Solving and Analyzing a Model with Two Lucas Trees

This note explores an analysis of an economy with two Lucas trees, which is studied in Cochrane, Longstaff and Santa-Clara ('Two Trees', The Review of Financial Studies, vol. 21 no. 1, 2008) (CLS) and Ian Martin, 'The Lucas Orchard,' Econometrica, January 2013 (see especially Figure 3). The model is particularly well suited to the study of projection and dynamic programming as methods for solving a dynamic model.

1 Model

Consider an economy with two trees, tree number 1 and tree number 2. Corresponding to these two trees there are the following two dividend processes,

$$D_{1t}, D_{2t}.$$

The time series representations are:

$$\frac{D_{1,t+1}}{D_{1,t}} = \varepsilon_{1,t+1}, \ \frac{D_{2,t+1}}{D_{2,t}} = \varepsilon_{2,t+1},$$

where the two shocks are iid over time and independent of each other.

The representative household takes $z_{1,0}$ and $z_{2,0}$ as given and chooses $\{z_{1,t}, z_{2,t}\}$ to maximize discounted utility,

$$E_0 \sum_{t=0}^{\infty} \beta^t u(C_t), \ u(C) = \frac{C^{1-\gamma}}{1-\gamma},$$

subject to

$$C_t + p_{1,t}(z_{1,t+1} - z_{1,t}) + p_{2,t}(z_{2,t+1} - z_{2,t}) \leq z_{1,t}D_{1,t} + z_{2,t}D_{2,t},$$

where $z_{i,t}$ denotes the stock of tree i owned at time t, $i = 1, 2$.

2 Equilibrium Conditions

In equilibrium, the quantity of trees purchased must be equal to the outstanding stock, so that

$$z_{1,t} = \alpha, \ z_{2,t} = 1 - \alpha, \text{ for all } t$$

and goods market clearing requires,

$$C_t = \alpha D_{1,t} + (1-\alpha) D_{2,t}.$$

Maximization by household leads to:

$$p_{1,t} = \beta E \left(\frac{C_t}{C_{t+1}}\right)^{\gamma} \left[D_{1,t+1} + p_{1,t+1}\right].$$

It is convenient to rewrite the previous expression. Thus, consider the (weighted) price consumption ratio:

$$\begin{aligned} P_{1,t} &\equiv \frac{\alpha p_{1,t}}{C_t}, \\ P_{2,t} &\equiv \frac{(1-\alpha) p_{2,t}}{C_t}. \end{aligned}$$

Then,

$$\begin{aligned} P_{1,t} &= \beta E \left(\frac{C_t}{C_{t+1}}\right)^{\gamma} \left[\frac{\alpha D_{1,t+1}}{C_{t+1}} + P_{1,t+1}\right] \frac{C_{t+1}}{C_t} \\ &= \beta E \left(\frac{C_{t+1}}{C_t}\right)^{1-\gamma} \left[x_{t+1} + P_{1,t+1}\right], \end{aligned}$$

where x_t denotes the share in consumption of earnings from tree 1 :

$$x_t \equiv \frac{\alpha D_{1,t}}{\alpha D_{1,t} + (1-\alpha) D_{2,t}}.$$

Also,

$$\begin{aligned} P_{2,t} &= \beta E \left(\frac{C_t}{C_{t+1}}\right)^{\gamma} \left[\frac{(1-\alpha) D_{2,t+1}}{C_{t+1}} + P_{2,t+1}\right] \frac{C_{t+1}}{C_t} \\ &= \beta E \left(\frac{C_{t+1}}{C_t}\right)^{1-\gamma} \left[1 - x_{t+1} + P_{2,t+1}\right], \end{aligned}$$

Note,

$$\begin{aligned} \frac{C_{t+1}}{C_t} &= \frac{\alpha D_{1,t+1} + (1-\alpha) D_{2,t+1}}{\alpha D_{1,t} + (1-\alpha) D_{2,t}} \\ &= \frac{\alpha \varepsilon_{1,t+1} D_{1,t} + (1-\alpha) \varepsilon_{2,t+1} D_{2,t}}{\alpha D_{1,t} + (1-\alpha) D_{2,t}} \\ &= \frac{\alpha D_{1,t}}{\alpha D_{1,t} + (1-\alpha) D_{2,t}} \varepsilon_{1,t+1} + \frac{(1-\alpha) D_{2,t}}{\alpha D_{1,t} + (1-\alpha) D_{2,t}} \varepsilon_{2,t+1} \\ &= x_t \varepsilon_{1,t+1} + (1-x_t) \varepsilon_{2,t+1}. \end{aligned}$$

It is convenient to derive an expression for x_{t+1} :

$$\begin{aligned}
x_{t+1} &= \frac{\alpha D_{1,t+1}}{\alpha D_{1,t+1} + (1-\alpha) D_{2,t+1}} \\
&= \frac{\alpha D_{1,t+1}}{\alpha D_{1,t}} \frac{\alpha D_{1,t}}{\alpha D_{1,t} + (1-\alpha) D_{2,t}} \frac{\alpha D_{1,t} + (1-\alpha) D_{2,t}}{\alpha D_{1,t+1} + (1-\alpha) D_{2,t+1}} \\
&= \frac{\varepsilon_{1,t+1} x_t}{x_t \varepsilon_{1,t+1} + (1-x_t) \varepsilon_{2,t+1}}.
\end{aligned}$$

It is interesting to think about what the ergodic distribution of x_t is. Surprisingly, perhaps, the distribution of x_t is not a function of α. Actually nothing of interest to us is a function of α.

Another object that is of interest is the price-dividend ratio:

$$\begin{aligned}
\frac{p_{1,t}}{D_{1,t}} &= \frac{\alpha p_{1,t}}{C_t} \frac{C_t}{\alpha D_{1,t}} = \frac{P_{1,t}}{x_t} \\
\frac{p_{2,t}}{D_{2,t}} &= \frac{(1-\alpha) p_{2,t}}{C_t} \frac{C_t}{(1-\alpha) D_{2,t}} = \frac{P_{2,t}}{1-x_t}.
\end{aligned}$$

Expressions suggest that the price-consumption ratios may be better behaved at the boundaries of x than is the price-dividend ratio.

It is of interest to compute the risk free interest rate, R_t :

$$1 = R_t \beta E_t \left[x_t \varepsilon_{1,t+1} + (1-x_t) \varepsilon_{2,t+1}\right]^{-\gamma}.$$

With no uncertainty,

$$1.0795 = \frac{1.02^{1.8}}{0.96} = \frac{\mu^{\gamma}}{\beta} = R_t.$$

The price-dividend ratio with no uncertainty would be

$$\begin{aligned}
P_1 &= \beta \mu^{1-\gamma} \left[\alpha + P_1\right] \\
P_1 &= \frac{\beta \mu^{1-\gamma} \alpha}{1 - \beta \mu^{1-\gamma}},
\end{aligned}$$

so that the price/dividend ratio is:

$$\frac{p_1}{D_1} = \frac{P_1}{x} = \frac{\beta \mu^{1-\gamma}}{1-\beta \mu^{1-\gamma}} = \frac{0.96 * 1.02\hat{}\,(1-1.8)}{1 - 0.96 * 1.02\hat{}\,(1-1.8)} = 17.$$

3 Shocks

Suppose the net growth rates of the two trees takes on three possible values as follows:

$$\varepsilon_1 - 1, \varepsilon_2 - 1 \in (\mu - \sigma, \mu, \mu + \sigma).$$

Let the 9 states be given by the 9 by 1 vector, s :

$$s = \begin{pmatrix} l,l \\ l,m \\ l,h \\ m,l \\ m,m \\ m,h \\ h,l \\ h,m \\ h,h \end{pmatrix}$$

Let π_1 and π_2 denote the Markov transition matrices for ε_1 and ε_2, respectively. The iid assumption implies that the rows of π_1 are all equal. Similarly for the rows of π_2. Let π denote the Markov transition matrix for s :

$$\begin{aligned} \pi &= \pi_1 \otimes \pi_2 \\ & \begin{bmatrix} \pi_1^{11}\pi_2 & \pi_1^{12}\pi_2 & \pi_1^{13}\pi_2 \\ \pi_1^{21}\pi_2 & \pi_1^{22}\pi_2 & \pi_1^{23}\pi_2 \\ \pi_1^{31}\pi_2 & \pi_1^{32}\pi_2 & \pi_1^{33}\pi_2 \end{bmatrix} \end{aligned}$$

We impose a symmetric structure on the rows of π_i, so that p_i denotes the probability of the low and high state and $1 - 2p_i$ denotes the probability of the middle state, $i = 1, 2$. Then, the standard deviation of the process is:

$$\sqrt{p_i\left[-\sigma\right]^2 + (1 - 2p_i)\, 0^2 + p_i\left[\sigma\right]^2} = \sigma\sqrt{2p_i},$$

for $0 < p_i \leq 1/2$.

4 Symmetry of Equilibrium Price Consumption Ratios

We posit that the solution is a set of functions, $P_1(x)$, $P_2(x)$, $x \in [0,1]$ that satisfy the following fixed point:

$$\begin{aligned} P_1(x) &= \beta E\left[x\varepsilon_1' + (1-x)\varepsilon_2'\right]^{1-\gamma}\left[x' + P_1(x')\right] && (1) \\ P_2(x) &= \beta E\left[x\varepsilon_1' + (1-x)\varepsilon_2'\right]^{1-\gamma}\left[1 - x' + P_2(x')\right], && (2) \end{aligned}$$

where

$$x' = \frac{\varepsilon_1' x}{x\varepsilon_1' + (1-x)\varepsilon_2'} \tag{3}$$

It is easy to see that $P_2(x) = P_1(1-x)$. To see this, note that P_1 satisfies the first of the above two functional equations when the argument, x, is replaced by $1-x$. For this, it is convenient to make x explicit everywhere:

$$P_1(x) = \beta E\left[x\varepsilon_1' + (1-x)\varepsilon_2'\right]^{1-\gamma}\left[\frac{\varepsilon_1' x}{x\varepsilon_1' + (1-x)\varepsilon_2'} + P_1\left(\frac{\varepsilon_1' x}{x\varepsilon_1' + (1-x)\varepsilon_2'}\right)\right]$$

Now, replace x by $1-x$:

$$P_1(1-x) = \beta E\left[(1-x)\varepsilon_1' + x\varepsilon_2'\right]^{1-\gamma}\left[\frac{\varepsilon_1'(1-x)}{(1-x)\varepsilon_1' + x\varepsilon_2'} + P_1\left(\frac{\varepsilon_1'(1-x)}{(1-x)\varepsilon_1' + x\varepsilon_2'}\right)\right]$$

Note that the right side has the form, $Ef(\varepsilon_1', \varepsilon_2')$. Writing this out:

$$Ef(\varepsilon_1', \varepsilon_2') = \sum_{i=1}^{3}\sum_{l=1}^{3}\pi_i^1\pi_l^2 f\left(\varepsilon_i^1, \varepsilon_l^2\right),$$

where π_i^1 and π_l^2 are the distributions of the first and second shock, respectively. Also, ε_l^i, $l = 1,2,3$ are the different possible realizations of the i^{th} shocks, $i = 1,2$. But, in our setting, $\varepsilon_l^1 = \varepsilon_l^2$ for $l = 1,2,3$. In addition, $\pi_l^1 = \pi_l^2$ for $l = 1,2,3$. Thus, ε^1 and ε^2 are the same random variables, so that

$$Ef(\varepsilon_1', \varepsilon_2') = Ef(\varepsilon_2', \varepsilon_1').$$

Using this,

$$\begin{aligned} P_1(1-x) &= \beta E\left[(1-x)\varepsilon_2' + x\varepsilon_1'\right]^{1-\gamma}\left[\frac{\varepsilon_2'(1-x)}{(1-x)\varepsilon_2' + x\varepsilon_1'} + P_1\left(\frac{\varepsilon_2'(1-x)}{(1-x)\varepsilon_2' + x\varepsilon_1'}\right)\right] \\ &= \beta E\left[x\varepsilon_1' + (1-x)\varepsilon_2'\right]^{1-\gamma}\left[1 - x' + P_1(1-x')\right], \end{aligned}$$

where x' is defined by (3). We conclude that $P_1(1-x)$ satisfies the functional equation, (2), so that

$$P_2(x) = P_1(1-x).$$

We conclude that whenever the distributions of ε^1 and ε^2 are identical, then we need only solve for one of the pricing functions.

5 Solving the Model

Our functional equations can be written:

$$\begin{aligned} P_1(x) - \beta \sum_{j=1}^{N} \pi_{ij} \left[x\varepsilon_1(j) + (1-x)\varepsilon_2(j) \right]^{1-\gamma} \left[x'(j) + P_1(x'(j)) \right] &= 0 \\ P_2(x) - \beta \sum_{j=1}^{N} \pi_{ij} \left[x\varepsilon_1(j) + (1-x)\varepsilon_2(j) \right]^{1-\gamma} \left[1 - x'(j) + P_2(x'(j)) \right] &= 0, \end{aligned}$$

for all $0 \leq x \leq 1$, where

$$x'(j) = \frac{\varepsilon_1'(j)\, x}{x\varepsilon_1'(j) + (1-x)\varepsilon_2'(j)} \tag{4}$$

One approach is to construct a Chebyshev polynomial approximation to P_1 and P_2. The domain of these functions is $[0,1]$, but the domain of the Chebyshev polynomial is $[-1,1]$. Thus, we require a mapping,

$$\varphi : [0,1] \to [-1,1],$$

and the following serves our purposes:

$$\varphi(x) = 2x - 1.$$

We approximate P_1 and P_2 with $M-1^{th}$ ordered Chebyshev polynomials, with basis functions, $T_i(\varphi(x))$, for $i = 0, 1, ..., M-1$.[1] In particular, let

$$T(x) = [T_0(\varphi(x)), T_1(\varphi(x)), ..., T_{M-1}(\varphi(x))]'.$$

Let a and b denote two $M \times 1$ vectors of parameters. Then, one strategy for approximating the solutions is:

$$\hat{P}_1(x; a) = a'T(x), \;\; \hat{P}_2(x; b) = b'T(x).$$

[1] The Chebyshev polynomials are defined as follows: $T_0(x) \equiv 1$, $T_1(x) = x$, and $T_i(x) = 2xT_{i-1}(x) - T_{i-2}(x)$, for $i \geq 2$.

The M zeros of the M^{th} order Chebyshev polynomial, T_M, are

$$r_j = \cos\left(\frac{\pi(j-0.5)}{M}\right),\ j = 1, \ldots, M,$$

and let

$$x_j = \varphi^{-1}(r_j) = \frac{r_j + 1}{2},\ j = 1, ..., M. \tag{5}$$

Another possibility is to use a version of the finite element approach to approximating the equilibrium price-consumption functions. We fixed a set of grid points for x and then the parameters, a and b, represent the values of the functions, $\hat{P}_1$ and $\hat{P}_2$, at the grid points. The functions were made continuous by spline interpolation using the MATLAB function, interp1.[2]

Define the error functions, E_1 and E_2 :

$$E_1(x; a) = \hat{P}_1(x; a) - \beta \sum_{j=1}^{N} \pi_{ij} \left[x\varepsilon_1(j) + (1-x)\varepsilon_2(j)\right]^{1-\gamma} \left[x'(j) + \hat{P}_1(x'(j); a)\right] \tag{6}$$

$$E_2(x; b) = \hat{P}_2(x; b) - \beta \sum_{j=1}^{N} \pi_{ij} \left[x\varepsilon_1(j) + (1-x)\varepsilon_2(j)\right]^{1-\gamma} \left[1 - x'(j) + \hat{P}_2(x'(j); b)\right]. \tag{7}$$

Given the grid points for x and the parameters of the parametric functions, a collocation approach solves $2M$ unknowns and $2M$ equations. Solving these using a nonlinear equation solver is a classic Collocation approach. Dynamic programming (or, policy function iteration) is also a convenient way to solve these equations. This involves positing functions, $\hat{P}_1$ and $\hat{P}_2$, on the right of the equality in (6) and (7) and then computing new functions, $\hat{P}_1$ and $\hat{P}_2$, that satisfies the two equalities. The solution we seek is a fixed point of this mapping. At first, dynamic programming may seem computationally inefficient. However, it is easy to see that the computations grow linearly with M while in the case of equation solving by Newton methods the number of computations grow with M^3. Thus, if one wants to work with a large value of M, dynamic programming may be preferred. Another way to work with a large value of M is to do the Galerkin method. This sets M large but then chooses parameters of the policy functions to minimize a relatively small set of linear combinations of the Euler errors. This approach is discussed in the last section below.

From an economic standpoint, it is interesting to study the ratio of consumption to 'wealth'.

[2] In practice, the upper and lower grid points line interior to the unit interval, which is the domain of x. To evaluate the pricing functions at points below the lowest grid point and points above the highest grid point (but, of course, always interior to the set, $[0, 1]$) we included the parameter, 'extrap', in the call to interp1.

We can define wealth as w_t, where

$$w_t = z_{1,t}\left(D_{1,t} + p_{1,t}\right) + z_{2,t}\left(D_{2,t} + p_{2,t}\right).$$

Note:

$$\begin{aligned}\frac{w_t}{C_t} &= \frac{\alpha p_{1,t} + (1-\alpha)\, p_{2,t}}{C_t} + \frac{\alpha D_{1,t} + (1-\alpha)\, D_{2,t}}{C_t} \\ &= P_{1,t} + P_{2,t} + 1.\end{aligned}$$

Write

$$C = f(x)\, w,$$

so that

$$f(x) = \frac{1}{P_1(x) + P_2(x) + 1}.$$

The Friedman idea was that $f(x) = (1-\beta)/\beta$, *i.e.*, that consumption is a fixed proportion of wealth. This is often used as a simple rule of thumb for predicting what will happen to consumption after a stock market collapse (e.g., the collapses in 2000 and 2008). This simple rule of thumb usually does now fare well in these calculations because it predicts that consumption will drop by the same percent that wealth drops and consumption in fact falls by less. The theory here predicts that consumption is a fixed fraction of wealth only when $\gamma = 1$, but not when $\gamma \neq 1$. We find that when $\gamma > 1$ then $f(x)$ has an inverted 'U' shape when graphed against x, with a peak at $x = 1/2$ (actually, we find that the inverted U is perfectly symmetric, a property that is probably easy to prove). If we want to see consumption relatively smooth after a wealth shift, then there must be an appropriate change in x. Given the shape of $f(x)$, this means that x must either rise in the interval $(0, 1/2)$ or fall in the interval $(1/2, 1)$.

We now turn to returns. The ex post realized return on tree 1 is

$$\begin{aligned}R_{1,t+1} &= \frac{p_{1,t+1} + D_{1,t+1}}{p_{1,t}} = \frac{\frac{1}{\alpha}P_{1,t+1}C_{t+1} + \frac{1}{\alpha}C_{t+1}x_{t+1}}{\frac{1}{\alpha}P_{1,t}C_t} \\ &= \frac{P_{1,t+1} + x_{t+1}}{P_{1,t}}\left[x_t\varepsilon_{1,t+1} + (1-x_t)\,\varepsilon_{2,t+1}\right].\end{aligned}$$

Also,

$$\begin{aligned}R_{2,t+1} &= \frac{p_{2,t+1} + D_{2,t+1}}{p_{2,t}} = \frac{\frac{1}{1-\alpha}P_{2,t+1}C_{t+1} + \frac{1}{1-\alpha}C_{t+1}\left(1 - x_{t+1}\right)}{\frac{1}{1-\alpha}P_{2,t}C_t} \\ &= \frac{P_{2,t+1} + 1 - x_{t+1}}{P_{2,t}}\left[x_t\varepsilon_{1,t+1} + (1-x_t)\,\varepsilon_{2,t+1}\right]\end{aligned}$$

Then, given approximate solutions for the pricing functions, returns are given by:

$$\begin{aligned}\hat{R}_1(x,j;a) &= \frac{\left[x'(j)+\hat{P}_1(x'(j);a)\right]}{\hat{P}_1(x;a)}\left[x\varepsilon_1'+(1-x)\varepsilon_2'\right] \\ \hat{R}_2(x,j;a) &= \frac{\left[1-x'(j)+\hat{P}_2(x'(j);a)\right]}{\hat{P}_1(x;a)}\left[x\varepsilon_1'+(1-x)\varepsilon_2'\right],\end{aligned}$$

where $x'(j)$ is given in (4).

The risk free rate is computed as follows:

$$1=R(x)\beta E_t\left[x\varepsilon_1'+(1-x)\varepsilon_2'\right]^{-\gamma}.$$

Define the mean returns (conditional on the state, x) as follows:

$$\begin{aligned}M_1(x;a) &= \sum_{j=1}^{N}\pi_{ij}\hat{R}_1(x,j;a) \\ M_2(x;b) &= \sum_{j=1}^{N}\pi_{ij}\hat{R}_2(x,j;b),\end{aligned}$$

where the value of i can be anything between 1 and N because of the independence assumption. Finally,

$$\begin{aligned}Cov(x) &= \sum_{j=1}^{N}\pi_{i,j}\left[\hat{R}_1(x,j;a)-M_1(x;a)\right]\left[\hat{R}_2(x,j;a)-M_2(x;a)\right] \\ V_1(x) &= \sum_{j=1}^{N}\pi_{i,j}\left[\hat{R}_1(x,j;a)-M_1(x;a)\right]^2 \\ V_2(x) &= \sum_{j=1}^{N}\pi_{i,j}\left[\hat{R}_2(x,j;b)-M_2(x;b)\right]^2 \\ \rho(x) &= \frac{Cov(x)}{\sqrt{V_1(x)V_2(x)}}.\end{aligned}$$

6 Results

We set the following parameter values:

$$\sigma^2=0.06,\ \beta=0.96,\ \mu=0.02,$$

and the probability of a low, medium or high realization of the ε's is $1/3$. We considered two values of γ, 1 and 2. In the following exercise we used 100 equally-spaced grid points and dynamic programming. That the points are equally spaced may be seen in two of the graphs in Figure 1. Note that there are some Euler errors near the boundaries of zero and unity.

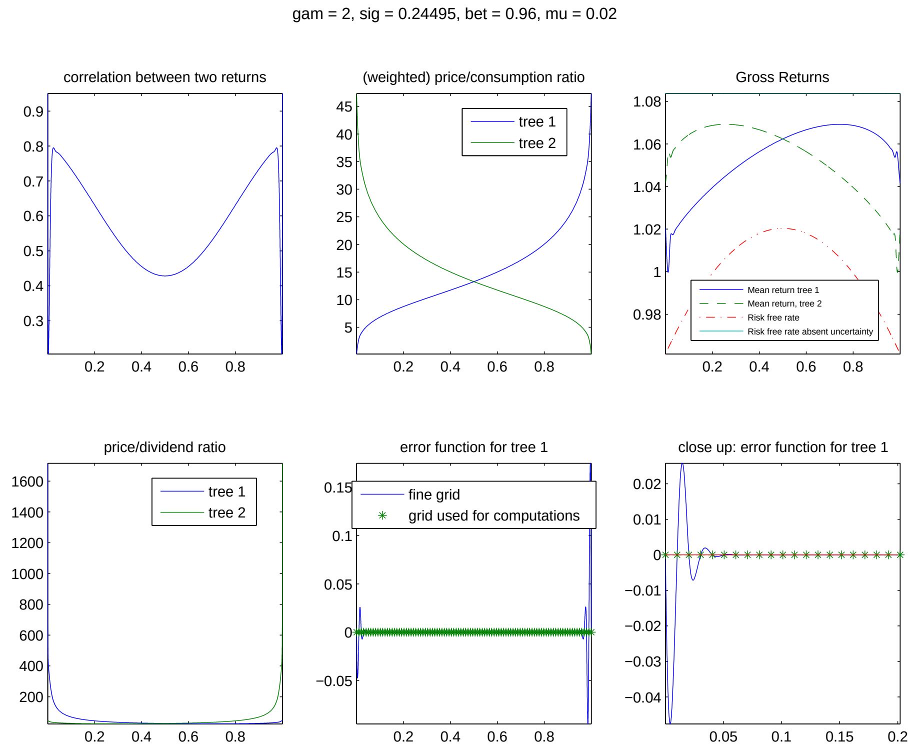

Figure 1

For comparison we also performed the computations when the grid was composed of 100 points constructed from Chebyshev zeros. These are displayed in Figure 2. Note how the grid points are distributed more heavily in the tails. The Euler errors are now somewhat smaller, because relatively more grid points have been allocated in the regions where they tend to be large.

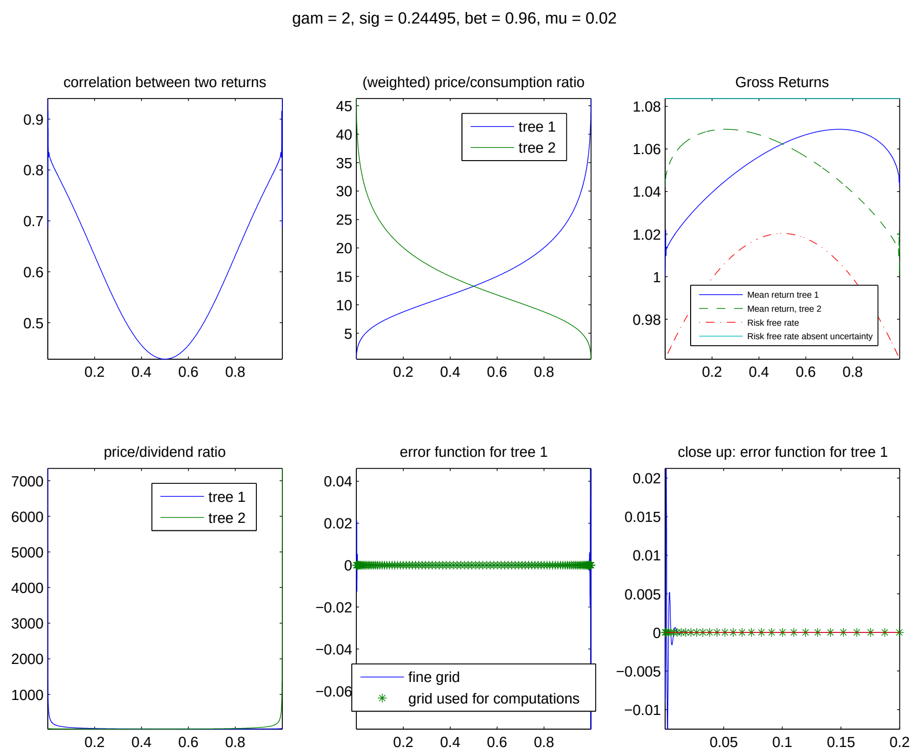

Figure 2

Comparing the results in Figures 1 and 2 we see that with the Chebyshev zeros, the correlations (top left figure) are smoother in the tail areas, as are the mean returns (top right). The price-consumption ratios look similar. The convergence criterion was 0.1×10^{-8}. Converge was slightly faster with the Chebyshev grid, being 28 seconds with that grid and 40 with the equally spaced grid.

Next, we considered the same two experiments considered above, with 500 grid points. Figure 3 displays results with equally-spaced grid points.

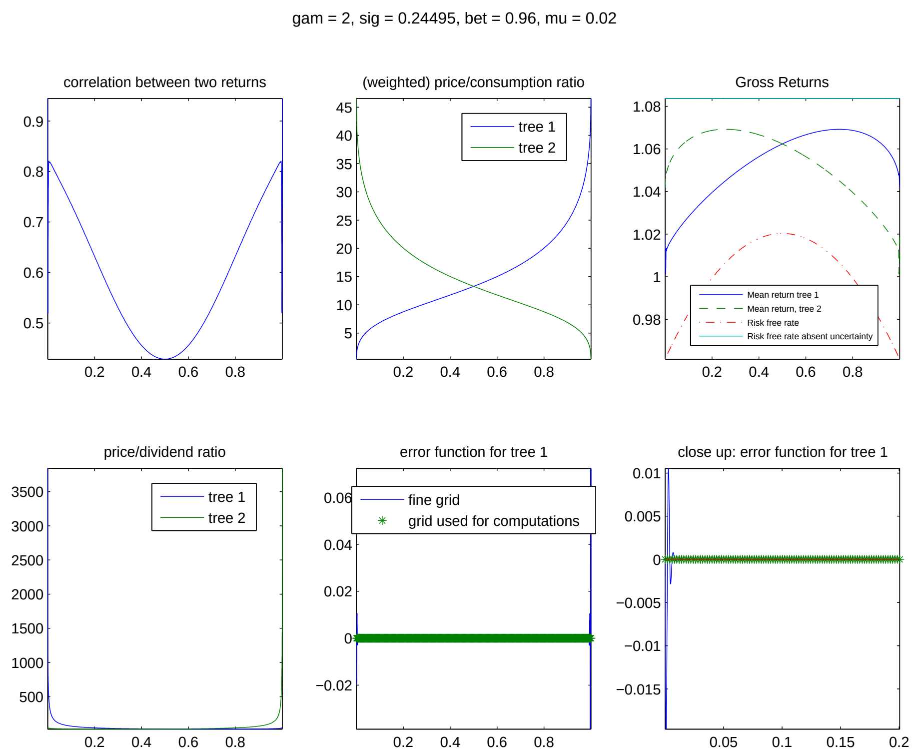

Figure 3

Figure 4 displays the results with Chebyshev grid points.

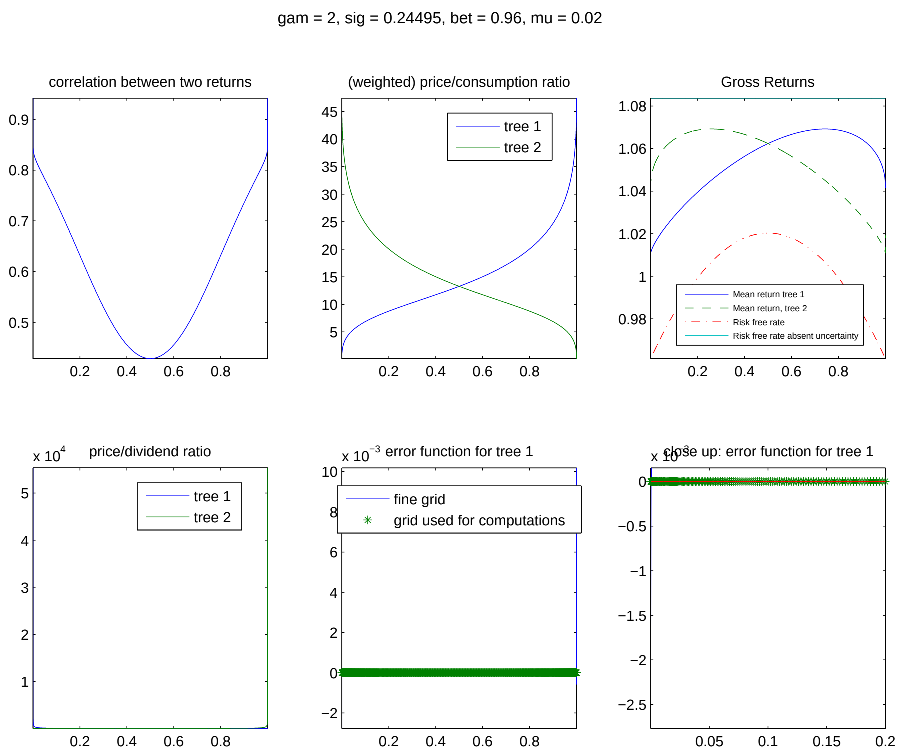

Figure 4

Note that the errors are smaller again for the Chebyshev grid points. However, the results are virtually indistinguishable. One difference is in the price to dividend ratio. With the solution based on the Chebyshev grid, this price ratio rises much higher near the boundaries than it does with the grid with equally-spaced points.

It is interesting that the model predicts such extreme behavior in the price-dividend ratio near the boundaries. To see the implication of this, consider The return on assets:

$$
\begin{aligned}
R_{1,t+1} &= \frac{D_{1,t+1} + p_{1,t+1}}{p_{1,t}} = \frac{1 + \frac{p_{1,t+1}}{D_{1,t+1}}}{\frac{p_{1,t}}{D_{1,t}}} \varepsilon_{1,t+1} \\
R_{2,t+1} &= \frac{1 + \frac{p_{2,t+1}}{D_{2,t+1}}}{\frac{p_{2,t}}{D_{2,t}}} \varepsilon_{2,t+1}
\end{aligned}
$$

Recall that $p_{i,t}/D_{i,t}$ is a function of x_t alone. From this expression, we see that if the price dividend ratio remains constant as x varies, then the returns on the two assets are not correlated. Only if the mapping from x to the price dividend ratio is non-trivial is it possible that the correlation is non-zero. The curvature of the mapping from x to the price dividend ratio is obviously very great and this presumably has something to do with the relatively high correlations between the asset returns. This curvature is connected to the fact that the price dividend ratio is shooting off to infinity at the boundaries. This per se is highly counterfactual, of course. An interesting question is whether this counterfactual property is essential to get the rate of returns to be correlated.

We also performed a simulation. We simulated x's using (4) and then computed various variables corresponding to x at each date and the solution computed using 500 grid points computed using Chebyshev zeros (i.e., the solution reported in Figure 4). The results are displayed in Figure 5.

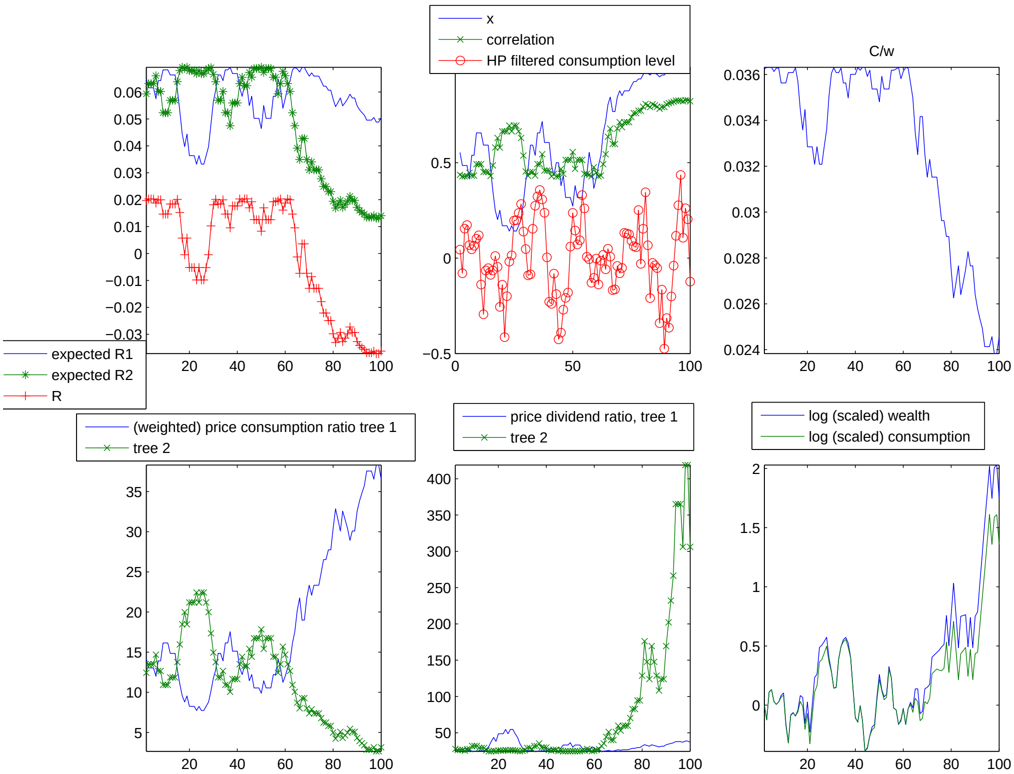

Figure 5

7 The FEM Galerkin Method

Here we describe the Finite Element Method, Galerkin, method as discussed in Christiano-Fisher (2000, Appendix C2, especially footnote 36, also the NBER Technical Working Paper No. 218, issued in October 1997). We consider a finite element method with linear interpolation. We consider a grid of values of x :

$$x_1 < x_2, ..., x_{M-1} < x_M,$$

where $x_1 = 0$ and $x_M = 1$. Consider the basis functions, $L_i(x)$, $i = 1, ..., M$, for a piecewise linear function defined by the values of the function on the grid points ('node points'). For $i = 2, ..., M-1$:

$$L_i(x) = \begin{cases} \frac{x - x_{i-1}}{x_i - x_{i-1}} & x_{i-1} \le x \le x_i \\ \frac{x_{i+1} - x}{x_{i+1} - x_i} & x_i \le x \le x_{i+1} \\ 0 & \text{elsewhere} \end{cases} .$$

This is a sequence of 'tents', each starting at zero for x_{i-1}, peaking at unity for x_i, and returning linearly to zero at x_{i+1}. The successive tents overlap partially. For $i = 1$:

$$L_1(x) = \begin{cases} \frac{x_2 - x}{x_2 - x_1} & x_1 \le x \le x_2 \\ 0 & \text{elsewhere} \end{cases} ,$$

and for $i = M$:

$$L_M(x) = \begin{cases} \frac{x - x_{M-1}}{x_M - x_{M-1}} & x_{M-1} \le x \le x_M \\ 0 & \text{elsewhere} \end{cases} .$$

Let $a = \begin{bmatrix} a_1 & a_M \end{bmatrix}$ and $b = \begin{bmatrix} b_1 & b_M \end{bmatrix}$ denote parameters, as before. These are the values taken on by the piecewise linear functions, $\hat{P}_1(x;a)$ and $\hat{P}_2(x;b)$, at $x = x_i$ for $i = 1, ..., M$. The functions, defined over $x \in [0,1]$ are as follows:

$$\begin{aligned} \hat{P}_1(x;a) &= \sum_{i=1}^{M} a_i L_i(x) \\ \hat{P}_2(x;a) &= \sum_{i=1}^{M} b_i L_i(x) \end{aligned}$$

The Galerkin finite element method makes the error function orthogonal to each of the M basis functions of the piecewise linear functions. That is, it sets to zero the following M objects:

$$\begin{aligned} \int_{x_{i-1}}^{x_{i+1}} E_1(x;a) L_i(x) dx &= 0, \ i = 2, ..., M-1 \\ \int_{x_1}^{x_2} E_1(x;a) L_1(x) dx &= 0 \\ \int_{x_{M-1}}^{x_M} E_1(x;a) L_M(x) dx &= 0. \end{aligned}$$

It does the same for E_2, so there are $2M$ objects to set to zero and $2M$ parameters, a and b. Note that in contrast to Collocation, the criterion is now sensitive to the behavior of E_1 and E_2 over the whole range of x's.

Each of the above M integrals are evaluated by Gauss-Legendre $m-$point quadrature approximation. Thus,

$$\int_A^B E_1(x;a) L_i(x) dx \simeq \sum_{j=1}^{m} E_1(x_j;a) L_i(x_j) v_j,$$

where $A = x_{i-1}$ and $B = x_{i+1}$ for $i = 2, ..., M-1$; $A = x_1$ and $B = x_2$ for $i = 1$; $A = x_{M-1}$ and $B = x_M$ for $i = M$. The number of quadrature points, m, is chosen by the modeler. The quadrature weights, v_j, are defined below. To compute the x_j's belonging to $[A, B]$, we first solve for r_j, $j = 1, ..., m$, the m zeros of the m^{th} order Legendre polynomial, $P_m(y)$. Legendre polynomials are defined as $P_i(y) : [-1, 1] \to [-1, 1]$, for $i = 0, ..., m$, where

$$P_i(y) = 1 + \alpha_1^i y + \alpha_2^i y^2 + ... + \alpha_i^i y^i,$$

where the α's are defined by the requirement,

$$P_0(y) = 1, \quad \int_{-1}^{1} P_i(y) P_j(y) dy, \text{ for } j = 0, ..., i-1 \text{ and } i \geq 0.$$

Then, $x_j = [r_j(B-A) + (A+B)]/2$, $j = 1, ..., m$.