Chapter 7

TESTING LINEAR FORECASTING MODELS

Some economic models imply that a linear function of a variable X_t is a forecaster of Y_{t+1} in the sense that

$$E(Y_{t+1}|\mathrm{I}_t) = a + bX_t, \tag{7.1}$$

where a and b are constants, and I_t is an information set. Typically, I_t is the information set available to economic agents at date t, and includes the current and past values of X_t and Y_t. Equation (7.1) is called a linear forecasting model. In some cases, a linear function of a variable X_t is a forecaster of Y_{t+s} :

$$E(Y_{t+s}|\mathrm{I}_t) = a + bX_t. \tag{7.2}$$

If Y_{t+s} is in I_{t+s}, (7.2) is a multi-period linear forecasting model. In this chapter, we discuss some standard methods to test these linear forecasting models.

7.1 Forward Exchange Rates

In usual market transactions (called *spot transactions*), transactions are carried out immediately. In forward contracts, two parties agree to carry out transactions at a specified future date. In foreign exchange forward contracts, a party agrees to deliver specified units of a currency to another party who agrees to pay a specified price.

Let $F_{t,1}$ be the forward exchange rate at date t of a foreign currency to be delivered at date $t+1$: at date t a contract is made in which $F_{t,1}$ units of the domestic currency is promised to be paid when one unit of the foreign currency is delivered at date $t+1$. Let S_t be the spot exchange rate at date t which is expressed as the price of one unit of the foreign currency in terms of the domestic currency. Assume that the domestic investors are risk neutral. For now, assume that risk neutrality is defined about gambles involving the domestic currency. Given that preferences are defined over goods rather than currencies, risk neutrality should be defined about gambles involving goods. The assumption of risk neutrality over the domestic currency leads to Siegel's (1972) Paradox as discussed below. Under this assumption,

$$F_{t,1} = E(S_{t+1}|\mathrm{I}_t) \tag{7.3}$$

should hold in equilibrium, where I_t is the information set available at date t. To see this relation suppose that $F_{t,1} > E(S_{t+1}|\mathrm{I}_t)$. Then the domestic investors' expected profit is positive when they sell the foreign currency with forward contracts. The supply of foreign currency will be infinite, and therefore equilibrium cannot be attained. If $F_{t,1} < E(S_{t+1}|\mathrm{I}_t)$, then the domestic investors' expected profit is positive when they buy the foreign currency with forward contracts.

Let $F_{t,s}$ be the forward exchange rate at date t of a foreign currency to be delivered at date $t+s$. Then with a similar argument,

$$F_{t,s} = E(S_{t+s}|\mathrm{I}_t) \tag{7.4}$$

should hold. Relation (7.4) is implied by the uncovered interest parity (UIP) if we assume the covered interest parity (CIP)[1], and Relation (7.3) is a special case of UIP when $s = 1$.

Given (7.4), a natural way to test UIP is to consider a regression

$$S_{t+s} = a + bF_{t,s} + e_t. \tag{7.5}$$

Then (7.4) implies that $E(e_t|\mathrm{I}_t) = 0$ when $a = 0$ and $b = 1$. Since $F_{t,s}$ is in I_t, if S_{t+s} and $F_{t,s}$ are stationary, then the asymptotic theory of OLS in Chapter 5 applies to this regression. In the data of exchange rates, it is often observed that the first difference of $\ln(S_{t+s})$, the first difference $\ln(F_{t,s})$, and $\ln(S_{t+s}) - \ln(F_{t,s})$ appear to be stationary. However, S_{t+s} and $F_{t,s}$ do not appear to be stationary, hence the asymptotic theory of Chapter 5 does not apply to (7.5). One solution found in the literature is to apply cointegration to (7.5) or the log version of (7.5).[2]

We consider transforming the data to obtain a regression with stationary variables. For this purpose, we first take the natural log of both sides of (7.4) to obtain an approximate relation

$$\ln(F_{t,s}) = a + E(\ln(S_{t+s})|\mathrm{I}_t), \tag{7.6}$$

where a is a constant. This relation is an approximation because the log of the expected value of a random variable is not the expected value of the log of the variable.

The approximation error for (7.6) can be significant and may lead to the rejection of the model when the exchange rate is conditionally heteroskedastic even when UIP holds. This problem may be serious because conditional heteroskedasticity

[1] CIP says that $1 + i_{t+s} = \frac{F_{t,s}}{S_t}(1 + i^*_{t,s})$ while UIP says that $1 + i_{t+s} = \frac{E(S_{t,s}|I_t)}{S_t}(1 + i^*_{t,s})$, where $i_{t,s}$ and $i^*_{t,s}$ are interest rates on domestic deposit and foreign deposit, respectively.

[2] This solution is, however, problematic for the purpose of testing UIP as discussed in Chapter 14.

is detected for most exchange rate series especially when high frequency data (e.g., weekly or daily) are used. To illustrate this point, assume that S_{t+s} is log normally distributed conditional on I_t, then $\ln(E(S_{t+s}|\mathrm{I}_t)) = E(\ln(S_{t+s})|\mathrm{I}_t) + \frac{1}{2}Var(\ln(S_{t+s})|\mathrm{I}_t)$.

Hence (7.6) is exact with $a = \frac{1}{2}Var(\ln(S_{t+s}))$ if $\ln(S_{t+s})$ is conditionally homoskedastic. However, if $\ln(S_{t+s})$ is conditionally heteroskedastic, $Var(\ln(S_{t+s})|\mathrm{I}_t)$ is not constant. Hence (7.6) is an approximation, and the approximation error is more important for data with stronger conditional heteroskedasticity effects.

Assuming that this approximation error is negligible, we consider a regression

$$\ln(S_{t+s}) - \ln(F_{t,s}) = a + \mathbf{X}_t'\mathbf{b} + e_t, \tag{7.7}$$

where $\mathbf{X}_t$ is a stationary random vector that is in I_t. Then (7.6) implies that $\mathbf{b} = \mathbf{0}$ and $E(e_t|\mathrm{I}_t) = 0$ (note that $a \neq 0$ here). Assuming that $\ln(S_{t+s}) - \ln(F_{t,s})$ is stationary, the asymptotic theory in Chapter 5 applies to (7.7). For example, $\mathbf{X}_t$ is a vector of lagged values of $\ln(S_{t+s}) - \ln(F_{t,s}) : \mathbf{X}_t = (\ln(S_t) - \ln(F_{t-s,s}), \ln(S_{t-1}) - \ln(F_{t-1-s,s}), \cdots, \ln(S_{t-k}) - \ln(F_{t-k-s,s}))'$. UIP can be tested by testing the null hypothesis $H_0 : \mathbf{b} = \mathbf{0}$.

The assumption of risk neutrality over the domestic currency leads to *Siegel's Paradox*. Assume that foreign investors are risk neutral over their currency. Then the same argument made for (7.3) for the domestic investors imply

$$\frac{1}{F_{t,1}} = E(\frac{1}{S_{t+1}}|\mathrm{I}_t). \tag{7.8}$$

Since $\frac{1}{X}$ is a convex function, (7.3) and (7.8) cannot hold at the same time. This property is known as Siegel's Paradox.[3]

[3] Because preferences are defined over goods, risk neutrality should be defined over goods. Siegel's Paradox is a result of defining risk neutrality over currencies. In order to illustrate this point, imagine

7.2 The Euler Equation

Consider an economy with a single good, in which the current and past values of a random vector $\mathbf{X}_t$ generate the information set I_t, which is available to the economic agents. The random vector $\mathbf{H}_t = [\mathbf{X}_0', \mathbf{X}_1', \cdots, \mathbf{X}_t']'$ summarizes I_t. Let $Prob(\mathbf{H}_t)$ denote the probability of $\mathbf{H}_t$. For simplicity, we assume that the economy ends at date T, and that there exist N possible values of $\mathbf{H}_T$. With this notation, $\mathbf{H}_T$ plays the role of the state of the world s in Chapter 2, and $\mathbf{H}_t$ specifies the subset in the partition of S at date t. The history notation is more convenient for the purpose of this section to ensure that consumption is in the information available at date t.

We assume that the representative consumer maximizes the lifetime utility function

(7.9) $$U = \sum_{t=0}^{T} \sum_{\mathbf{H}_t} Prob(\mathbf{H}_t)\beta^t u(C_t(\mathbf{H}_t)),$$

where β is a discount factor, $u(\cdot)$ is the utility function, and $C_t(\mathbf{H}_t)$ is the consumption at date t with history $\mathbf{H}_t$. As a bench mark case, we assume that there exists a complete set of contingent security markets at date 0. Assuming that there are N states of the world, and the contingent security for one unit of $C_t(\mathbf{H}_t)$ costs $P_t(\mathbf{H}_t)$ in terms of the good at date 0, the lifetime budget constraint is

(7.10) $$\sum_{t=0}^{T} \sum_{\mathbf{H}_t} P_t(\mathbf{H}_t)C_t(\mathbf{H}_t) = \sum_{t=0}^{T} \sum_{\mathbf{H}_t} P_t(\mathbf{H}_t)C_t^e(\mathbf{H}_t),$$

where $C_t^e(\mathbf{H}_t)$ is the endowment. Let λ be the Lagrange multiplier for the budget constraint (7.10). Then the first order conditions for the consumer's maximization

that there are two consumption goods in the world economy: a good purchased with the domestic currency, and another good purchased with the foreign currency. The real version of (7.3), expressed in terms of the domestic good, and the real version of (7.8), expressed in terms of the foreign good, are not subject to Siegel's Paradox (see, e.g. Frankel, 1979, 1980). Engel (1984) empirically tests the absence of expected real profits from forward market speculation and shows that Siegel's paradox is not empirically important in this case.

problem include

$$\beta^t Prob(\mathbf{H}_t) mu(C_t(\mathbf{H}_t)) = \lambda P_t(\mathbf{H}_t), \tag{7.11}$$

where $mu(\cdot)$ is the derivative of the utility function (marginal utility). Hence

$$\frac{\beta^{t+1} Prob(\mathbf{H}_{t+1}) mu(C_{t+1}(\mathbf{H}_{t+1}))}{\beta^t Prob(\mathbf{H}_t) mu(C_t(\mathbf{H}_t))} = \frac{P_{t+1}(\mathbf{H}_{t+1})}{P_t(\mathbf{H}_t)}, \tag{7.12}$$

which we call the *state-by-state intertemporal first order condition.* This type of condition is useful in testing for complete risk sharing as we will discuss in Chapter 17.

The first order condition (7.12) does not necessarily hold when markets are incomplete. We derive the asset pricing equation and Euler equation, which can be shown to hold for some incomplete market models, from this first order condition. For this purpose, imagine that a security pays off $D_{t+1}(\mathbf{H}_{t+1})$ units of the good at date $t+1$ when the history $\mathbf{H}_{t+1}$ is realized. Let $V_t(\mathbf{H}_t)$ be the price of the security in terms of the good at date t when the history $\mathbf{H}_t$ is realized. Then an arbitrage condition gives

$$V_t(\mathbf{H}_t) = \frac{\sum_{\mathbf{H}_{t+1}|\mathbf{H}_t} P_{t+1}(\mathbf{H}_{t+1}) D_{t+1}(\mathbf{H}_{t+1})}{P_t(\mathbf{H}_t)}, \tag{7.13}$$

where the summation in the numerator sums up all $\mathbf{H}_{t+1}$'s that follow $\mathbf{H}_t$. The numerator is the price of the security in terms of the good at date 0, and the denominator is the price of the good at date t, so that the security price is expressed in terms of the good at date t. Substituting (7.12) into (7.13) yields

$$V_t(\mathbf{H}_t) = \frac{\sum_{\mathbf{H}_{t+1}|\mathbf{H}_t} \beta \; prob(\mathbf{H}_{t+1}) mu_{t+1} D_{t+1}(\mathbf{H}_{t+1})}{prob(\mathbf{H}_t) mu_t}, \tag{7.14}$$

where mu_t denotes $mu(C_t(\mathbf{H}_t))$. Noting that $\frac{Prob(\mathbf{H}_{t+1})}{Prob(\mathbf{H}_t)}$ is the probability of $\mathbf{H}_{t+1}$

conditional on $\mathbf{H}_t$, we can rewrite (7.14) as

$$V_t = \frac{E(\beta\ mu_{t+1} D_{t+1} | \mathrm{I}_t)}{mu_t}, \tag{7.15}$$

which we call the *asset pricing equation*.

Dividing both sides of the asset pricing equation (7.15) by V_t yields

$$\frac{E(\beta\ mu_{t+1} R_{t+1} | \mathrm{I}_t)}{mu_t} = 1, \tag{7.16}$$

which is called the *Euler equation* where $R_{t+1} = \frac{D_{t+1}}{V_t}$ is the real gross asset return. It should be noted that the asset pricing equation and the Euler equation hold for any asset while the state-by-state intertemporal first order condition only holds for the contingent securities since $P_t(\mathbf{H}_t)$ in (7.12) is the price of contingent security rather than any other security.

7.3 The Martingale Model of Consumption

Consider a bond that pays one unit of the good at date $t+1$ without any uncertainty, which we call the real risk free bond. Let R^f_{t+1} be the real gross asset return on the real risk free bond. Then $R^f_{t+1} - 1$ is the real interest rate. Assume that the real interest rate is constant, and that $\beta R^f_{t+1} = 1$. Then the Euler equation (7.16) implies

$$E(mu_{t+1} | \mathrm{I}_t) = mu_t. \tag{7.17}$$

Therefore, under these assumptions, the marginal utility is a martingale adapted to I_t. This implication is testable when the intra-period utility function is parameterized, so that mu_t is related to consumption.

Hall (1978) assumes that the intra-period utility function is quadratic:

$$u(C_t) = -\alpha(C_t - \gamma)^2, \tag{7.18}$$

where α and γ are positive constants. Then $mu_t = -2\alpha(C_t - \gamma)$, and (7.17) implies

$$E(C_{t+1}|\mathrm{I}_t) = C_t. \tag{7.19}$$

Thus Euler equation implies that consumption is a martingale adapted to I_t. Therefore, this model is called the *martingale model of consumption*. With an additional assumption that consumption is conditionally homoskedastic, (7.19) implies that consumption is a random walk. For this reason, some authors prefer to call this model the *random walk model of consumption*.

This martingale (or random walk) hypothesis can be tested by applying OLS to

$$C_{t+1} - C_t = a + \mathbf{X}_t'\mathbf{b} + e_t \tag{7.20}$$

where $\mathbf{X}_t$ is a stationary random vector which is in I_t. Then (7.19) implies that $a = 0$, $\mathbf{b} = \mathbf{0}$, and $E(e_t|\mathrm{I}_t) = 0$.

7.4 The Linearized Euler Equation

It should be noted that the random walk model of consumption is derived under the assumptions of a quadratic utility function and a constant real interest rate. These assumptions are not attractive. There exists some evidence that real interest rates are not even stationary (see, Rose, 1988). The quadratic utility function has an implication that both absolute and relative risk aversion coefficients increase with consumption. The intertemporal elasticity of substitution is the reciprocal of the relative risk aversion coefficient for the time-separable expected utility function, and the quadratic utility function implies that the elasticity decreases as consumption increases. These implications are counterintuitive to most people upon introspection, and there is empirical evidence against them (see Chapter 17????????).

Masao needs to check this!

Most researchers agree that the isoelastic utility function,

$$u(C) = \frac{1}{1-\alpha}[C^{1-\alpha} - 1]$$

is more reasonable than the quadratic utility function. For this utility function, the relative risk aversion coefficient is α (a constant), and the absolute risk aversion coefficient decreases with consumption. The intertemporal elasticity of substitution is $\frac{1}{\alpha}$. With this utility function, $mu_t = C_t^{-\alpha}$, and (7.16) implies

$$E(\beta R_{t+1} C_{t+1}^{-\alpha} | \mathrm{I}_t) = C_t^{-\alpha}. \tag{7.21}$$

With an assumption that R_t and C_t are jointly log normally distributed conditional on I_t, we obtain

$$E(\ln(R_{t+1}) - \alpha \ln(C_{t+1}) | \mathrm{I}_t) = -\ln(\beta) - \frac{1}{2} Var(\ln(R_{t+1} C_{t+1}^{-\alpha}) | \mathrm{I}_t) - \alpha \ln(C_t). \tag{7.22}$$

Further assuming that $\ln(R_{t+1} C_{t+1}^{-\alpha})$ is conditionally homoskedastic with respect to I_t, we obtain the linearized version of the Euler equation (7.21):

$$E(\ln(R_{t+1}) - \alpha \ln(C_{t+1}) | \mathrm{I}_t) = b - \alpha \ln(C_t), \tag{7.23}$$

where $b = -\ln(\beta) - \frac{1}{2} Var(\ln(R_{t+1} C_{t+1}^{-\alpha}) | \mathrm{I}_t)$ is a constant. Note that the linearized Euler equation (7.23) holds for any asset return under the stated assumptions.

With an additional assumption that the real interest rate is constant as in Section 7.3, we can obtain a result similar to the random walk hypothesis. In this case, (7.23) implies

$$E(\ln(C_{t+1}) | \mathrm{I}_t) = c + \ln(C_t), \tag{7.24}$$

where $c = -\frac{b}{\alpha} + \frac{1}{\alpha} \ln(R_{t+1})$. As in the previous section, we can test this model by applying OLS to

$$\ln(C_{t+1}) - \ln(C_t) = c + \mathbf{X}_t' \mathbf{b} + e_t \tag{7.25}$$

where $\mathbf{X}_t$ is a stationary random vector which is in I_t. Equation (7.24) implies that $\mathbf{b} = \mathbf{0}$, and $E(e_t|\mathrm{I}_t) = 0$.

The linearized Euler equation (7.23) has been used by many researchers without the additional assumption of the constant real interest rate. Hansen and Singleton (1983) apply the maximum likelihood estimation method to (7.23). Hall (1988) estimates the intertemporal elasticity of substitution from

$$\ln(C_{t+1}) - \ln(C_t) = d + \frac{1}{\alpha}\ln(R_{t+1}) + e_t. \tag{7.26}$$

Equation (7.23) implies that $d = -\frac{b}{\alpha}$ and $E(e_t|\mathrm{I}_t) = 0$. Since $\ln(R_{t+1})$ is not in I_t, OLS cannot be applied to (7.26). Any stationary variable in I_t, however, can be used as an instrumental variable for (7.26). Hansen and Singleton (1996) also apply an IV method to (7.23).

7.5 Optimal Taxation

The method to derive the martingale property of consumption can be applied to other optimization problems. A good example is the optimal taxation model tested by Barro (1981), Sahasakul (1986), Kingston (1984), Mankiw (1987), and Bizer and Durlauf (1990) among others.

Assume that the government minimizes the following quadratic cost function at date t:

$$E_t \sum_{j=0}^{\infty} \beta^j (c_0 \tau_{t+j} + \frac{c_1}{2}\tau_{t+j}^2), \tag{7.27}$$

subject to the budget constraint

$$B_{t+1} = R[B_t + g_t - \tau_t], \quad B_t \text{ bounded for all } t \tag{7.28}$$

by choosing $\{\tau_{t+j}, B_{t+j}\}_{j=0}^{\infty}$. Here $\{g_t\}_{t=j}^{\infty}$ is a stochastic process describing the ratio of government spending to GDP, τ_t is the tax collected as a percentage of GDP, B_t is the real value of a one-period risk free bond to be repaid at date $t+1$ as a percentage of GDP, and R is the gross real interest rate, which is assumed to be constant. We assume that $\beta R = 1$.

The Euler equation for the maximization problem is

$$E(\tau_{t+1}|\mathrm{I}_t) = \tau_t. \tag{7.29}$$

As in the consumption case, this martingale hypothesis can be tested by applying OLS to

$$\tau_{t+1} - \tau_t = a + \mathbf{X}_t'\mathbf{b} + e_t, \tag{7.30}$$

where $\mathbf{X}_t$ is a stationary random vector which is in I_t. Then (7.29) implies that $a = 0$, $\mathbf{b} = \mathbf{0}$, and $E(e_t|\mathrm{I}_t) = 0$. Barro (1981), Kingston (1984), and Mankiw (1987) have found that movements of U.S. tax rates over time are roughly consistent with the martingale hypothesis. On the other hand, Sahasakul (1986) reports that U.S. tax rates are predictably related to wars and recessions, which is evidence against the martingale hypothesis. Bizer and Durlauf (1990) report evidence against the hypothesis based on a frequency- domain based test (see Section 16.3??????? below).

Masao needs to check this!

7.6 Tests of Forecast Accuracy

Tests of forecast accuracy can be used to test economic models. A prominent example is tests for exchange rate models in the literature that started by Meese and Rogoff (1983) who compared predictions of exchange rate models with predictions of the random walk model.

7.6.1 The Monetary Model of Exchange Rates

A multi-period forecasting formulation in Mark (1995) is motivated by the the monetary model of the monetary model of Frenkel (1976), Mussa (1976), and Bilson (1978). The monetary model implies the present value relationship

$$s_t = (1 - \beta)E(\sum_{i=1}^{\infty} \beta^i f_{t+i} | \mathrm{I}_t). \tag{7.31}$$

where s_t is the log exchange rate, $f_t = m_t - m_t^* - \gamma(y_t - y_t^*)$ where m_t is the log domestic money supply, y_t is the log domestic income. We call f_t fundamentals. Here, γ is the income elasticity of money demand, $\beta = \alpha/(1 + \alpha)$ where α is the interest semi-elasticity of money demand. If f_t is a driftless random walk, then the present value relationship implies $s_t = f_t$, and the log exchange rate is a random walk. However, deviations from the log exchange rate from the fundamentals are known to be persistent. These considerations motivated Mark to investigate the projection of the k-period-ahead change in the log exchange rate on its current deviation from the fundamental value

$$s_{t+k} - s_t = a + bX_t + e_1 t, \tag{7.32}$$

where $X_t = f_t - s_t$ and $e_1 t$ is a forecast error. On the other hand, if the log exchange rate is a driftless random walk, $a = b = 0$, then

$$s_{t+k} - s_t = e_2 t, \tag{7.33}$$

where $e_2 t$ be the forecast error of the random walk model. Tests described in this section compare forecast accuracy based on the differences in mean squared prediction errors (MSPEs) from these two models. Unlike the previous works that had found that the economic model does not improve forecast accuracy over the random walk model, Mark (1995) found evidence in favor of the economic model for long-horizon changes (large values of k). However, Kilian (1999) pointed out problems with Mark's (1995) bootstrap procedure. With a corrected bootstrap procedure, Kilian found no evidence of increased long-horizon predictability. In panel data that combine time series of 19 industrialized countries, Mark and Sul (2001) found evidence of increased long-horizon predictability. Thus the evidence is mixed for the monetary model compared with the random walk model.

Engel and West (2005) showed analytically that in present value models such as Equation (7.31), the log exchange rate manifests near-random walk behavior if the first difference of fundamentals is stationary and β is near one. Their result helps explain that fundamentals provide little help in predicting changes in the log exchange rates.

7.7 The Taylor Rule Model of Exchange Rates

Many recent papers have explored various aspects of exchange rate models with the Taylor Rule (see, e.g., Mark, 2005; Engel and West, 2005, 2006; Clarida and Waldman, 2008; Kim and Ogaki, 2009). Molodtsova, Nikolsko-Rzhevskyy, and Papell (2008) and Molodtsova and Papell (2009) find strong evidence of exchange rate predictability using the Taylor rule model. In this model, under the assumption that uncovered interest parity holds, exchange rate movements are related to the differential of short-term nominal interest rates between two countries. In each country, the nominal interest rate is in turn set by the central bank that follows a policy rule proposed by Taylor (1993). According to Taylor's original specification, the home central bank adjusts the nominal interest rate in response to changes in the domestic inflation and output gap:

$$\tilde{i}_t = \pi_t + \phi(\pi_t - \tilde{\pi}) + \gamma y_t + \tilde{r},$$

where $\tilde{i}_t$ is the target rate of the short-term nominal interest rate, π_t is inflation, $\tilde{\pi}$ is the inflation target, y_t is the output gap, and $\tilde{r}$ is the equilibrium level of the real interest rate. The Taylor rule for small open economies may include the real exchange rate s_t (Clarida, Galí, and Gertler, 1998):

$$\tilde{i}_t = \pi_t + \phi(\pi_t - \tilde{\pi}) + \gamma y_t + \delta s_t + \tilde{r}. \tag{7.34}$$

Empirical studies (e.g., Clarida, Galí, and Gertler, 1998, 2000) find that central banks engage in interest rate smoothing so that the observed nominal rate i_t is a partial adjustment of its lagged value and the target rate:

$$i_t = (1 - \rho)\tilde{i}_t + \rho i_{t-1} + v_t. \tag{7.35}$$

Suppose the foreign central bank follows an analogous policy rule:

$$i_t^* = (1 - \rho^*)\tilde{i}_t^* + \rho^* i_{t-1}^* + v_t^*. \tag{7.36}$$

Taking the difference between the policy reaction function of home country (7.35) and that of foreign country (7.36) yields the interest rate differential:

$$i_t - i_t^* = \beta + \beta_\pi \pi_t - \beta_\pi^* \pi_t^* + \beta_y y_t - \beta_y^* y_t^* + \beta_s s_t + \beta_s^* s_t + \rho i_{t-1} - \rho^* i_{t-1}^* + \eta_t, \tag{7.37}$$

where $\eta_t = v_t - v_t^*$, $\beta = (\tilde{r} - \phi\tilde{\pi})(1 - \rho)$, $\beta_\pi = (1 + \phi)(1 - \rho)$, $\beta_y = \gamma(1 - \rho)$, and $\beta_s = \delta(1 - \rho)$. Analogous definitions apply for foreign coefficients denoted by a star. Note that since $s_t = -s_t^*$, we have $\beta_s s_t + \beta_s^* s_t$.

Assume that uncovered interest rate parity holds: $\Delta e_{t+1} = i_t - i_t^*$ where e_t is the log of the nominal exchange rate defined as the domestic currency price of foreign currency. Equating the UIP condition with the interest rate differential (7.37) yields the Taylor rule model of exchange rates:

$$(7.38)\quad \Delta e_{t+1} = \beta + \beta_\pi \pi_t - \beta_\pi^* \pi_t^* + \beta_y y_t - \beta_y^* y_t^* + \beta_s s_t + \beta_s^* s_t + \rho i_{t-1} - \rho^* i_{t-1}^* + \eta_t.$$

Molodtsova and Papell (2009) evaluate out-of-sample forecasts of one-month-ahead exchange rate movements using the Taylor-rule model for the monthly U.S. exchange rates against 12 OECD countries. The data spans from March 1973 to June 2006 (December 1998 for the European Monetary Union countries). The predictive performance of the model is evaluated using the CW test statistics for the null hypothesis that the exchange rate follows a random walk against the alternative hypothesis that it is predictable by the model (7.38).

They find that the Taylor rule model exhibits strong evidence of short-term exchange rate predictability, especially when the real exchange rates are excluded from equation (7.38). For that specification, the model outperforms the random walk for 10 out of 12 currencies at the 10% significance level - four of them at the 1% level and additional six at the 5% level - using one of the three output gap specifications they consider (the linear trend, the quadratic trend, and the HP-filter). By contrast, using the same dataset and the inference method, they find much less evidence of exchange rate predictability with conventional models of exchange rates (the UIP model of Clark and West, 2006; the monetary model of Mark, 1995; and the PPP model of Mark and Sul, 2001). Even after combining the results from these three models, they find statistically significant evidence of exchange rate predictability at the 5% level for only 3 of the 12 currencies and for an additional currency at the 10%

level.

7.7.1 Diebold and Mariano (1995)

One of the commonly used methods for testing forecast accuracy is the test of equal accuracy proposed by Diebold and Mariano (1995, the DM test).

Consider two competing forecast series y_{1t} and y_{2t} of the time series y_t, with associated forecast errors e_{1t} and e_{2t}, $t = 1, \cdots, T$, respectively. The DM test is applicable to a wide variety of accuracy measures. Here, as in many applications, we compare forecast accuracy based on the differences in mean squared prediction errors (MSPEs) from the two series. In this case, the DM test evaluates the null hypothesis that the population mean of the MSPE differences is 0, $E(e_{1t}^2 - e_{2t}^2) = 0$, against the alternative hypothesis $E(e_{1t}^2 - e_{2t}^2) \neq 0$.

Let $\bar{d}$ denote the sample mean of the MSPE differential:

$$\bar{d} = \frac{1}{T}\sum_{t=1}^{T}(e_{1t}^2 - e_{2t}^2).$$

If the MSPE differential is covariance stationary and short memory, then $\sqrt{T}\bar{d}$ is asymptotically normally distributed with mean zero and variance $2\pi f_d(0)$ where $f_d(0) = \frac{1}{2\pi}\sum_{\tau=-\infty}^{\infty}\gamma_d(\tau)$ is the spectral density of the MSPE differential at frequency 0, and $\gamma_d(\tau) = E[(e_{1t}^2 - e_{2t}^2)(e_{1t-\tau}^2 - e_{2t-\tau}^2)]$ is the autocovariance of the MSPE differential. The DM statistic is given by

$$DM = \frac{\bar{d}}{\sqrt{\frac{2\pi\hat{f}_d(0)}{T}}},$$

where $2\pi\hat{f}_d(0)$ is a consistent estimator of $2\pi f_d(0)$. It is obtained by a weighted sum of the sample autocovariances,

$$2\pi\hat{f}_d(0) = \sum_{\tau=-(T-1)}^{(T-1)} 1\left(\frac{\tau}{S(T)}\right)\hat{\gamma}_d(\tau),$$

where $1(\tau/S(T))$ is the lag window, $S(T)$ is the truncation lag, and $\hat{\gamma}_d(\tau) = \frac{1}{T}\sum_{t=|\tau|+1}^{T}(d_t - \bar{d})(d_{t-|\tau|} - \bar{d})$ with $d_t \equiv e_{1t}^2 - e_{2t}^2$. Diebold and Mariano (1995) suggest the use of the uniform lag window,

$$1\left(\frac{\tau}{S(T)}\right) = \begin{cases} 1 & \text{for } \left|\frac{\tau}{S(T)}\right| \le 1, \\ 0 & \text{otherwise} \end{cases}$$

and the truncating lag $S(T) = (k-1)$ since optimal k-step-ahead forecast errors are at most $(k-1)$-dependent.

7.7.2 Clark and West (2006) and Clark and West (2007)

Now suppose we wish to compare out-of-sample forecast accuracy of two nested models. One example as in the exchange rate forecasting literature above is the case where a linear econometric model (model 2) is compared to a random walk model (model 1):

$$\text{Model 1:} \quad y_t = e_t$$

$$\text{Model 2:} \quad y_t = \mathbf{X}_t'\boldsymbol{\beta} + e_t,$$

where e_t in both models is a zero mean martingale difference which may be conditionally heteroskedastic.

Let $T+1$ be the sample size of y_t and $\mathbf{X}_t$ which is divided into two subsamples $T+1 = R+P$. For illustration, suppose we are comparing one-period-ahead forecasts, y_{t+1}.[4] Model 2 is estimated using data prior to t to generate P predictions for y_{t+1}, $t = R, R+1, \cdots T$. The out-of-sample MSPEs of the two models are,

$$\text{Model 1:} \quad \hat{\sigma}_1^2 \equiv P^{-1}\sum_{t=T-P+1}^{T} y_{t+1}^2,$$

$$\text{Model 2:} \quad \hat{\sigma}_2^2 \equiv P^{-1}\sum_{t=T-P+1}^{T} (y_{t+1} - \mathbf{X}_{t+1}'\hat{\boldsymbol{\beta}}_t)^2.$$

[4]For multi-horizon predictions, see Clark and West (2006).

Recall that the DM test is based on the assumption that the difference in sample MSPEs from two models is asymptotically zero. However, Clark and West (2006) and Clark and West (2007) show that this is not the case when the two models are nested. To see this, write:

$$\hat{\sigma}_1^2 - \hat{\sigma}_2^2 = 2\left(P^{-1}\sum_{t=T-P+1}^{T} y_{t+1}\mathbf{X}_{t+1}'\hat{\boldsymbol{\beta}}_t\right) - \left[P^{-1}\sum_{t=T-P+1}^{T}(\mathbf{X}_{t+1}'\hat{\boldsymbol{\beta}}_t)^2\right]. \tag{7.39}$$

Under the null hypothesis of equal predictive accuracy, y_t follows a martingale difference ($\boldsymbol{\beta} = \mathbf{0}$) as in model 1. Therefore, $y_{t+1} = e_{t+1}$ and $Ee_{t+1}\mathbf{X}_{t+1}'\hat{\boldsymbol{\beta}}_t = 0$, and thus the first term in equation (16.11) is expected to be approximately zero. However, the second term is $-P^{-1}\sum_{t=T-P+1}^{T}(\mathbf{X}_{t+1}'\hat{\boldsymbol{\beta}}_t)^2 < 0$, and thus the MSPE from model 2 is expected to be greater than that of model 1:

$$\hat{\sigma}_1^2 - \hat{\sigma}_2^2 \xrightarrow{p} -E(\mathbf{X}_{t+1}'\hat{\boldsymbol{\beta}}_t)^2 < 0.$$

The DM statistics, while appropriate for non-nested models, do not adjust for this shift, and result in non-normal test statistics when the models are nested. Therefore, hypothesis tests based on standard normal critical values are usually poorly sized, failing to reject the null hypothesis when it should (McCracken, 2004; Clark and McCracken, 2001, 2005). This is particularly problematic for tests of our-of-sample predictability of financial data for which the null hypothesis is a random walk.

Clark and West (2006) and Clark and West (2007) propose an asymptotically normal test for two nested models that properly adjusts the difference in MSPEs by a consistent estimate of $E(\mathbf{X}_{t+1}'\hat{\boldsymbol{\beta}}_t)^2$. This test is applicable when $\boldsymbol{\beta}_t$ is estimated from rolling regressions using data from $t - R + 1$ to t.[5]

[5]Clark and West (2007) consider a general parametric specification of the null model (model 1) that is smaller than the alternative model (model 2). Thus, model 2 reduces to model 1 if some of the parameters in model 2 are zero.

The bias-adjusted difference of the sample mean MSPEs is given by,

$$\begin{aligned}\bar{f} &\equiv \hat{\sigma}_1^2 - \left[\hat{\sigma}_2^2 - P^{-1}\sum_{t=T-P+1}^{T}(\mathbf{X}'_{t+1}\hat{\boldsymbol{\beta}}_t)^2\right]. \\ &= P^{-1}\sum_{t=T-P+1}^{T}\hat{f}_{t+1}.\end{aligned}$$

where $\hat{f}_{t+1} \equiv y_{t+1}^2 - [(y_{t+1} - \mathbf{X}'_{t+1}\hat{\boldsymbol{\beta}}_t)^2 - (\mathbf{X}'_{t+1}\hat{\boldsymbol{\beta}}_t)^2]$. Under some mild conditions, $\sqrt{P}\bar{f}$ is asymptotically normally distributed with mean zero and variance $V \equiv 4E(y_{t+1}\mathbf{X}'_{t+1}\hat{\boldsymbol{\beta}}_t)^2$. The adjusted test statistic is

$$CW = \frac{\bar{f}}{\sqrt{\frac{\hat{V}}{P}}},$$

where $\hat{V} \equiv 4P^{-1}\sum_{t=T-P+1}^{T}(y_{t+1}\mathbf{X}'_{t+1}\hat{\boldsymbol{\beta}}_t)^2 = P^{-1}\sum_{t=T-P+1}^{T}(\hat{f}_{t+1} - \bar{f})^2$ is a consistent estimator of V. Clark and West (2006) present simulation results showing that inferences of the CW statistics using normal critical values are properly sized. Note that the alternative hypothesis of this test is that y_t is linearly predictable ($\boldsymbol{\beta} \neq \mathbf{0}$) as in model 2, implying that the population MSPE of model 2 is smaller than that of model 1. Therefore, this test is one-sided, and the null hypothesis is rejected when the CW test statistic is significantly positive.

References

BARRO, R. J. (1981): "On the Predictability of Tax-Rate Changes," NBER Working Paper No. 636.

BILSON, J. F. O. (1978): "Rational Expectations and the Exchange Rate," in *The Economics of Exchange Rates: Selected Studies*, ed. by J. A. Frenkel, and H. G. Johnson, pp. 75–96. Addison-Wesley, Reading, MA.

BIZER, D. S., AND S. N. DURLAUF (1990): "Testing the Positive Theory of Government Finance," *Journal of Monetary Economics*, 26, 123–141.

CLARIDA, R. H., J. GALÍ, AND M. GERTLER (1998): "Monetary Policy Rules in Practice: Some International Evidence," *European Economic Review*, 42(6), 1033–1067.

——— (2000): "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics*, 115(1), 147–180.

Clarida, R. H., and D. Waldman (2008): "Is Bad News About Inflation Good News for the Exchange Rate? And, If So, Can That Tell Us Anything about the Conduct of Monetary Policy?," in *Asset Prices and Monetary Policy*, ed. by J. Y. Campbell, pp. 371–396. The University of Chicago Press.

Clark, T. E., and M. W. McCracken (2001): "Tests of Equal Forecast Accuracy and Encompassing for Nested Models," *Journal of Econometrics*, 105(1), 85–110.

——— (2005): "Evaluating Direct Multistep Forecasts," *Econometric Reviews*, 24(4), 369–404.

Clark, T. E., and K. D. West (2006): "Using Out-of-Sample Mean Squared Prediction Errors to Test the Martingale Difference Hypothesis," *Journal of Econometrics*, 135(1–2), 155–186.

——— (2007): "Approximately Normal Tests for Equal Predictive Accuracy in Nested Models," *Journal of Econometrics*, 138(1), 291–311.

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

Engel, C., and K. D. West (2005): "Exchange Rates and Fundamentals," *Journal of Political Economy*, 113(3), 485–517.

——— (2006): "Taylor Rules and the Deutschmark-Dollar Real Exchange Rate," *Journal of Money, Credit, and Banking*, 38(5), 1175–1194.

Engel, C. M. (1984): "Testing for the Absence of Expected Real Profits from Forward Market Speculation," *Journal of International Economics*, 17, 299–308.

Frankel, J. A. (1979): "The Diversifiability of Exchange Risk," *Journal of International Economics*, 9, 379–393.

——— (1980): "Tests of Rational Expectations in the Forward Exchange Market," *Southern Economic Journal*, 46, 1083–1101.

Frenkel, J. A. (1976): "A Monetary Approach to the Exchange Rate: Doctrinal Aspects and Empirical Evidence," *Scandinavian Journal of Economics*, 78(2), 200–224.

Hall, R. E. (1978): "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy*, 86, 971–987.

——— (1988): "Intertemporal Substitution in Consumption," *Journal of Political Economy*, 96, 339–357.

Hansen, L. P., and K. J. Singleton (1983): "Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns," *Journal of Political Economy*, 91, 249–265.

——— (1996): "Efficient Estimation of Linear Asset-Pricing Models with Moving Average Errors," *Journal of Business and Economic Statistics*, 14, 53–68.

Kilian, L. (1999): "Exchange Rates and Monetary Fundamentals: What Do We Learn from Long-Horizon Regressions?," *Journal of Applied Econometrics*, 14(5), 491–510.

Kim, H., and M. Ogaki (2009): "Purchasing Power Parity and the Taylor Rule," Working Paper No. 09-03, Department of Economics, Ohio State University.

Kingston, G. H. (1984): "Efficient Timing of Income Taxes," *Journal of Public Economics*, 24(2), 271–280.

Mankiw, N. G. (1987): "The Optimal Collection of Seigniorage: Theory and Evidence," *Journal of Monetary Economics*, 20(2), 327–341.

Mark, N. C. (1995): "Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability," *American Economic Review*, 85, 201–218.

——— (2005): "Changing Monetary Policy Rules, Learning, and Real Exchange Rate Dynamics," NBER Working Paper No. 11061.

Mark, N. C., and D. Sul (2001): "Nominal Exchange Rates and Monetary Fundamentals: Evidence from a Small Post-Bretton Woods Panel," *Journal of International Economics*, 53(1), 29–52.

McCracken, M. W. (2004): "Parameter Estimation and Tests of Equal Forecast Accuracy between Non-nested Models," *International Journal of Forecasting*, 20(3), 503–514.

Meese, R. A., and K. Rogoff (1983): "Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?," *Journal of International Economics*, 14(1–2), 3–24.

Molodtsova, T., A. Nikolsko-Rzhevskyy, and D. H. Papell (2008): "Taylor Rules with Real-Time Data: A Tale of Two Countries and One Exchange Rate," *Journal of Monetary Economics*, 55(S1), S63–S79.

Molodtsova, T., and D. H. Papell (2009): "Out-of-Sample Exchange Rate Predictability with Taylor Rule Fundamentals," *Journal of International Economics*, 77(2), 167–180.

Mussa, M. (1976): "The Exchange Rate, the Balance of Payments, and Monetary and Fiscal Policy under a Regime of Controlled Floating," *Scandinavian Journal of Economics*, 78(2), 229–248.

Rose, A. K. (1988): "Is the Real Interest Rate Stable?," *Journal of Finance*, 43(5), 1095–1112.

Sahasakul, C. (1986): "The U.S. Evidence on Optimal Taxation over Time," *Journal of Monetary Economics*, 18(3), 251–275.

Siegel, J. J. (1972): "Risk, Interest Rates and the Forward Exchange," *Quarterly Journal of Economics*, 86(2), 303–309.

Taylor, J. B. (1993): "Discretion versus Policy Rules in Practice," *Carnegie-Rochester Conference Series on Public Policy*, 39, 195–214.