	Torben Voetmann: Changes in the Bid-Ask Componets Around Earnings Announcemtents: Evidence from the Copenhagen Stock Exchange.
WP 2000-7	Henrik Lando: The Optimal Standard of Proof in Criminal Law When Both Fairness and Deterrence Are Social Aims.
WP 2000-8	Jesper Rangvid and Carsten Sørensen: Convergence in the ERM and Declining Numbers of Common Stochastic Trends.
WP 2000-9	Claus Munk and Carsten Sørensen: Optimal Consumption and Investment Strategies with Stochastic Interest Rates.
WP 2000-10	Henrik Lando and Caspar Rose: On Specific Performance in Civil Law and Enforcement Costs.
WP 2000-11	Henrik Lando: Ny lov om jordforurening i økonomisk belysning.

WP 2001-1	Michael Møller, Claus Parum og Thomas Sørensen: Den ny pensionsafkastbeskatningslov.
WP 2001-2	Bjarne Astrup Jensen: Mean variance efficient portfolios by linear programming: A review of some portfolio selection criteria of Elton, Gruber and Padberg.
WP 2001-3	Caspar Rose: Impact of Investor Meetings/Presentations on Share

Prices, Insider Trading and Securities Regulation.

WP 2001-4	Caspar Rose: Corporate Financial Performance and the Use of Takeover Defenses.
WP 2001-5	Shubhashis Gangopadhyay and Clas Wihlborg: The Impact of Bankruptcy Rules on Risky Project Choice and Skill Formation under Credit Rationing.
WP 2001-6	Claus Munk, Carsten Sørensen & Tina Nygaard Vinther: Portfolio Choice under Inflation: Are Popular Recommendations Consistent with Rational Behaviour?
WP 2001-7	Ken L. Bechmann: Evidence on the Limits of Arbitrage: Short Sales, Price Pressure, and the Stock Price Response to Convertible Bond Calls.
WP 2001-8	Michael Møller & Caspar Rose: Legal pre-emption rights as call-options, redistribution and efficiency loss.
WP 2001-9	Peter Raahauge: Empirical Rationality in the Stock Market.

WP 2002-1	Bjarne Astrup Jensen: On valuation before and after tax in no arbitrage models: Tax neutrality in the discrete time model.
WP 2002-2	Ken L. Bechmann: Price and Volume Effects Associated with Changes in the Danish Blue-Chip Index – The KFX Index.
WP 2002-3	Steen Thomsen and Caspar Rose: Foundation ownership and financial performance. Do companies need owners?
WP 2002-4	Martin Richter and Carsten Sørensen: Stochastic Volatility and Seasonality in Commodity Futures and Options: The Case of Soybeans.
WP 2002-5	Caspar Rose: Aktiemarkedets reaktion på indførelsen af incitaments-programmer.
WP 2002-6	Caspar Rose: Impact of Takeover Defenses on Managerial Incentives.
WP 2002-7	Ken L. Bechmann og Peter Løchte Jørgensen: Optionsaflønning i danske børsnoterede selskaber.
WP 2002-8	Jesper Rangvid: Output and Expected Returns – a multicountry study.
WP 2002-9	Jonas Aziz Bhatti og Michael Møller: Pensionsafkastbeskatning og optimal porteføljesammensætning.

WP 2003-1 Bjarne Florentsen, Michael Møller and Niels Chr. Nielsen: Reimbursement of VAT on written-off Receivables.

WP 2003-2: Ken L. Bechmann and Peter Løchte Jørgensen: The Value and Incentives of Option-based Compensation in Danish Listed Companies.

WP 2004-1 Ken L. Bechmann and Johannes Raaballe: The Difference Between Stock Splits and Stock Dividends – Evidence from Denmark.

WP 2004-2 Caspar Rose: Bestyrelsessammensætning og finansiel performance i danske børsnoterede virksomheder - Er Nørbyrapportens anbefalinger til gavn for aktionærerne?

WP 2004-3 Jens Lunde: Lack of balance in after-tax returns - lack of tenure neutrality. The Danish case.

WP 2004-4 Peter Raahauge: Upper Bounds on Numerical Approximation Errors.

WP 2004-5 Peter Raahauge: Higher-Order Finite Element Solutions of Option Prices.

WP 2004-6 Anders Holst & Morten Nalholm: Modelling Callable Annuity Bonds with Interest-Only Optionality.

WP 2004-7 Bent Jesper Christensen & Peter Raahauge: Latent Utility Shocks in a Structural Empirical Asset Pricing Model.