Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.

External Habit and the Cyclicality of Expected Stock Returns

Thomas D. Tallarini, Jr. and Harold H. Zhang

2005-27

External Habit and the Cyclicality of Expected Stock Returns*

Thomas D. Tallarini, Jr.† Harold H. Zhang‡

May 2005

Abstract

We estimate an equilibrium asset pricing model in which agents' preferences have an unobserved external habit using the efficient method of moments (EMM). Given the estimated structural parameters we examine the cyclical behavior of expected stock returns in the model. We find that the estimated structural parameters imply countercyclical expected stock returns as documented in existing empirical studies. The model, however, is still rejected at the one percent level. Detailed examination of the moment conditions in our estimation indicates that the model performs reasonably well in matching the mean of returns, but it fails to capture the higher order moments.

JEL Classification: G12

Keywords: external habit, expected returns, asset pricing, Efficient Method of Moments

*We thank Ronald Gallant and George Tauchen for providing us the EMM code, John Campbell, John Cochrane, Dave DeJong, Wayne Ferson, Rene Garcia, Lars Hansen, Paul Harrison, John Heaton, Ming Liu, Adrian Pagan, Amir Yaron, Stan Zin, and the seminar participants at Carnegie Mellon University, the Chinese University of Hong Kong, and the University of Pittsburgh for helpful comments. The paper was presented at the Second International Conference on Computing in Economics and Finance in Geneva, Switzerland, the Far Eastern Meetings of the Econometric Society in Hong Kong, and the North American Summer Meetings of the Econometric Society in Montreal. The standard disclaimer applies. The views presented are solely those of the authors and do not necessarily represent those of the Federal Reserve Board or its staff.

†Corresponding author. Division of Research and Statistics, Federal Reserve Board, Washington, DC 20551. Phone: (215) 898-3609. Email: thomas.d.tallarini@frb.gov.

‡Kennan-Flagler Business School, University of North Carolina, Chapel Hill, NC 27599. Phone: (919) 843-8340. Email: zhangha@kenan-flagler.unc.edu.

1 Introduction

Expected stock returns are related to the business cycle as shown in papers by Fama and French (1989), Fama (1990), Kandel and Stambaugh (1990), and Harrison and Zhang (1999), and others. In fact, expected returns are countercyclical: higher in recessions and lower in booms. Much of this research has focused on identifying the cyclicality of expected returns without attempting to explain it. 'Standard' models with time separable utility and exogenous endowment processes tend to generate procyclical returns. This is the finding reported in Kandel and Stambaugh (1990). This result is robust to a variety of extensions. Balvers, Cosimano, and McDonald (1990) introduce production, and Zhang (1997) introduces heterogeneous agents and incomplete markets with short-sale constraints into the model examined by Kandel and Stambaugh (1990), both finding that expected stock returns remain procyclical.

Campbell and Cochrane (1999) show that relaxing time separability in preferences can generate countercyclical returns. The mechanism they propose is a slow-moving, non-linear external habit. The representative agent's utility is now a function of current private consumption as well as current and past aggregate consumption. Since the habit is external, individual agents do not consider the effects of current consumption on future utility. The set-up is similar in spirit to Abel's (1990) 'catching up with the Joneses' framework. The habit process moves more slowly than consumption so that in a downturn, consumption falls faster than the habit resulting in an increase in local risk aversion. In an expansion, the opposite happens. The countercyclical effect on risk aversion results in a countercyclical pattern for risk premia and therefore countercyclical expected returns. This explanation is consistent with Black (1990) who argues that risk aversion should be higher in recessions when wealth is low. However, Ljungqvist and Uhlig (1999) show that the Campbell and Cochrane specification implies consumption bunching. This is the result of the habit moving negatively with consumption with these preferences when consumption is endogenous. More standard 'catching up with the Joneses' preferences have consumption and habit moving together. In a Campbell and Cochrane world, a benevolent government counteracts the externality by inducing cycles while in a more standard catching-up world, the benevolent government stabilizes the economy (Ljungqvist and Uhlig (2000)).

In this paper, we estimate the model proposed by Campbell and Cochrane (1999). We use the efficient method of moments (EMM) proposed by Gallant and Tauchen (1996). The EMM is in the same spirit to the 'Indirect Inference' method proposed by Smith (1993) and Gourieroux, Monfort, and Renault (1993). This not only allows us to pin down the structural parameters of interest, but also permits a detailed assessment of the performance of the economic model in matching observed stock returns. The estimation method is based on simulation and allows us to handle the unobserved external habit with relative ease. In a similar situation, Eichenbaum and Hansen (1990) propose that an initial guess for the unobserved variable be used and then the Generalized Method of Moments (GMM) be applied. Following this approach in our case may lead to poor estimates and statistical inference. In particular, the external habit is extremely persistent. Given the small sample size, initial conditions will have a strong effect on the estimates of the other parameters. Since the EMM estimator is simulation-based, we can remove the effect of initial conditions on the persistent habit process by discarding a long series of simulated realizations of returns before we start to collect observations used in our EMM estimation. Specifically, for numerical tractability, we take a two-step approach by first estimating the exogenous driving processes consisting of consumption growth and dividend growth using a vector autoregression. We then estimate the structural parameters such as the discount factor, the risk aversion coefficient, and the persistence parameter for the surplus consumption ratio using the EMM and the exogenous driving processes estimated in the first step.

Using the estimated structural parameters, we then investigate the cyclicality of expected stock returns implied by the economic model at various holding intervals. We first provide some benchmark results on the cyclical behavior of expected returns of the S&P 500 index portfolio at various holding intervals from one quarter to five years. This is done in two steps. In the first step, we obtain the expected holding returns by regressing compounded stock returns on a set of information variables such as the dividend yield, earning-price ratio, default premium, and term premium. In the second step, we regress the expected holding returns on a business cycle proxy as in Harrison and Zhang (1999) to quantify the comovements with business cycles. To examine the cylicality of expected stock returns implied by the economic model, we obtain long series of the expected holding returns at the same holding intervals as for the data using Monte Carlo integration. We then regress the expected holding returns on the business cycle dummy variables constructed for the economic model using the methodology suggested in Rouwenhorst (1995) to quantify their cyclical behavior.

We find that the estimated subjective discount factor is slightly and statistically significantly above one. This result supports the finding reported in Kocherlakota (1990). The estimated persistence parameter for the surplus consumption ratio is statistically significantly above 0.9 indicating the existence of a slow-moving external habit. The economic model with the external habit is rejected at the 1 percent level (but not at the 0.1 percent level). Detailed examination of the moment conditions (both the scores and the conventional moments) indicates that the economic model matches reasonably well the mean stock returns but it fails to match the higher order moments such as variance, skewness, and kurtosis. The external habit implied by the estimated structural parameters generates countercyclical expected stock returns at longer horizons such as one year, two years, and five years. This is driven by the fact that the surplus consumption ratio reacts strongly positively to the contemporaneous consumption growth rate. This implies that the external habit moves at a much slower rate than consumption. The consumption and habit differential thus decreases and agents' local risk aversion increases in recessions. High expected returns are needed to induce agents to hold stocks. Finally, the curvature parameter (the coefficient of relative risk aversion in a time separable environment) has a point estimate of 6.27 which is more than three times the value chosen by Campbell and Cochrane and outside what is generally considered to be the reasonable range for this parameter.

Consistent with the finding that the economic model can match the mean stock returns reasonably well, we also find that the expected simulated stock returns during non-recession periods are comparable to the expected S&P 500 index returns at various holding horizons ranging from one quarter to five years. The spread in expected returns between recession and non-recession periods implied in the economic model is much smaller than the counterpart for the S&P 500 index. This is consistent with the finding that the economic model is unable to match the variance of the observed stock returns.

The rest of the paper is organized as follows. Section 2 presents the model and defines the law of motion for exogenous forcing variables. Section 3 presents and discusses the EMM estimation results and evaluates the performance of the model. Section 4 investigates the cyclicality of expected returns implied in the model, and Section 5 concludes the paper.

2 Preferences with an External Habit

There is an infinitely-lived representative agent who derives utility from consuming a single consumption good. Consumer's preferences depend on both the current consumption and some habit level. Let C_t be the representative agent's consumption at time t and X_t be the agent's habit level at time t. Following Campbell and Cochrane, we assume that X_t depends on economy wide per capita consumption rather than the agent's own consumption though in equilibrium the two are equal. The agent's objective is to maximize the expected sum of discounted future utility given as follows:

$$\max E\left[\sum_{t=0}^{\infty}\beta^t\frac{(C_t-X_t)^{1-\gamma}-1}{1-\gamma}\right] \tag{1}$$

where β is the subjective discount factor and γ is the curvature parameter.[1]

The habit level X_t will evolve as specified by Campbell and Cochrane. It will be treated by individual agents as beyond their control and therefore a function of aggregate per capita consumption. It will follow a non-linear process in order to keep the habit below consumption ($X_t < C_t$). Let S_t denote the surplus consumption ratio:

$$S_t = \frac{C_t - X_t}{C_t}. \tag{2}$$

We will use lowercase letters to denote the natural logs of variables ($x = \log X$). As in Campbell and Cochrane, we assume that the surplus consumption ratio evolves according to:

$$s_{t+1} = (1-\phi)\bar{s} + \phi s_t + \lambda(s_t)(\bar{c}_{t+1} - \bar{c}_t - g) \tag{3}$$

where ϕ determines the persistence of the surplus consumption ratio, $\bar{s}$ is the steady state surplus consumption ratio, $\lambda(s_t)$ is the sensitivity function of the surplus consumption ratio to the contemporaneous per capita consumption growth rate, $\bar{c}_t$ is the log of per capita consumption which equals the consumption of the agent in equilibrium, and g is the average economy wide per capita consumption growth rate.

The sensitivity function, $\lambda(s_t)$, takes the same form as in Campbell and Cochrane:

$$\lambda(s_t) = \begin{cases} (1/\bar{S})\sqrt{1 - 2(s_t - \bar{s})} - 1, & s_t < s_{max} \\ 0 & s_t \geq s_{max} \end{cases} \tag{4}$$

where $\bar{S}$, the steady state surplus consumption ratio, is given by $\sigma\sqrt{\frac{\gamma}{1-\phi}}$ where σ is the standard deviation of consumption growth rate and s_{max}, the upper limit for the surplus consumption ratio, is given by

$$s_{max} = \bar{s} + \frac{1}{2}(1 - \bar{S}^2). \tag{5}$$

The implication of the above specification is that the risk-free rate is constant when the consumption growth rate is a log normal *i.i.d.* random variable, and the habit moves non-negatively with consumption.

We consider a stock with dividend process D_t. The equilibrium stock price, denoted by

[1]Because of the presence of the external habit, the agent's local risk aversion coefficient is no longer γ but rather a function of consumption and habit level.

P_t, can then be expressed as:

$$P_t = E_t\left[\beta(\frac{C_{t+1}-X_{t+1}}{C_t - X_t})^{-\gamma}(P_{t+1}+D_{t+1})\right] = E_t\left[\beta(\frac{S_{t+1}}{S_t})^{-\gamma}(\frac{C_{t+1}}{C_t})^{-\gamma}(P_{t+1}+D_{t+1})\right]. \quad (6)$$

As in Campbell and Cochrane, we model consumption and dividends as separate processes. Since the growth rates of dividends and consumption are only weakly correlated in US data, it can be important to model dividends and consumption separately.

To induce stationarity, we normalize the stock price by dividends (D_t). The normalization yields

$$\tilde{P}_t = E_t\left[\beta(\frac{S_{t+1}}{S_t})^{-\gamma}(\frac{C_{t+1}}{C_t})^{-\gamma}(\tilde{P}_{t+1}+1)(\frac{D_{t+1}}{D_t})\right] \quad (7)$$

where $\tilde{P}_t = P_t/D_t$.

Comparing the pricing function above with the one without the external habit, we have introduced the term $(S_{t+1}/S_t)^{-\gamma}$ which decreases as the growth of the surplus consumption ratio increases. Since the growth of the surplus consumption ratio enters the pricing function exactly as consumption growth, we expect the former to affect the stock price and expected stock returns in the same way as the latter does. The important issue is how the surplus consumption ratio evolves as a function of contemporaneous consumption growth. If the growth of the surplus consumption ratio is highly correlated with per capita consumption growth, the stock price will be strongly procyclical and expected stock returns countercyclical.

The exogenous forcing variables in the model consist of the per capita consumption growth rate ($\Delta c_t = \bar{c}_t - \bar{c}_{t-1}$) and dividend growth rate (Δd_t). We assume that the consumption and dividend growth rates follow a bivariate vector autoregressive (VAR) process with one lag,[2] i.e.,

$$\begin{bmatrix} \Delta c_{t+1} \\ \Delta d_{t+1} \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \Delta c_t \\ \Delta d_t \end{bmatrix} + \begin{bmatrix} \epsilon_{1t+1} \\ \epsilon_{2t+1} \end{bmatrix} \quad (8)$$

where $[\epsilon_{1t}, \epsilon_{2t}]' \sim N(0, \Sigma)$. The sufficient state space for the above problem thus consists of consumption and dividend growth rates, and the surplus consumption ratio of the previous period. Denote by $z_t = [\Delta c_t, \Delta d_t]'$ the vector of exogenous state variables. The one-step

[2]Campbell and Cochrane (1999) suggest that it would be better to make consumption and dividends cointegrated. They report that imposing cointegration does not significantly change their results while complicating some of calculations by introducing an additional state variable. Campbell and Cochrane assume that consumption and dividend growth rates are i.i.d., potentially correlated and have identical means. We relax the i.i.d. and identical mean assumptions in our specification, but we do not impose cointegration.

ahead conditional density for z_t is therefore a Gaussian with mean

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \Delta c_t \\ \Delta d_t \end{bmatrix}$$

and covariance matrix Σ. We can thus represent the state space by (z_t, S_{t-1}). The stock price can then be written as

$$\begin{aligned}\tilde{P}(z_t, S_{t-1}) = E\{\beta[\frac{S(z_{t+1}, S(z_t, S_{t-1}))}{S(z_t, S_{t-1})}]^{-\gamma} \exp(\Delta c_{t+1})^{-\gamma} \\ [\tilde{P}(z_{t+1}, S(z_t, S_{t-1})) + 1] \exp(\Delta d_{t+1}) | z_t, S_{t-1}\}.\end{aligned} \quad (9)$$

To estimate the bivariate VAR specification for the exogenous state variables, we construct quarterly per capita consumption growth rate of non-durables and services and per capita dividend growth rate adjusted by the population growth (age 16 and above) using the CITIBASE data from 1947:Q2 to 1995:Q3. Table 1 presents the estimation results. We make the following observations. Consumption growth at time t is positively related to both consumption and dividend growth rates at time $t-1$. The estimates are statistically significant at the 5 percent level. While the dividend growth rate is also positively related to the lagged consumption and dividend growth rates, the estimates are not statistically significant at the conventional test level. In the next section, we formally estimate the economic model with the external habit and provide a statistical assessment of its performance in matching the observed stock returns.

3 EMM Estimation of Structural Parameters

3.1 The EMM Estimator

Our goal here is to jointly estimate the Euler equation and the law of motion for the external habit for a given exogenous driving force:[3]

$$\tilde{P}_t = E_t\left[\beta(\frac{S_{t+1}}{S_t})^{-\gamma}(\frac{C_{t+1}}{C_t})^{-\gamma}(\tilde{P}_{t+1} + 1)\frac{D_{t+1}}{D_t}\right], \quad (10)$$

$$s_{t+1} = (1-\phi)\bar{s} + \phi s_t + \lambda(s_t)(\Delta \bar{c}_{t+1} - g), \quad (11)$$

[3]The vector autoregressive process for the exogenous forcing variables is fixed at the point estimates reported in Table 1 when we estimate the structural parameters governing preferences. Thus, sampling errors in the first step estimation are not explicitly accounted for.

$$\lambda(s_t) = \begin{cases} (1/\bar{s})\sqrt{1-2(s_t-\bar{s})}-1, & s_t < s_{max} \\ 0, & s_t \geq s_{max}. \end{cases} \quad (12)$$

Because the external habit is not observed, to apply the GMM as in Eichenbaum and Hansen (1990), one needs to jointly estimate the initial surplus consumption ratio (s_0) and the other structural parameters. Using the S&P 500 value weighted returns and the returns of ten decile portfolios in combination with the risk-free rate as the moment conditions, we find that the curvature parameter (γ) estimate is in general quite small (less than 1.0 in most cases). It gets larger when the initial surplus consumption is fixed at the steady state level but it is still usually less than 3.0. All the models have a high p-value consistently above 10 percent implying that the model with an external habit cannot be rejected at conventional test levels. However, this may simply reflect that the model is poorly estimated when the surplus consumption ratio process is highly persistent and the sample size is too small. Indeed, the estimated persistence parameter (ϕ) is consistently above 0.98 implying strong persistence in the surplus consumption ratio process. The high persistence parameter estimate coupled with a relatively small sample size thus render the GMM estimation unreliable in this case.[4]

A simulation based generalization of GMM proposed in Bansal, Gallant, Hussey and Tauchen (1994) and Gallant and Tauchen (1996), the EMM, is well suited to the problem at hand. It allows us to remove the effect of the initial value of the surplus consumption ratio when the habit process is highly persistent and the sample size is small. This is achieved by first simulating a long series of exogenous variables, the consumption and dividend growth rates, and then recursively computing the surplus consumption ratio and the stock returns starting from an initial guess for the surplus consumption ratio. We discard the first several thousand simulated stock returns before we collect data for our EMM estimation. The idea of the EMM is to use the expectation under the structural model of the scores from an auxiliary model (called a *score generator*) as the vector of moment conditions. The scores are the derivatives of the log likelihood function of the auxiliary model with respect to its parameters. The estimator is defined as follows.

Let $\{\tilde{y}_t, \tilde{x}_{t-1}\}_{t=1}^n$ denote the observed data set (in our case, $\{\tilde{y}_t\}$ represents the S&P 500 index return series), where $\tilde{x}_{t-1} = (\tilde{y}_{t-L}, \cdots, \tilde{y}_{t-1})', L \geq 1$. Let $f(\tilde{y}_t|\tilde{x}_{t-1}, \theta), \theta \in \Theta \subset \Re^{l_\theta}$, be the one-step conditional density function which is the auxiliary model that generates scores for the economic model to match. Let $\rho \in \varrho \subset \Re^{l_\rho}$ be the vector of structural parameters to be estimated, and $\{\hat{y}_\tau(\rho), \hat{x}_{\tau-1}(\rho)\}_{\tau=1}^N$ be the counterpart of the observed data simulated from

[4]The GMM results are not reported to save space and are available upon request. Interestingly, the results are broadly consistent with the findings reported in Ferson and Constantinides (1991) for a model with an internal habit formation using quarterly data.

the structural model given the structural parameters ρ. The simulated data can be obtained by numerically solving equations (10), (11), and (12) jointly for a given exogenous driving force. We choose to implement the parameterized expectations algorithm proposed by Den Haan and Marcet (1990), and applied in Bansal, Gallant, Hussey and Tauchen (1993) in connection with the simulated method of moments estimation. Specifically, we parameterize the right-hand-side of equation (10) as follows:

$$E_t\left[\beta(\frac{S_{t+1}}{S_t})^{-\gamma}(\frac{C_{t+1}}{C_t})^{-\gamma}(\tilde{P}_{t+1}+1)\frac{D_{t+1}}{D_t}\right]=\exp[h(z_t,S_{t-1};\psi)] \tag{13}$$

where $h(z_t,S_{t-1};\psi)$ is a polynomial in state variables $(z_t,S_{t-1})'$ and ψ is a vector of parameters for the polynomial. In our application, we find the following quadratic polynomial to work very well:

$$h(z_t,S_{t-1};\psi)=\psi_0+\psi_1\Delta c_t+\psi_2\Delta d_t+\psi_3 S_{t-1}+\psi_4\Delta c_t^2+\psi_5\Delta d_t^2+\psi_6 S_{t-1}^2. \tag{14}$$

We first estimate the auxiliary parameter θ using quasi-maximum likelihood estimation:

$$\tilde{\theta}_n=argmax_{\theta\in\Theta}\frac{1}{n}\sum_{t=1}^{n}\ln f(\tilde{y}_t|\tilde{x}_{t-1},\theta). \tag{15}$$

This is achieved by solving the following set of first-order conditions:

$$\frac{1}{n}\sum_{t=1}^{n}(\partial/\partial\theta)\ln f[\tilde{y}_t|\tilde{x}_{t-1},\tilde{\theta}_n]=0. \tag{16}$$

The basic idea of EMM estimation is that under the null hypothesis that the structural model is correctly specified, there exists a true parameter vector $\rho^0\in\varrho$ such that the simulated data $\{\hat{y}_\tau(\rho^0),\hat{x}_{\tau-1}(\rho^0)\}_{\tau=1}^{N}$ also satisfy the above first-order conditions.

We thus introduce the moment conditions as follows:

$$m_n(\rho,\tilde{\theta}_n)=\frac{1}{N}\sum_{\tau=1}^{N}(\partial/\partial\theta)\ln f[\hat{y}_\tau(\rho)|\hat{x}_{\tau-1}(\rho),\tilde{\theta}_n]. \tag{17}$$

The EMM estimator of the structural parameter vector is then defined analogously to the GMM estimator as follows:

$$\hat{\rho}_n=argmin_{\rho\in\varrho}m_n'(\rho,\tilde{\theta}_n)(\tilde{\mathcal{I}}_n)^{-1}m_n(\rho,\tilde{\theta}_n) \tag{18}$$

where $(\tilde{\mathcal{I}}_n)^{-1}$ is the weighting matrix and is given by the outer-product-of-gradient formula

$$\tilde{\mathcal{I}}_n = \frac{1}{n}\sum_{t=1}^{n}[(\partial/\partial\theta)\ln f(\tilde{y}_t|\tilde{x}_{t-1},\tilde{\theta}_n)][(\partial/\partial\theta)\ln f(\tilde{y}_t|\tilde{x}_{t-1},\tilde{\theta}_n)]'. \tag{19}$$

Gallant and Tauchen (1996) show that the EMM estimator defined above is consistent and asymptotically normal, and the normalized objective value

$$\tilde{s}_n(\hat{\rho}_n) = nm_n'(\hat{\rho}_n,\tilde{\theta}_n)(\tilde{\mathcal{I}}_n)^{-1}m_n(\hat{\rho}_n,\tilde{\theta}_n) \tag{20}$$

has an asymptotic chi-square distribution with degrees of freedom $l_\theta - l_\rho$, where l_θ and l_ρ are the numbers of parameters in auxiliary and economic models, respectively.

Because the conditional density function, which serves as the auxiliary model, is usually unknown, in practice, it needs to be estimated. We employ the Gallant and Tauchen (1989) seminonparametric (SNP) method to estimate $f(\tilde{y}_t|\tilde{x}_{t-1},\theta)$ along with the auxiliary parameter vector, θ. Next, we provide the empirical estimation results.

3.2 Empirical Results

3.2.1 SNP Density Estimate of Real Stock Returns

In their baseline formulation, Campbell and Cochrane (1999) choose the sensitivity function $\lambda(s_t)$ so that the risk-free rate is constant and the habit moves nonnegatively with consumption. In an alternative formulation, they allow the risk-free rate to vary with the state variable, for instance, let the risk-free rate be a linear function of $(s_t - \bar{s})$. However, Campbell and Cochrane show that adding interest rate variation in this way has very little effect on the stock market results. Following Campbell and Cochrane, we focus on the stock returns.[5] Specifically, the series that we choose to match for our economic model is the quarterly real returns of the S&P 500 index portfolio. The data spans the period from the second quarter of 1947 to the third quarter of 1995. The *ex post* real returns are obtained by adjusting the nominal returns of the S&P 500 index portfolio by inflation. We estimate the conditional density function for the real returns using the SNP method which we briefly discuss below.

The SNP method is based on the notion that a Hermite expansion can be used as a general purpose approximation to a density function. This basic approach can be adapted to the estimation of the conditional density of the return series, denoted $\{y_t\}$, that has a Markovian structure – where the conditional density of y_t given the entire history $\{y_{t-l}\}_{l=1}^{\infty}$ depends only on L lags from the most recent past. Collecting these lags together in a single

[5] The implications for the risk-free rate will be discussed below.

vector gives a L-vector denoted as x_{t-1}:

$$x_{t-1} = (y_{t-L}, \cdots, y_{t-2}, y_{t-1})'.$$

The SNP approximation of a density takes the following form:

$$f(y_t|x_{t-1}, \theta) = \frac{1}{\xi}[P(z_t, x_{t-1})]^2 n(y_t|\mu_{x_{t-1}}, \sigma_{x_{t-1}}) \tag{21}$$

where ξ is a scalar that makes the density integrate to one,[6] $z_t = \sigma^{-1}_{x_{t-1}}(y_t - \mu_{x_{t-1}})$ is an innovation, $P(z_t, x_{t-1})$ denotes a polynomial in z_t of degree K_z whose coefficients are polynomials of degree K_x in x_{t-1}, $n(y_t|\mu_{x_{t-1}}, \sigma_{x_{t-1}})$ is a normal distribution with mean $\mu_{x_{t-1}}$ (the location function) and standard deviation $\sigma_{x_{t-1}}$ (the scale function) whose values depend on x_{t-1}. The constant term of the polynomial is put to one to obtain a unique representation. This normalization means that the leading term of the entire expansion is $n(y_t|\mu_{x_{t-1}}, \sigma_{x_{t-1}})$.[7] The location function $\mu_{x_{t-1}}$ is given by an autoregression

$$\mu_{x_{t-1}} = b_0 + Bx_{t-1}. \tag{22}$$

It is assumed to depend on $L_\mu \leq L$ lags. The scale function $\sigma_{x_{t-1}}$ is given by

$$\sigma_{x_{t-1}} = \rho_0 + P|e^*_{t-1}| \tag{23}$$

where $e^*_{t-1} = [(y_{t-L_r} - \mu_{x_{t-L_r-1}}), \cdots, (y_{t-1} - \mu_{x_{t-2}})]$, and $|\cdot|$ denotes elementwise absolute value. The scale function depends on L_r lagged (unnormalized) innovations $(y_t - \mu_{x_{t-1}})$ and $(L_\mu + L_r) \leq L$ lagged y_t in total. This is an ARCH–type process akin to that proposed by Nelson (1991).

The Hermite polynomial $P(z_t, x_{t-1})$ is given by

$$P(z_t, x_{t-1}) = \sum_{\alpha=0}^{K_z} (\sum_{\beta=0}^{K_x} a_{\alpha\beta} x^\beta_{t-1}) z^\alpha_t \tag{24}$$

where $\alpha \leq K_z$ and $\beta \leq K_x$ are non-negative integers. It is assumed that the polynomial depends on $L_p \leq L$ lags of y from x.

[6]Therefore $\xi = 1/\int [P(s,x)]^2 \phi(s) ds$.

[7]The vector θ of $f(y_t|x_{t-1}, \theta)$ consists of the coefficients of the polynomial plus $\mu_{x_{t-1}}$ and $\sigma_{x_{t-1}}$ and is estimated by maximum likelihood. Equivalent to maximum likelihood, but more stable numerically, is to estimate θ in a sample of size n by minimizing $s_n(\theta) = -\frac{1}{n}\sum_{t=1}^{n} log[f(y_t|x_{t-1}, \theta)]$. If the number of parameters p_θ grows with the sample size n, then the true density, its derivatives, and moments are estimated consistently as shown in Gallant and Nychka (1987).

When K_z is positive, the resulting density function is a modification of the Gaussian due to the multiplication by the polynomial $\{P(z_t)\}^2$. When K_x is positive, the shape of the density will depend on x_{t-1}. Thus, all moments can depend on x_{t-1} and the density can approximate any form of conditional heterogeneity (Gallant and Tauchen, 1989). The shape modifications are rich enough to accurately approximate densities from a large class that includes densities that have fat tails, thin tails, or are skewed.[8]

For notational convenience, hereafter, the hierarchical SNP structure is denoted as SNP(L_μ, L_r,L_p,K_z,K_x). To illustrate, consider first the model with $L_\mu = 4$, $L_r = 4$, $L_p = 0$, $K_z = 4$, and $K_x = 0$. The polynomial is of the form

$$P(z_t) = \sum_{\alpha=0}^{4} a_\alpha z_t^\alpha \tag{25}$$

where the $a_0, a_1, \cdots, a_4$ are the polynomial coefficients with the constant term $a_0 = 1$ to achieve a unique representation. Both $\mu_{x_{t-1}}$ and $\sigma_{x_{t-1}}$ are linear in $y_{t-1}, \cdots, y_{t-4}$. The model has fourteen free parameters: the four free polynomial parameters, the intercept and four slope parameters in $\mu_{x_{t-1}}$, and the intercept and four slope parameters of $\sigma_{x_{t-1}}$.

Now consider $L_p = 1$ and $K_x = 1$ but everything else the same. The polynomial becomes

$$P(z_t, x_{t-1}) = \sum_{\alpha=0}^{4} (a_{0\alpha} + a_{1\alpha} y_{t-1}) z_t^\alpha. \tag{26}$$

The normalization is $a_{00} = 1$. The polynomial has nine free parameters, yielding nineteen free parameters in total.

To select the optimal SNP model for the real returns, the following strategy is adopted. We start with a VAR process, SNP(10100), and gradually expand L_μ until certain model selection criterion reaches a minimum. Three model selection criteria are calculated for each SNP fit: the Schwarz criterion $[s_n + \frac{p_\theta}{2n}\ln(n)]$, the Hannan-Quinn criterion $[s_n + \frac{p_\theta}{n}\ln[\ln(n)]]$, and the Akaike criterion $[s_n + \frac{p_\theta}{n}]$. We then introduce ARCH by increasing L_r, introduce non-Gaussian ARCH by increasing K_z, and finally bring in general nonlinear processes by increasing L_p and K_x.

The preferred models are then subject to a battery of diagnostic tests to determine the goodness-of-fit. The diagnostic tests entail checking for predictability in the residuals from

[8] Fenton and Gallant (1996) assess the qualitative behavior of SNP in finite samples using the Marron and Wand (1992) test suite. For each of the fifteen densities proposed by Marron and Wand, Fenton and Gallant generate samples at sizes 400, 900, 1600, 2500, and 5625. To each they fit a kernel and SNP. For both kernel and SNP, they compute a Riemann sum of the absolute value of the distance between the estimated density and the true density on 1024 points evenly spaced over the interval $[-3, 3]$. Their results suggest that the SNP estimator for a univariate series is both qualitatively and asymptotically similar to the kernel estimator.