CHAPTER 2

MACROECONOMIC INFORMATION IN BOND PRICES

This chapter constructs the programming framework necessary to build an equilibrium model of the term structure of interest rates. The well-known first-order conditions are inverted to recover information about aggregate real consumption growth from real bond yields. Two econometric methodologies are proposed to deal with the separation of the interest rate and consumption variables. The first method, suggested by Hansen and Singleton (1983), involves making strong assumptions on the distribution of the joint consumption and returns process. With these assumptions, the parameters can be estimated linearly. The second strategy developed by Hansen (1982) and Hansen and Singleton (1982) involves weaker assumptions and allows for consistent estimates of the parameters of the non-linear first-order conditions with the Generalized Method of Moments (GMM) technique.

2.1 The Consumer's Planning Problem

Consider a pure exchange economy with a representative agent with additively separable utility receiving a stochastic endowment. This agent can choose to consume this endowment or invest in a portfolio of j–period bonds, $\{B_{j,t}: \quad j = 1, \ldots, k\}$. The bonds are assumed to be default free. Expectations at time t are conditioned on the information set $\mathbf{F}_t$ – which contains all the information about the environment available at time t. Consumption C_t is required to be measurable at t with respect to $\mathbf{F}_t$. The consumer maximizes the following objective:

$$\max_{\{C_t, B_{j,t}\}_{t=0}^{\infty} \quad j=1,\ldots,k} \quad \sum_{t=0}^{\infty} \delta^t E_0 U(C_t); \qquad 0 < \delta < 1,$$

subject to:

$$C_t^N + \sum_{j=1}^{k} B_{j,t} \leq Y_t^N + \sum_{j=1}^{k} B_{j,t-j}(1 + R_{j,t-j}^N), \tag{2.1}$$

where C_t is the agent's real consumption, C_t^N represents nominal consumption, $R_{j,t-j}^N$ is the nominal yield on a j–period bond bought at time $t-j$, E_t is the expectation operator conditioned on information $\mathbf{F}_t$, Y_t^N is the nominal endowment[1] and δ is the consumer's constant time discount factor. The first-order necessary conditions[2] are:

$$E\left[\delta^j \frac{U'(C_{t+j})}{U'(C_t)}(1 + R_{j,t}) - 1 \,\middle|\, \mathbf{F}_t\right] = 0, \qquad \text{for} \quad j = 1, \ldots, k, \tag{2.2}$$

where $R_{j,t}$ represents the real yield on a j–period bond. To keep the consumer's problem well behaved, the arbitrary rolling over of debt is ruled out. This effectively imposes a limitation on the terminal value of the consumer's debt holding. Condition (2.2) is necessary to characterize an optimal plan. The Euler equation (2.2) provides the basis for the intertemporal consumption-based asset pricing model. Sufficient conditions are given in Rubinstein (1976), Breeden and Litzenberger (1978), Lucas (1978), Brock (1982) and Breeden (1986).

[1] The conditional expectation $E_t[Y_{t+j}^N]$ is assumed to exist for all $j \geq 1$.

[2] See appendix A for a more detailed discussion of the problem and solution.

Equation (2.2) depicts a non-linear relation between the marginal rate of substitution and interest rates. Note that the real interest rate, $R_{j,t}$, represents the total return over the period t to $t + j$. If this value is known at time t and the parameterization of the utility function is also known, then it is possible to solve for the expected marginal rate of substitution. Under some utility specifications, the marginal rate of substitution can be linked to the growth rate in consumption. With this type of specification, the real interest rate should forecast future economic growth.

In practice, the real interest rate is not known at time t. The idea of this paper is to look at expected real interest rates and to test if they contain information about future growth in the economy. The next two sections outline methodologies that allow for the estimation of consumption forecasting equations.

2.2 The Linear Specification

Let utility be represented by the constant relative risk aversion or isoelastic class:[3]

$$U(C,\alpha) = \begin{cases} \frac{C^{1-\alpha}-1}{1-\alpha}, & \text{if } \alpha > 0,\ \alpha \neq 1; \\ \log(C), & \text{if } \alpha = 1. \end{cases} \tag{2.3}$$

With this convenient form, we can rewrite the initial first–order conditions as:

$$E_t\left[\delta^j \left\{\frac{C_t}{C_{t+j}}\right\}^\alpha (1+R_{j,t})\right] = 1 \qquad j = 1, \ldots k. \tag{2.4}$$

Following Hansen and Singleton (1983), suppose that the process that characterizes the marginal rates of substitution and the returns is stationary jointly lognormally distributed. Then (2.4) can be re-written:

$$\log E_t\left[\delta^j \left\{\frac{C_t}{C_{t+j}}\right\}^\alpha (1+R_{j,t})\right] = E_t \log\left[\delta^j \left\{\frac{C_t}{C_{t+j}}\right\}^\alpha (1+R_{j,t})\right] + \frac{1}{2}\text{var}_t \log\left[\delta^j \left\{\frac{C_t}{C_{t+j}}\right\}^\alpha (1+R_{j,t})\right] = 0. \tag{2.5}$$

The RHS of (2.5) can be rearranged to bring expected consumption growth to the LHS.

$$E_t\left[\log\frac{C_{t+j}}{C_t}\right] = \frac{j}{\alpha}\log\delta + \frac{v_j}{2\alpha} + \frac{1}{\alpha}E_t\left[\log(1+R_{j,t})\right], \tag{2.6}$$

where v_j is the variance term in (2.5) which is assumed constant. Equation (2.6) is be estimated by least squares in the form:

[3] The empirical section also considers negative exponential utility which implies constant absolute risk aversion.

$$\log \frac{C_{t+j}}{C_t} = \beta_0^j + \beta_1 E_t \left[\log\left(1 + R_{j,t}\right)\right] + \epsilon_{j,t+j}. \tag{2.7}$$

The coefficients should equal

$$\beta_0^j = \frac{j}{\alpha} \log \delta + \frac{v_j}{2\alpha},$$

$$\beta_1 = \frac{1}{\alpha}.$$

The β_1 coefficient can be considered an elasticity as well as one over the relative risk aversion. In the Life Cycle–Permanent Income Hypothesis literature, this coefficient is sometimes referred to as the *elasticity of intertemporal substitution*. It can be interpreted as the sensitivity of consumption growth to changes in expected real rates. Recently, Hall (1985) has argued that this elasticity is very small and perhaps even zero. This implies that there is no information in the expected real rate that is relevant for forecasting real consumption growth. This paper will provide an alternate way of estimating the inverse of the risk aversion parameter or the elasticity of substitution by looking at the term structure rather than a single short term interest rate. Evidence is presented that suggests that there is some information in the term structure that is useful in forecasting consumption growth.

Note that the least squares specification contains an expected value for a regressor. A generated regressor will often lead to an errors in the variables problem. This problem is usually addressed by using an instrumental variables technique. The generated regressor vanishes if β_1 is assumed to be unity (logarithmic utility). It is immediate that expected inflation will cancel from both sides of equation (2.7). Since the nominal rate is known at the beginning of the period, the expectation operator can be dropped. The empirical section documents the results of both the ordinary least squares regressions and instrumental variables estimation for the case of an unrestricted coefficient of relative risk aversion. Results are also presented for case where logarithmic utility is assumed.

If $j > 1$, then the error process $\{\epsilon_{j,t+j} : t \geq 1\}$ will not be, in general, independently distributed due to an overlapping dependent variable. The standard errors on the regression coefficients need to be corrected for an induced moving average process in the residuals. Following Hansen (1982) and White (1980), all standard errors are corrected for the moving average process and are heteroskedastic consistent. [4]

2.3 The Generalized Method of Moments Approach[5]

While the assumption of joint lognormality for the consumption-returns process provides a simple way to obtain a consumption forecasting equation, it places a strong restriction on the behavior of the data that may not be realistic. If the assumption is violated, the parameter estimates will not be consistent. This causes obvious problems in obtaining forecasts. Hansen's (1982) Generalized Method of Moments (GMM) technique allows for the consistent estimation of the parameters of the first-order conditions with far weaker assumptions.

The GMM serves three uses here. First, the GMM based parameter estimates can be compared to the linear estimates. If the assumptions of the linear specification are true, then the two techniques should deliver the same parameter

---

[4] Although the theory implies homoskedasticity, the use of heteroskedastic consistent variance-covariance matrix should not over turn any of the large sample results.

[5] For a detailed description of this technique see Hansen (1982) and Hansen and Singleton (1982).

estimates in a large sample. If the estimates are similar, then it is less problematic to use the linear version for forecasting. If the estimates diverge, it could indicate that the assumptions of the linearized version are violated. Second, the instrumental variables estimation of the linear specification can use the linear version of the GMM. The *generalized* instrumental variables estimation provides consistent parameter estimates and the standard errors are adjusted to take any serial correlation into account. Third, the GMM allows one to test the specification of the consumption-based model. If the model is rejected, then it is questionable whether the model can be used for forecasting.

The following is a brief description of how the technique works. Consider the first-order conditions:

$$E\left[\delta^j \left\{\frac{C_t}{C_{t+j}}\right\}^{\alpha}(1+R_{j,t})-1 \;\middle|\; \mathbf{F}_t\right] = h(x_{t+j},\theta_0) = 0 \tag{2.8}$$

where $\mathbf{F}_t$ is the market information set, x_{t+j} is the data and $\theta_0 = \{\alpha_0, \delta_0\}$ is the parameter vector selected from a compact ℓ–dimensional parameter space. It is assumed that the error process has finite second moments. Condition (2.8) implies that errors are uncorrelated with variables in the market information set $\mathbf{F}_t$. Now consider an instrument vector, $\mathbf{Z}_t$, which is part of the the market information set, i.e. $\mathbf{Z}_t \subseteq \mathbf{F}_t$. Condition (2.8) implies:

$$E\left[h(x_{t+j},\theta_0) \;\middle|\; \mathbf{F}_t\right]\mathbf{Z}_t = 0\mathbf{Z}_t = \mathbf{0}. \tag{2.9}$$

By the law of iterated expectations, (2.9) implies:

$$E\left[h(x_{t+j},\theta_0)\mathbf{Z}_t\right] = \mathbf{0}. \tag{2.10}$$

This expression is written in terms of unconditional expectations. From (2.10), it is possible to construct an estimator of θ_0 as long as the number of orthogonality conditions (instruments), r, is greater than or equal to the number of parameters to be estimated, ℓ.

Let

$$\mathbf{G}_0(\theta) = E\left[h(x_{t+j},\theta)\mathbf{Z}_t\right]. \tag{2.11}$$

Note that $\mathbf{G}_0(\theta)$ has a zero at $\theta = \theta_0$. The method of moments estimator for $\mathbf{G}_0$ is:

$$\mathbf{G}_T(\theta) = \frac{1}{T}\sum_{t=1}^{T} h(x_{t+j},\theta)\mathbf{Z}_t. \tag{2.12}$$

At $\theta = \theta_0$, $\mathbf{G}_T(\theta_0)$ should be close to zero as T gets large. Equation (2.12) provides the foundation for the GMM technique. The objective is to search for parameters that force (2.12) to be as close as possible to the zero vector. The parameter vector θ_T is chosen by minimizing the quadratic form:

$$J_T(\theta) = \mathbf{G}_T(\theta)'\mathbf{W}_T\mathbf{G}_T(\theta), \tag{2.13}$$

where $\mathbf{W}_T$ is a symmetric non-singular weighting matrix that defines the metric used to make $\mathbf{G}_T$ as close to zero as possible. Hansen (1982) shows that if, among other assumptions, the parameter space is compact, $\frac{\partial h}{\partial \theta}$ is continuous and the stochastic process $\{(x_{t+j}, z_t) : t \geq 1\}$ is stationary and ergodic, then the weighting matrix, $\mathbf{W}_T$, will almost surely converge to a constant, $\mathbf{W}_0$. This implies that θ_T will almost surely converge to θ_0. This guarantees strong consistency and asymptotic normality of the estimator.

If $j = 1$, then the $r \times r$ weighting matrix is computed by estimating:

$$\mathbf{W}_T^* = \left[\sum_{t=1}^{T} [h(x_{t+1}, \theta_T)\mathbf{Z}_t)][h(x_{t+1}, \theta_T)\mathbf{Z}_t]' \right]^{-1}. \tag{2.14}$$

If $j > 1$, then the error process will be serially correlated. Appendix B demonstrates how to construct the weighting matrix in this case. Note that an estimate of θ_T is necessary in order to solve for W_T^*. The standard estimation strategy proceeds in two stages. First, a sub-optimal choice of W_T, such as the identity matrix, is used in the minimization of the objective function (2.13). As a result of this minimization, an initial parameter vector θ_T obtains. In the second stage, the initial parameter vector is used to solve for the optimal weighting matrix, W_T^* in (2.14). This matrix is used in the objective function and the final parameter vector θ_T^* is solved for.

The limiting variance-covariance matrix of the GMM estimator is consistently estimated by:

$$\Sigma_T^* = \left[\left\{ \frac{1}{T} \sum_{t=1}^{T} \left[\frac{\partial h(x_{t+1}, \theta_T^*)}{\partial \theta_T^*} \mathbf{Z}_t \right] \right\} \mathbf{W}_T^* \left\{ \frac{1}{T} \sum_{t=1}^{T} \left[\frac{\partial h(x_{t+1}, \theta_T^*)}{\partial \theta_T^*} \mathbf{Z}_t \right] \right\}' \right]^{-1}. \tag{2.15}$$

Furthermore, the number of observations times the minimized value of the objective function in (2.13) is distributed χ^2 with $r - \ell$ (the number of orthogonality conditions less the number of parameters) degrees of freedom. This statistic provides a test of the *over-identifying* restrictions[7] in (2.12).

There are a number of advantages in using the GMM procedure to estimate the non-linear first-order conditions. The strong distributional assumption of stationary joint lognormality need not be made. The GMM only requires that process

[7] Consistent estimates can be obtained with the $r = \ell$. If $r > \ell$, then there are $r - \ell$ over-identifying restrictions that can be tested.

$\{(x_{t+j}, z_t) : t \geq 1\}$ be stationary and ergodic. The linear representation forces the conditional covariance between returns and marginal rates of substitution to be constant through time. The GMM does not impose this restriction. Finally, the error process can be allowed to be conditionally heteroskedastistic. It is not necessary to characterize the dependence of the conditional variances when using the GMM technique.

2.4 Real Interest Rates and Yield Spreads

It is also of some interest to examine a measure of the term structure: the spread between two annualized yields of different maturity. Kessel (1965) documented the cyclical nature of the term structure of interest rates. By subtracting the one–period version of (2.6) from the j–period formulation, consumption growth can be linked to the slope of the yield curve:

$$E_t\left[\log\frac{C_{t+j}}{C_{t+1}}\right] = \frac{j-1}{\alpha}\log\delta + \frac{v_j - v_1}{2\alpha} + \frac{1}{\alpha}E_t\left[\log\frac{1+R_{j,t}}{(1+R_{1,t})^j}\right] + \frac{j-1}{\alpha}E_t\left[\log(1+R_{1,t})\right]. \tag{2.16}$$

This can be estimated in the form:

$$\log\frac{C_{t+j}}{C_{t+1}} = \beta_0^{j-1} + \beta_1 E_t\left[\log\frac{1+R_{j,t}}{(1+R_{1,t})^j}\right] + \beta_2 E_t\left[\log 1 + R_{1,t}\right] + \epsilon_{j-1,t+j}. \tag{2.17}$$

The coefficients should equal:

$$\beta_0^{j-1} = \frac{j-1}{\alpha} \log \delta + \frac{v_j - v_1}{2\alpha},$$

$$\beta_1 = \frac{1}{\alpha},$$

$$\beta_2 = \frac{j-1}{\alpha}.$$

As with the interest rate specification, (2.17) contains generated regressors. Both the expected yield spread and the expected real rate must be estimated. The technique of instrumental variables is used for the estimation as well as ordinary least squares. Note that with the assumption of logarithmic utility, the expected inflation terms cancel from both sides of the equation and (2.17) can be estimated in nominal terms.

There may be an advantage in using the yield spread specification. In the expected real interest rate formulation (2.7), the intercept contains v_j. This variance is assumed to be constant but in practice it may change through time. The spread formulation (2.7) has the difference between v_j and v_1 in the intercept term. It is possible that this difference is closer to a constant than the levels. The time series behavior of the difference in the variances is investigated in the empirical section.

Two types of information available from bond prices are examined: real interest rates and yield spreads. The intertemporal consumption-based asset pricing model provides a framework whereby these financial variables can be linked to macroeconomic fluctuations. Tests are conducted in chapter 4 to determine if these variables have the explanatory power that the theory suggests.