KEIO/KYOTO JOINT
GLOBAL CENTER OF EXCELLENCE PROGRAM
Raising Market Quality-Integrated Design of "Market Infrastructure"

KEIO/KYOTO GLOBAL COE DISCUSSION PAPER SERIES

DP2010-007

The GEL Estimates Resolve the Risk-free Rate Puzzle in Japan

Mikio Ito*
Akihiko Noda**

Abstract

We show the nonexistence of the well-known risk-free rate puzzle in the Japanese financial markets. This result crucially depends on the accurate estimates of the two basic parameters: the subjective discount factor and the degree of risk aversion, appearing in the standard consumption-based capital asset pricing model (CCAPM). We estimate these parameters by the recently developed method, generalized empirical likelihood (GEL) estimation; we also confirm our results by comparing mean squared errors (MSEs) based on higher order biases and first order asymptotic variances of the estimates.

*Mikio Ito
Faculty of Economics, Keio University
**Akihiko Noda
Keio Advanced Research Centers, Keio University

KEIO/KYOTO JOINT GLOBAL COE PROGRAM

Raising Market Quality-Integrated Design of "Market Infrastructure"

Graduate School of Economics and Graduate School of Business and Commerce,
Keio University
2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan

Institute of Economic Research,
Kyoto University
Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501, Japan

The GEL Estimates Resolve the Risk-free Rate Puzzle in Japan *

Mikio Ito† and Akihiko Noda‡

First Draft: January 27, 2010
This Draft: February 22, 2011

Abstract

We show the nonexistence of the well-known risk-free rate puzzle in the Japanese financial markets. This result crucially depends on the accurate estimates of the two basic parameters: the subjective discount factor and the degree of risk aversion, appearing in the standard consumption-based capital asset pricing model (CCAPM). We estimate these parameters by the recently developed method, generalized empirical likelihood (GEL) estimation; we also confirm our results by comparing mean squared errors (MSEs) based on higher order biases and first order asymptotic variances of the estimates.

JEL classification numbers: G12; E21; C13
Keywords: Risk-free Rate Puzzle; GEL; GMM; Higher Order Biases; CCAPM

1 Introduction

This paper shows the nonexistence of the well-known risk-free rate puzzle in the Japanese financial markets, unlike those in the U.S. The result crucially depends on our accurate estimates of the two basic parameters: the discount factor and the degree of risk aversion, appearing in the standard consumption-based capital asset pricing model (CCAPM). Instead of using the generalized moment method (GMM, Hansen (1982)), we estimate these parameters by a recently developed method, generalized empirical likelihood (GEL, Smith (1997) and Newey and Smith (2004)) estimation. Our empirical results show that estimates of the parameters of CCAPM strongly depend on the method chosen, GMM or GEL, and on which instruments are selected. Furthermore, the GEL estimates provide better small sample properties, consistent with the nonexistence of the risk-free rate puzzle in Japan.

*We would like to thank Kohtaro Hitomi, Taisuke Otsu and Naoya Sueishi for their helpful comments and suggestions. We also would like to thank Colin McKenzie and Kosuke Oya for helpful discussions. All data and programs used for this paper are available on request.

†Faculty of Economics, Keio University, 2-15-45 Mita, Minato-ku, Tokyo, 108-8345, Japan (E-mail: ito@econ.keio.ac.jp)

‡Corresponding Author: Keio Advanced Research Centers, Keio University, 2-15-45 Mita, Minato-ku, Tokyo, 108-8345, Japan (E-mail: noda@2007.jukuin.keio.ac.jp).

Initially, Mehra and Prescott (1985) point out a puzzle, the inability of standard intertemporal economic models, such as CCAPMs, to rationalize the statistics that have characterized the U.S. financial markets over the past century. Specifically, they show that the models fail to explain the difference between the average returns of risky and safe assets in U.S. financial markets. This puzzle, called the equity premium puzzle, comes from an equation describing the intertemporal rational behavior of participants in financial markets; we can easily check the puzzle with several statistics calculated from financial data and with estimates of the discount factor and the degree of risk aversion. Inspired by this puzzle, Weil (1989) points out another puzzle, a variant of the equity premium puzzle. In turn, economists confront the inability of the models to explain the average return of safe assets. These puzzles are still puzzles for the U.S. and other industrialized countries, including Japan (see Kocherlakota (1996), Mehra and Prescott (2003), and Nakano and Saito (1998)).

In order to resolve the discrepancy between model prediction and empirical data, some economists modify their theoretical models by introducing additional settings, such as habit formation, imperfect market, or trading cost, while others try different types of consumer preferences, such as the Kreps–Porteus utility function. Some researchers focus on the data used in their empirical work. In fact, there are preceding studies making the same assertion as ours. However, the authors of the studies introduce additional assumptions or radically modify the standard CCAPM in order to produce convenient estimates: Bakshi and Naka (1997) use an asset pricing model with habit formation, Maki and Sonoda (2002) consider a trading cost and Basu and Wada (2006) estimate CCAPM considering the international risk sharing between the U.S. and Japan.

While attention has been paid to methods of estimating the parameters in CCAPM, few economists have been interested in alternative methods to estimate the parameters in the underlying asset pricing model. Since Hansen and Singleton (1982), GMM has been used in estimating the basic parameters in CCAPM with moment restrictions. GMM provides a general framework enabling us to handle semi-parametric models specified by moment restrictions, such as the Euler equations under uncertainty in CCAPM. Furthermore, GMM can be understood as a generalization of conventional methods of estimation, such as ordinary least squares, instrumental variables, and maximum likelihood; however, it is more flexible than these estimators in that it only requires a few assumptions regarding moment restrictions.

Reported drawbacks to applying GMM to estimate the parameters of CCAPM include the problem of weak identification and that of too many moment conditions. We can understand both of these problems in the context of the small sample property of GMM: the GMM estimates have non-negligible small sample bias if the appropriate instruments are not chosen, and too many moment conditions could fail to extract information from the available data. As Stock and Wright (2000) report, the analysis of conventional GMM procedures of CCAPM and linear instrumental variable regression breaks down when some or all of the parameters are weakly identified (see also Stock et al. (2002) for details). Therefore, more and more econometricians who are interested in CCAPM have given up applying GMM to macroeconomic time series data in the U.S., although some economists still use GMM to test the performance of CCAPM for time series data in several European countries (for examples, see Hyde et al. (2005) and Engsted and Møller (2010), who report that CCAPMs have some explanatory power for European financial markets).

In an effort to improve the poor performance of GMM in the case of small samples, a number of alternative estimators have been suggested. A class of GEL estimators is attracting the attention of many econometricians because of their better performance than that of GMM. All of the members of the class of GEL estimators have the same asymptotic distributions as GMM estimators; Newey and Smith (2004) find the theoretical advantage of GEL estimators by comparing the higher order asymptotic bias of the estimators of GEL and GMM. Some economists have reported advantages of such estimators. For instance, Noda and Sugiyama (2010) use continuous updating GMM (CU-GMM, Hansen et al. (1996)) in place of two-step GMM (2S-GMM) to estimate the parameters of CCAPM; they find that the estimating procedure successfully identifies the parameters of CCAPM for the Japanese data. When Yogo (2008) estimates the parameters in an asset pricing model with habit formation, he uses the same procedure and reports its validity as Noda and Sugiyama (2010).

Following their examples, in place of GMM, we attempt to use estimators in the GEL to estimate the parameters of the standard CCAPM. We compute higher order biases empirically by using the method of Newey and Smith (2004) and the mean squared error (MSE) suggested by Donald and Newey (2001), the sum of the square of the second order bias and the first order asymptotic variance, to compare the accuracy of estimates using GMM and those using GEL. We consider MSE as a measure of total reliability of estimates of a moment restriction model; Donald and Newey (2001) propose MSE for selecting appropriate instruments. Our results suggest that the CU-GMM and GEL estimators, the continuous updating estimator (CUE, Newey and Smith (2004)), the empirical likelihood (EL, Owen (1988), Qin and Lawless (1994), and Imbens (1997)), and the exponential tilting (ET, Kitamura and Stutzer (1997) and Imbens et al. (1998)) perform better than the 2S-GMM estimators, as Newey and Smith (2004) theoretically point out. The higher order biases of the GMM estimates depend on the preliminary choice of estimates, whereas those of GEL do not. We also find that CUE, belonging to the GMM estimators, successfully identifies the model parameters when we apply CUE to Japanese macroeconomic time series data, as Noda and Sugiyama (2010) show, while its empirically estimated MSE is not always small. Finally, when we substitute the estimates that our reliability measure suggests are the best, the EL estimates, into the formula of Kandel and Stambaugh (1991), we conclude the nonexistence of the risk-free rate puzzle in the Japanese financial markets.

Section 2 presents a review of the standard CCAPM and the GMM and GEL estimators that we compare. We include a brief discussion of higher order bias of GMM and GEL estimators following Newey and Smith (2004) and Anatolyev (2005). Section 3 describes the data used. Section 4 shows empirical results: (1) possibility for overcoming the difficulty concerning small sample biases in estimating CCAPM, and (2) the nonexistence of the well-known risk-free rate puzzle in the Japanese financial markets. Section 5 is for concluding remarks.

2 Model and Empirical Method

In this section, we present the standard CCAPM and empirical methods to estimate the parameters.

2.1 CCAPM

We assume that the representative consumer at time 0 chooses his/her life-time consumption and holding of several assets to maximize his/her expected utility subject to the budget constraint. The utility function is given by

$$\text{Max } E_0 \left[\sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\gamma} - 1}{1-\gamma} \right], \quad 0 < \beta < 1, \quad 0 < \gamma, \tag{1}$$

$$\text{s.t. } C_t + \sum_{i=1}^{N} p_{i,t} A_{i,t} = \sum_{i=1}^{N} [p_{i,t} + d_{i,t}] A_{i,t-1} + Y_t, \quad i = 1, 2, \cdots, N, \tag{2}$$

where the subscript t indicates time, C_t is real per capita consumption, $p_{i,t}$ is the price of the ith asset, $d_{i,t}$ is the dividend of the ith asset, $A_{i,t}$ is the amount of the per capita holdings of the ith asset, Y_t is real per capita labor income, N is the number of assets, β is the subjective time discount factor, γ is the relative risk aversion (RRA) and $E_t[\cdot]$ is the expectation operator conditional on the information available at time t. In equation (1), we assume that the utility function is of the constant relative risk aversion (CRRA) class.

Solving the above utility maximization problem, we can derive the following Euler equations:

$$E_t \left[\beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} (1 + r_{i,t+1}) - 1 \right] = 0, \quad i = 1, 2, \cdots, N, \tag{3}$$

where $r_{i,t+1}$ is the real return of the asset at time $t+1$, which is defined as

$$r_{i,t+1} = \frac{p_{i,t+1} + d_{i,t+1}}{p_{i,t}} - 1, \quad i = 1, 2, \cdots, N. \tag{4}$$

At this stage, we do not assume any data generating process (DGP) for the C_t's and $r_{i,t}$'s. Later in this section, we will show how to estimate the parameters in this Euler equation using the conditional expectation operator and how to conduct statistical inferences using estimates of these parameters.

2.2 Moment Restriction Model

We present a framework, called the moment restriction model, that allowing us to cope generally with statistical models where the distribution of data is not specified. Many elaborate estimators, such as the GMM estimator of the solution to equation (1), proposed by Hansen (1982) and the GEL estimators by Newey and Smith (2004) can be discussed in terms of the moment restriction model. In particular, we transform equation (3) into one without a conditional expectation operator in order to estimate its parameters.

Let us define an N error vector $\mathbf{u}_{t+1}(\theta)$ as

$$\mathbf{u}_{t+1}(\theta) = \left[\beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \cdot (\mathbf{1} + \mathbf{r}_{t+1}) \right] - \mathbf{1},$$

where $\mathbf{r}_{t+1} = (r_{1,t+1}, r_{2,t+1}, \cdots, r_{N,t+1})$, $\mathbf{1} = (1, 1, \cdots, 1)$ and $\theta = (\beta, \gamma)'$. Let $\mathbf{z}_t$ be a m vector of instruments known at time t, and define an $m = N \cdot K$ vector $g_t(\theta)$ as

$$g_t(\theta) = \mathbf{u}_{t+1}(\theta) \otimes \mathbf{z}_t. \tag{5}$$

Then the Euler equation implies

$$E[g_t(\theta)] = \mathbf{0}, \tag{6}$$

where $E[\cdot]$ is the unconditional expectation operator. We call this equation a moment restriction model.

Generally, let $y_t, (t = 1, 2, \cdots, T)$, denote observations on a finite dimensional process, which is usually assumed to be stationary and strongly mixing (see Smith (2004)). The right-hand side in equation (5), $g_t(\theta)$, is called a moment indicator, which is a function with respect to the parameters concerned but also depends on the data, y_t, and potentially on the instruments, $\mathbf{z}_t$'s. For economic notation, we omit the dependence of g on y_t and $\mathbf{z}_t$ in g.

In the next subsection, we present two methods to estimate the parameters in moment restriction models: GMM and GEL.

2.3 GMM and GEL

This subsection provides a brief review of the estimators of a moment restriction model, following Newey and Smith (2004). The model in subsection 2.2 presented above is one with m moment restrictions. Before reviewing the estimators of the parameters in such a model, let us define some notation. $\mathbf{x}_t, (t = 1, 2, \cdots, T)$ denotes i.i.d. observations of the data, and p (equal to 2 in our model) denotes the number of parameters to be estimated. We sometimes write $g(\mathbf{x}, \theta)$ as a m vector of functions of the data and the parameters. We assume that $m \geq p$ and that the model has a true parameter θ_0 satisfying the following condition:

$$E[g(\mathbf{x}, \theta)] = 0,$$

where E is expectation taken with respect to the distribution of x_t's. In order to explain the 2S-GMM estimator of Hansen (1982), let $g_t(\theta) = g(x_t, \theta)$, $\hat{g}(\theta) = T^{-1}\sum_{i=1}^{T} g_t(\theta)$, and $\hat{\Omega}(\theta) = T^{-1}\sum_{i=1}^{T} g_t(\theta)g_t(\theta)'$. 2S-GMM requires some preliminary estimator, given by $\tilde{\theta} = \arg\min_{\theta\in\Theta} \hat{g}(\theta)'\hat{W}^{-1}\hat{g}(\theta)$, where Θ denotes the parameter space, and $\hat{W}$ is a random matrix with a desired property. Alternatively, sometimes the identity matrix is used for $\hat{W}$. The 2S-GMM estimator is defined as

$$\hat{\theta}_{2S-GMM} = \arg\min_{\theta\in\Theta} \hat{g}(\theta)'\hat{\Omega}(\tilde{\theta})^{-1}\hat{g}(\theta). \tag{7}$$

There is another GMM estimator. Instead of taking the weighting matrix as given in each optimization step of the GMM estimation, we also consider an estimator, called the CU-GMM estimator of Hansen et al. (1996), in which the covariance matrix W is continuously updated as the estimate changes in the optimization. Formally,

$$\hat{\theta}_{CU-GMM} = \arg\min_{\theta\in\Theta} \hat{g}(\theta)'\hat{\Omega}(\theta)^{-1}\hat{g}(\theta). \tag{8}$$

According to Newey and Smith (2004), the 2S-GMM is second order asymptotically equivalent; they also show that the higher order asymptotic bias, which will be described below, of the CU-GMM estimator is smaller. This difference is due to the randomness of W; this property will be verified in Section 4.

In general, the GMM estimator is

$$\hat{\theta}_{GMM} = \arg\min_{\theta\in\Theta} \hat{g}(\theta)'\hat{\Omega}(\theta^*)^{-1}\hat{g}(\theta), \tag{9}$$

where θ^* is a convergent estimate of θ_0, and $\hat{\Omega}(\theta^*)$ is an heteroskedasticity and autocorrelation consistent (HAC) matrix like the one proposed by Newey and West (1987), which depends on a kernel and its bandwidth that can be chosen using the procedure of Andrews (1991). In this paper, we use his method to obtain the desired bandwidth of the kernel. Several applied econometricians have used Hansen (1982)'s J test of overidentifying restrictions to confirm the goodness-of-fit of the model. Under the null hypothesis that equation (6) is true, T times the minimized value of $\hat{g}(\theta)'\hat{\Omega}(\theta^*)^{-1}\hat{g}(\theta)$ is asymptotically distributed as χ^2_{m-p}, where p is the number of parameters.

However, it is widely known that the GMM have poor small sample properties. Many econometricians have made an effort to improve these small sample properties and have suggested several alternative estimators. These include the EL estimator of Owen (1988), CUE of Newey and Smith (2004), and the ET estimator of Kitamura and Stutzer (1997). The EL and ET estimators belong to a class of GEL estimators; the CU-GMM of Hansen et al. (1996) is also a member of this class, as shown by Newey and Smith (2004).

To explain GEL estimators, let $\rho(v)$ be a concave function on a real open interval $\mathcal{V}$ containing zero. The GEL estimator is defined as

$$\hat{\theta}_{GEL} = \arg\min_{\theta\in\Theta} \sup_{\lambda\in\hat{\Lambda}_t(\theta)} \sum_{t=1}^{T} \rho(\lambda' g_t(\theta)), \tag{10}$$

where Θ denotes the parameter space, and $\hat{\Lambda}_t(\theta) = \{\lambda : \lambda' g_t(\theta) \in \mathcal{V}\}$. Alternative estimators to the GMM estimator can be obtained by specifying $\rho(v)$. As Kitamura and Stutzer (1997) show, in the case of the empirical likelihood (EL) estimator, $\rho(v) = \ln(1 - v)$, and in the case the exponential tilting (ET) estimator, $\rho(v) = -e^v$. Furthermore, when $\rho(v) = -(1 + v)^2/2$, the GEL estimator is equivalent to the CUE estimator,

$$\hat{\theta}_{CUE} = \arg\min_{\theta\in\Theta} \hat{g}(\theta)'\hat{\Omega}(\theta)^{-}\hat{g}(\theta), \tag{11}$$

where A^- denotes any generalized inverse of a matrix A (see Theorem 2.1 in Newey and Smith (2004)).

For convenience, we impose a normalization on $\rho(v)$ as follows. Let $\rho_j(v) = \partial^j \rho(v)/\partial v^j$ and $\rho_j = \rho_j(0)$ for each j. We assume that $\rho_1 = \rho_2 = -1$. Associated with each GEL estimator are implied probabilities for the observed data. Since these probabilities are used in our empirical analysis, we verify brief review them. Consider $\rho(v)$, an associated GEL estimator $\hat{\theta}$, and $\hat{g}_t(\hat{\theta})$. The implied probabilities are

$$\hat{\pi}_t = \frac{\rho_1(\hat{\lambda}'\hat{g}_t)}{\sum_{s=1}^{T} \rho_1(\hat{\lambda}'\hat{g}_s)}, \quad t = 1, 2, \cdots, T, \tag{12}$$

where $\hat{\lambda} = \arg\max_\lambda \sum_{t=1}^{T} \rho(\lambda'\hat{g}_t)/T$. For any function $f(x, \theta)$ and GEL estimator θ, an efficient estimator of $E[f(x, \theta_0)]$, $\sum_{t=1}^{T} \hat{\pi}_t f(x_t, \hat{\theta})$, can be derived, as shown in Brown and Newey (1998).

Similar to the J-statistic in GMM estimation, we can use a J-statistic to test the overidentifying restrictions when a GEL estimator employed. The J-statistic for using GEL estimator is computed using the kernel-smoothed moment indicator (see Section 4 in Smith (2004) for details). Under the null hypothesis that equation (6) is true, the statistic is asymptotically distributed as χ^2_{m-p}, where p is the number of parameters.

2.4 Higher Order Bias of the GEL and the GMM Estimators

This subsection reviews how to calculate empirically biases of the estimators of moment restriction models GMM and GEL based on Newey and Smith (2004), who find the asymptotic bias and higher order variance using stochastic expansions. We have MSE to compare the accuracy of the estimates of the CCAPM under the assumption that each estimate is true, that is, without first order bias. We give formulas to calculate the asymptotic bias and higher order variance empirically using the implied probabilities, which assumed to be $1/T$ for the GMM estimates.

From Theorem 3.1 in Newey and Smith (2004),

$$Bias(\hat{\theta}_{GMM}) = B_I + B_G + B_\Omega + B_W, \tag{13}$$

$$B_I = H(-a + E[G_t H g_t])/T,$$
$$B_G = -\Sigma E[G_t P g_t]/T,$$
$$B_\Omega = H E[g_t g_t' P g_t]/T,$$
$$B_W = -H \sum_{j=1}^{p} \bar{\Omega}_{\theta_j} (H_W - H)' e_j / T,$$

where

$$G_t(\theta) = \frac{\partial g_t(\theta)}{\partial \theta},$$

and we define

$$G = E[G_t(\theta_0)],$$

both $p \times m$ matrices.

$$\Omega = E[g_t(\theta_0) g_t(\theta_0)']$$

is a $p \times p$ positive definite matrix. Now we define

$$\Sigma = (G'\Omega^{-1}G)^{-1}, \quad H = \Sigma G' \Omega^{-1}, \quad P = \Omega^{-1} - \Omega^{-1} G \Sigma G' \Omega^{-1},$$

$$H_W = (G'W^{-1}G)^{-1} G' W^{-1}.$$

The matrix W appearing in the definition of H_W depends on the initial weighting matrix $\hat{W}$ for the GMM estimation. Assumption 4 in Newey and Smith (2004) states that there exists W and $\xi(z)$ such that $\hat{W} = W + \sum_{t=1}^{T} \xi(z_t)/T + O_p(T^{-1})$, W is positive definite, $E[\xi(z_t)] = 0$, and $E[||\xi(z_t)||^6] < \infty$. The reader should note that the higher order bias depends on the preliminary estimator $\tilde{\theta}$ only through the limit W and the functions $\xi(z_t)$. Let

$$\bar{\Omega}_{\theta_j} = E[\partial(g_t(\theta_0) g_t(\theta_0)')/\partial \theta_i], \quad i = 1, 2, \cdots, p,$$

and let a be a m vector such that

$$a_j = \mathrm{tr}(\Sigma E[\partial^2 g_{tj}(\theta_0)/\partial\theta\partial\theta'])/2, \quad j = 1, 2, \cdots, m,$$

where $g_{tj}(\theta)$ is the jth element of $g_t(\theta)$, and e_j is a unit vector.

As Newey and Smith (2004) explain, the terms of $Bias(\hat{\theta}_{GMM})$ have the following interpretation. B_I is precisely the asymptotic bias for a GMM estimator with the optimal (asymptotic variance minimizing; Hansen (1982)) linear combination $G'\Omega^{-1}g(\theta)$. B_G is the bias arising from estimation of G, and similarly, B_Ω is that from Ω. The term B_W arises from the preliminary estimator chosen. For GEL, Newey and Smith (2004) derive the following bias formula,

$$Bias(\hat{\theta}_{GEL}) = B_I + \left(1 + \frac{\rho_3}{2}\right) B_\Omega. \tag{14}$$

In the special case with $\rho_3 = -2$, $Bias(\hat{\theta}_{GEL}) = B_I$.

When we calculate the bias for GMM and GEL estimates empirically, we need to take expectation, as shown in the definitions of G, a, and B_I. We use $\hat{\pi}_t$, $(t = 1, 2, \cdots, T)$, to compute expectation for GEL, and $1/T$ is used as identical implied probabilities for GMM.

Anatolyev (2005) extends the result of Smith (2004) to stationary time series. Considering serial correlation, he derives the bias formulas for kernel-smoothed GEL. However, his formulas are so complicated that it is impractical to calculate the biases in small sample cases. As will be mentioned in Section 4, the optimal kernel weights do not exceed one when we employ the kernel-smoothed GEL estimation. This suggests that the kernel smoothing does not function in our case and that there is little difference between the biases of Newey and Smith (2004) and Anatolyev (2005). Therefore, we adopt the bias formula of Newey and Smith (2004) in this paper.

Although Newey and Smith (2004) use higher order bias to compare several estimators, including the GMM and GEL estimators, we use their biases to compute empirically the MSE suggested by Donald and Newey (2001) in order to compare the accuracy of the estimates from our model. Their MSE is

$$MSE = \left(\text{2nd order bias of } \hat{\theta}\right)^2 + \left(\text{1st order asymptotic variance of } \hat{\theta}\right). \qquad (15)$$

Note that we ignore the first order bias when we adopt the MSE of Donald and Newey (2001) and that our main concern is the "quality" of the estimators, which is related to the curvature of the underlying parameter space.

We use the above tools to strengthen the identification of CCAPM in the following empirical analyses.

3 Data

In this paper, quarterly data from 1980Q3 to 2009Q4. The returns on the short-term instruments are employed as the return on risk-free asset and these were obtained from Nikko Financial Intelligence. The Fama-French's market portfolio returns are treated as the returns on the risky asset and these are obtained from Nikkei Portfolio Master.[1] Per capita consumption is computed as "Nondurable goods plus service consumption (benchmark year 2000)" divided by the estimates of the total population reported in the *Annual Report on National Accounts* in Japan. The per capita consumption data are seasonally adjusted using the X-12 ARIMA procedure.

To deflate all series, the "Nondurable plus service consumption" deflator published in the *Annual Report on National Accounts* is used.[2] Lagged values of the real return on the risk-free asset, the real return on the market portfolio, and the real consumption growth rate are used. For the GMM and GEL estimator, all variables that appear in the moment conditions should be stationary. To check whether the variables satisfy stationarity, we use the ADF test of Dickey and Fuller (1981). Table 1 provides some descriptive statistics and the results of the ADF tests. For all the variables, the ADF test rejects the null hypothesis that the variable contains a unit root at conventional significance levels.

[1]Fama-French's market factors in Japan are calculated by following Kubota and Takehara (2007).

[2]The "Nondurable plus service consumption" deflator is a weighted inflation rate using "Nondurable goods" and "Service" deflators that are also published in the *Annual Report on National Accounts*.

(Table 1 around here)

4 Empirical Results

In this section, we estimate the two basic parameters in CCAPM. Then, we check whether the risk-free rate puzzle is still a puzzle in the Japanese financial markets.

4.1 GMM and GEL Estimates

To confirm the accuracy of our estimates, we compare GMM and GEL estimates for CCAPM. Table 2 shows the empirical results with GMM estimators (2S-GMM and CU-GMM). In GMM estimations, we employ an appropriate HAC covariance matrix of Andrews (1991) to reduce estimation biases, which is the asymptotically optimal lag truncation/bandwidth for the quadratic spectral kernel estimator we used.

(Table 2 around here)

All estimates of β and γ are statistically significant at conventional levels. The estimates of β range from 0.9972 to 0.9985, which is plausible, but the estimates of γ range from 0.5600 to 1.0645, which implies lack of robustness. The p-values for Hansen's J test are large enough that we cannot reject the null that the moment conditions hold.

Table 2 also shows the empirical results with GEL estimators (CUE, EL, and ET). In GEL estimations, we choose the truncated kernel proposed by Kitamura and Stutzer (1997) and Smith (1997) to smooth the moment function (that is, equation (10) in our case) because Anatolyev (2005) demonstrates that, in the presence of correlation in the moment function, the smoothed GEL estimator is efficient.[3] In addition, we employ an appropriate HAC covariance matrix of Andrews (1991) to reduce estimation biases. The estimates of β and γ are statistically significant at conventional levels. The estimates of β range from 0.9981 to 0.9987; the estimates of γ range from 0.8026 to 0.8969. In contrast to the GMM estimators, the GEL estimates are very stable. The p-values for the Hansen's J test are large enough that we cannot reject the null that the moment conditions hold.

In addition, we calculate the higher order biases of Newey and Smith (2004) to investigate the asymptotic higher order properties of our estimates.[4] When the sample size is not so large that we cannot rely on the GMM estimates, we should suspect the reliability of the estimates. Table 3 shows higher order biases and MSEs for each estimate.

(Table 3 around here)

We find that higher order biases of the 2S-GMM estimates is are more than 20 times larger than those of the CU-GMM and the GEL estimates (CUE, EL, and ET). In particular,

[3]We employ the smoothed GEL estimator, but the optimal kernel weights do not exceed one. This suggests that the kernel smoothing has no effect.

[4]In the case of stationary time series, we should employ the formula in Anatolyev (2005). However, because there are no kernel-smoothing effects in our estimations, we employ the formula in Newey and Smith (2004).

biases of the preliminary weighting matrix estimator are huge in the 2S-GMM estimates. This suggests that both the GMM estimates are unreliable because of potential biases. We also find that the higher order MSEs of these estimates are also more than 20 times larger than those of the CU-GMM and the GEL estimates. We should note that the value of B_W in case of CU-GMM is zero, verifying the independence of the preliminary estimate of the covariance matrix W, as suggested in Section 2.3. For the same reason as in case of higher order biases, The GMM estimates other than CU-GMM are unreliable. This result corresponds to that of Noda and Sugiyama (2010), who compare the shapes of objective functions to be minimized for 2S-GMM and CU-GMM (see Figures 3 and 4 in Noda and Sugiyama (2010) for details).

Therefore, we conclude that the CU-GMM and the GEL estimates are indisputably better than the 2S-GMM estimates in asymptotic higher order properties when the sample size is small. We obtain the economically realistic parameters of the standard CCAPM when we employ the CU-GMM estimator and GEL estimators.

4.2 A Solution to the Risk-Free Rate Puzzle in Japan

Although several earlier studies attempted to resolve the risk-free rate puzzle in the Japanese financial markets, there is not yet a general consensus. For example, the estimates by Hamori (1992) lead to the conclusion that the risk-free rate puzzle does not exist, while Nakano and Saito (1998) report quite opposite results: the puzzle exists, as does the equity premium puzzle. We confirm that the Hamori's (1992) 2S-GMM estimates of β and γ lead to the conclusion when we substitute them into the formula of Kandel and Stambaugh (1991). However, we argue that his estimates by the 2S-GMM estimator is unreliable as we show that the higher order biases of the 2S-GMM estimates is quite large for samples with the size of around 100. Furthermore, he fails to avoid the problem of weak identification as Stock and Wright (2000) point out; Noda and Sugiyama (2010) show that the CU-GMM estimate of CCAPM applied to the Japanese financial data successfully identifies while that of 2S-GMM does not (see Figures 3 and 4 in Noda and Sugiyama (2010) for details). In turn, Nakano and Saito (1998) assert that their 2S-GMM estimates of β and γ by a single asset CCAPM with stock data lead to contradiction among the sample moments in three markets: stock, real estate, and call money, suggesting the existence of the risk-free rate puzzle. However, their analysis has two drawbacks. Their estimates are as unreliable as those of Hamori (1992), and estimates of a single asset CCAPM cannot produce a contradiction among several financial markets to lead to the puzzle.

Therefore, we investigate whether there is the risk-free rate puzzle in the Japanese financial markets when we use the CU-GMM and GELs estimates. Under the assumption of joint conditional lognormality and homoskedasticity of asset returns, Hansen and Singleton (1983) deliver a convenient equation:

$$0 = E_t[r_{i,t+1}] + \log\beta - \gamma E_t[\Delta C_{t+1}] + \frac{1}{2}\left(\sigma_i^2 + \gamma^2\sigma_C^2 - 2\gamma\sigma_{ic}\right), \tag{16}$$

where σ_i and σ_C are the standard deviations of the i'th asset and consumption, respectively, and σ_{ic} is the covariance between them. This equation implies the following equation shown by Kandel and Stambaugh (1991):

$$E[r_t^f] = -\log\beta + \gamma g - \frac{\gamma^2\sigma_c^2}{2}, \tag{17}$$

where $E[r_t^f]$ is the unconditional expectation of the risk-free interest rate, g is the mean growth rate of real consumption, and σ_c^2 is the variance of g.[5] When we substitute the estimates of β and γ on CCAPM into this equation and employ the CU-GMM and the GEL estimates (CUE, EL, and ET), we derive $E[r_t^f]$ in the range 0.0046 to 0.0054, which is close to 0.0050, the sample mean of the returns on the risk-free asset (see Table 1 for details). Thus, we conclude that the risk-free rate puzzle does not exist in Japan when if one adopts the appropriate empirical method.

5 Concluding Remarks

Following Noda and Sugiyama (2010), who use CU-GMM in place of 2S-GMM to estimate the parameters of CCAPM and find that CU-GMM successfully identifies the degree of risk aversion of CCAPM for the Japanese data, we obtain more accurate estimates of the parameters in CCAPM by using alternative estimators in the GELs and by computing their higher order biases following Newey and Smith (2004) to compare the MSEs of estimators of the GMMs and the GELs.

We can summarize the estimates of CCAPM as follow. First, the estimators of the CU-GMM and GEL (CUE, EL, and ET) perform better than the 2S-GMM estimator, as Newey and Smith (2004) suggest. Second, we find that CU-GMM apparently identifies the model parameters for the macroeconomic data in Japan, although its MSE is not always small compared to the estimators in the GELs. We confirm that the higher order MSEs of the 2S-GMM estimate are more than 20 times larger than those of the CU-GMM and the GEL estimates; this reflects the poor performance of GMM estimate when we focus on the higher order biases. Our empirical results suggest that there is some room for discussing the validity of CCAPM, which many applied econometricians have abandoned in favor of alternative models.

Substituting our estimates using the CU-GMM and GELs into the equation (17) for checking the resolution of the risk-free rate puzzle, we conclude that there is no puzzle in the Japanese financial markets, whereas it is still a puzzle for the U.S. data. It should be noted that even the CU-GMM and the GEL estimators fail to resolve the problem of weak identification: all these estimators gave us unstable estimates of parameters of CCAPM and the corresponding alternative models for the U.S. data. The difficulty relates to puzzles in the U.S. What the differences are between the financial markets in the U.S. and those in Japan is left to further research. Furthermore, as Hyde et al. (2005) and Engsted and Møller (2010) show, CCAPMs possibly have some explanatory power for European financial markets. Whether French, German, and Danish data, which the above authors analyzed, resolve the risk-free rate puzzle is an open question.

[5] The equation of Kandel and Stambaugh (1991) is a special case of the "mean-variance" representation of interest rates derived by Breeden (1986).

References

Akaike, H. (1973): "Information Theory and an Extension of the Maximum Likelihood Principle," in *Second International Symposium on Information Theory*, ed. by B. Petroc and F. Csake, Akademiai Kiado, 267–281.

Anatolyev, S. (2005): "GMM, GEL, Serial Correlation, and Asymptotic Bias," *Econometrica*, 73, 983–1002.

Andrews, D. (1991): "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817–858.

Bakshi, G. and A. Naka (1997): "An Empirical Investigation of Asset Pricing Models using Japanese Stock Market Data," *Journal of International Money and Finance*, 16, 81–112.

Basu, P. and K. Wada (2006): "Is Low International Risk Sharing Consistent with a High Equity Premium? A Reconciliation of Two Puzzles," *Economics Letters*, 93, 436–442.

Breeden, D. (1986): "Consumption, Production, Inflation and Interest Rates: A Synthesis," *Journal of Financial Economics*, 16, 3–39.

Brown, B. W. and W. K. Newey (1998): "Efficient Semiparametric Estimation of Expectations," *Econometrica*, 66, 453–464.

Dickey, D. and W. Fuller (1981): "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root," *Econometrica*, 49, 1057–1072.

Donald, S. and W. Newey (2001): "Choosing the Number of Instruments," *Econometrica*, 69, 1161–1191.

Engsted, T. and S. Møller (2010): "An Iterated GMM Procedure for Estimating the Campbell–Cochrane Habit Formation Model, with an Application to Danish Stock and Bond Returns," *International Journal of Finance & Economics*, 15, 213–227.

Hamori, S. (1992): "Test of C-CAPM for Japan: 1980-1988," *Economics Letters*, 38, 67–72.

Hansen, L. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029–1054.

Hansen, L., J. Heaton, and A. Yaron (1996): "Finite-Sample Properties of Some Alternative GMM Estimators," *Journal of Business & Economic Statistics*, 14, 262–280.

Hansen, L. and K. Singleton (1982): "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 50, 1269–1286.

——— (1983): "Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns," *Journal of Political Economy*, 91, 249–265.

Hyde, S., K. Cuthbertson, and D. Nitzsche (2005): "Resuscitating the C-CAPM: Empirical Evidence from France and Germany," *International Journal of Finance & Economics*, 10, 337–357.

Imbens, G. (1997): "One-Step Estimators for Over-Identified Generalized Method of Moments Models," *Review of Economic Studies*, 64, 359–383.

Imbens, G., R. Spady, and P. Johnson (1998): "Information Theoretic Approaches to Inference in Moment Condition Models," *Econometrica*, 66, 333–357.

Kandel, S. and R. Stambaugh (1991): "Asset Returns and Intertemporal Preferences," *Journal of Monetary Economics*, 27, 39–71.

Kitamura, Y. and M. Stutzer (1997): "An Information-Theoretic Alternative to Generalized Method of Moments Estimation," *Econometrica*, 65, 861–874.

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

Kubota, K. and H. Takehara (2007): "Farther Validation of Effectiveness of Fama-French Factor Model (Fama-French Model no Yukousei no Saikensho, in Japanese)," *Gendai Finance*, 22, 3–23.

Maki, A. and T. Sonoda (2002): "A Solution to the Equity Premium and Riskfree Rate Puzzles: An Empirical Investigation using Japanese Data," *Applied Financial Economics*, 12, 601–612.

Mehra, R. and E. Prescott (1985): "The Equity Risk Premium: A Puzzle," *Journal of Monetary Economics*, 15, 145–161.

——— (2003): "The Equity Premium Puzzle in Retrospect," in *Handbook of the Economics of Finance*, ed. by G. Constantinides, M. Harris, and R. Stulz, North Holland, vol. 1B, chap. 14, 887–936.

Nakano, K. and M. Saito (1998): "Asset Pricing in Japan," *Journal of the Japanese and International Economies*, 12, 151–166.

Newey, W. and R. Smith (2004): "Higher Order Properties of GMM and Generalized Empirical Likelihood Estimators," *Econometrica*, 72, 219–255.

Newey, W. and K. West (1987): "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

Noda, A. and S. Sugiyama (2010): "Measuring the Intertemporal Elasticity of Substitution for Consumption: Some Evidence from Japan," *Economics Bulletin*, 30, 524–533.

Owen, A. (1988): "Empirical Likelihood Ratio Confidence Intervals for a Single Functional," *Biometrika*, 75, 237–249.

Qin, J. and J. Lawless (1994): "Empirical Likelihood and General Estimating Equations," *Annals of Statistics*, 22, 300–325.

Smith, R. (1997): "Alternative Semi-Parametric Likelihood Approaches to Generalised Method of Moments Estimation," *Economic Journal*, 107, 503–519.

——— (2004): "GEL Criteria for Moment Condition Models," Mimeo.

Stock, J. and J. Wright (2000): "GMM with Weak Identification," *Econometrica*, 68, 1055–1096.

Stock, J., J. Wright, and M. Yogo (2002): "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal of Business & Economic Statistics*, 20, 518–529.

Weil, P. (1989): "The Equity Premium Puzzle and the Risk-Free Rate Puzzle," *Journal of Monetary Economics*, 24, 401–421.

Yogo, M. (2008): "Asset Prices under Habit Formation and Reference-Dependent Preferences," *Journal of Business & Economic Statistics*, 26, 131–143.

Table 1: Descriptive Statistics and Unit Root Tests

Variables	Mean	SD	Min	Max	ADF	Lag	$\mathcal{N}$
CG_t	1.0037	0.0091	0.9770	1.0312	-11.0593	0	
r_t^f	0.0050	0.0063	-0.0143	0.0207	-6.2968	0	118
r_t^m	0.0121	0.1041	-0.3335	0.2331	-7.5304	0	

"CG_t" denotes the gross real per capita consumption growth, "r_t^f" denotes the real return on risk-free asset, "r_t^m" denotes the real return on market portfolio, "SD" denotes the standard deviation, "ADF" denotes the Augmented Dickey-Fuller (ADF) test statistics, "Lag" denotes the lag order selected by Akaike (1973)'s information criterion, and "$\mathcal{N}$" denotes the number of observations. In computing the ADF test, a model with a time trend and a constant is assumed. The critical values at the 1% significance level for the ADF test is "-3.99". The null hypothesis that each variable has a unit root is clearly rejected at the 1% significance level.

Table 2: Empirical Results

	GMM		GEL		
	2S-GMM	CU-GMM	CUE	EL	ET
$\hat{\beta}$	0.9972	0.9985	0.9985	0.9987	0.9981
	[0.0007]	[0.0009]	[0.0008]	[0.0009]	[0.0008]
$\hat{\gamma}$	0.5600	1.0645	0.8502	0.8969	0.8026
	[0.1883]	[0.2612]	[0.2306]	[0.2403]	[0.2204]
p_J	0.9179	0.2195	0.6227	0.6184	0.6142
DF	6		8		

"$\hat{\beta}$" denotes the estimate of the subjective discount rate, "$\hat{\gamma}$" denotes the estimate of degree of the relative risk aversion, "p_J" denotes the p-value for Hansen's J and "DF" denotes the degrees of freedom for the Hansen's J test. The Andrews (1991) adjusted standard errors for each of the estimates are reported in brackets. R version 2.12.1 was used to compute the estimates, the starting values of the parameters are set equal to $\beta = 1$ and $\gamma = 1$.

Table 3: Higher Order Biases for Each Estimates

Estimates	B_I	B_G	B_Ω	B_W	B_T	MSE
$\hat{\beta}_{2S-GMM}$	0.0001	-0.0002	0.0001	0.0044	0.0044	1.3139
$\hat{\gamma}_{2S-GMM}$	0.0258	-0.0501	0.0075	1.1475	1.1307	
$\hat{\beta}_{CU-GMM}$	0.0002	-0.0004	0.0003	0.0000	0.0001	0.0684
$\hat{\gamma}_{CU-GMM}$	0.0590	-0.1054	0.0610	0.0000	0.0146	
$\hat{\beta}_{CUE}$	0.0002	-	0.0003	-	0.0005	0.0603
$\hat{\gamma}_{CUE}$	0.0418	-	0.0427	-	0.0845	
$\hat{\beta}_{EL}$	0.0002	-	-	-	0.0002	0.0601
$\hat{\gamma}_{EL}$	0.0489	-	-	-	0.0489	
$\hat{\beta}_{ET}$	0.0001	-	0.0003	-	0.0003	0.0518
$\hat{\gamma}_{ET}$	0.0381	-	0.0368	-	0.0565	

B_I denotes the asymptotic biases for a GMM estimator with the optimal linear combination, B_G denotes the estimation biases of G, B_Ω denotes the estimation biases of the second moment matrix, B_W denotes the estimation biases of preliminary weighting matrix, B_T denotes the total higher order biases, and MSE denotes the higher order mean squared error for each estimate. To compute the estimates, R version 2.12.1 was used.