Title	Inforamtion criteria for moment restriction models : an application of empirical cressie-read estimator for CCAPM
Sub Title	
Author	伊藤, 幹夫(Ito, Mikio) 野田, 顕彦(Noda, Akihiiko)
Publisher	Keio Economic Society, Keio University
Publication year	2010
Jtitle	Keio Economic Society discussion paper series Vol.10, No.1 (2010.)
JaLC DOI	
Abstract	We show nonexistence of the well known risk free rate puzzle in the Japanese financial markets. The result crucially depends on our accurate estimates of the two basic parameters of the discount factor and the degree of risk aversion appeared in a typical CCAPM. We estimate the parameters by the method recently developed, the generalized empirical likelihood estimation and by selecting instruments appropriately with a new information criterion.
Notes	非掲載ページあり
Genre	Technical Report
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AA10715850-00001001-0001

Keio Economics Society Discussion Paper Series
KESDP 10-1
Information Criteria for Moment Restriction Models:
An Application of Empirical Cressie-Read Estimator for CCAPM

Mikio Ito and Akihiko Noda

This Draft: March 31, 2010

INFORMATION CRITERIA FOR MOMENT RESTRICTION MODELS: AN APPLICATION OF EMPIRICAL CRESSIE-READ ESTIMATOR FOR CCAPM[1]

By Mikio Ito[2] and Akihiko Noda[3]

We show nonexistence of the well known risk free rate puzzle in the Japanese financial markets. The result crucially depends on our accurate estimates of the two basic parameters of the discount factor and the degree of risk aversion appeared in a typical CCAPM. We estimate the parameters by the method recently developed, the generalized empirical likelihood estimation and by selecting instruments appropriately with a new information criterion.

Keywords: Risk Free Rate Puzzle, GEL, GMM, Information Criterion, Weak Identification, CCAPM.

1. INTRODUCTION

This paper shows nonexistence of the well known risk free rate puzzle in the Japanese financial markets unlike in U.S. The result crucially depends on our accurate estimates of the two basic parameters of the discount factor and the degree of risk aversion appeared in a typical CCAPM. Instead of the generalized moment method (GMM, Hansen (1982)), we estimate the parameters by the method recently developed, the generalized empirical likelihood (GEL, Smith (1997) and Newey and Smith (2004)) estimation and by selecting instruments appropriately with a new information criterion, the empirical Cressie-Read information criterion (ECR-IC) proposed by Sueishi (2009). Our empirical results shows that estimates of the parameters of CCAPM strongly depend on the methods chosen, GMM or GEL, and on which instruments are selected. Furthermore, the GEL estimates, proved better small sample property, suggest nonexistence of the risk free rate puzzle in Japan.

Initially, Mehra and Prescott (1985) point a puzzle, the inability of standard intertemporal economic models such as CCAPM to rationalize the statistics that have characterized 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 concerning the intertemporal rational behavior of participants in financial markets; we can easily check the puzzle with several statistics calculated from financial data and with parameters estimated, the discount factor and the degree of risk aversion. Inspired by the puzzle, Weil (1989) point another puzzle, a derivative of the equity premium puzzle. In turn, economists confront the inability of the models to explain the average return of safe asset. The puzzles are still puzzles for 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, a number of economists modify theoretical models by introducing additional settings, such as habit formation, imperfect market or trading cost while others try different types of consumers' preferences such as the Kreps-Porteus utility function suggests. Some researchers pay attention to data used in their empirical work. There are preceding studies which make the same assertion as ours. Most of them introduce additional assumptions or extremely modify the basic CCAPM to have the convenient estimates: Bakshi and Naka (1997) use an asset pricing model with habit formation, Maki and Sonoda (2002) consider a trading costs and Basu and Wada (2006) estimate CCAPM considering the international risk sharing between U.S. and Japan.

While we pay attention to methods to estimate the parameters in CCAPM, few economists have been concerned over alternative methods to estimate the parameters in the underlying asset pricing model. Since Hansen and Singleton (1982), GMM have been used in estimating the basic parameters in CCAPM with moment restrictions. GMM provides a general framework which enable 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; it is even more flexible than the estimators since it only requires some assumptions about moment restrictions.

When we apply GMM to estimate the parameters of CCAPM, several drawbacks that arise from the method have been reported: (1) the problem of weak identification and (2) the one of many moment conditions. We can understand both the two problems in the context of the small sample property of GMM; the estimate of GMM has a non-negligible small sample bias when we failed to choose the appropriate instruments and too

[1]We would like to thank Taisuke Otsu for his helpfull comments. We also would like to thank Colin McKenzie and Shunsuke Sugiyama for their helpfull discussions.
[2]Corresponding Author: Fuculty of Economics, Keio University, 2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan.
[3]Keio Advanced Research Centers, Keio University, 2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan.

many moment conditions let us fail to extract information from the data available. As Stock and Wright (2000) report, the conventional GMM procedures of CCAPM and linear instrumental varables regression bring us the breakdown of analysis when some or all of parameters are wealkly identified (see also Stock, Wright, and Yogo (2002) for details). Therefore, more and more econometricians who are interested in CCAPM have given up applying GMM to macroeconomic time series data.

In an effort to improve the poor performance of GMM in case of small samples, a number of alternative estimators have been suggested. A class of GEL estimators is attracting many econometricians because of their better performance than GMM's. All of the members of the class of GEL estimators have the same asymptotic distributions as GMM; 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 report the advantage of such estimators. For instance, Noda and Sugiyama (2010) use the continuous updating estimator (CUE, Hansen, Heaton, and Yaron (1996)) in place of two-step GMM (2S-GMM) to estimate the parameters of CCAPM; they find that CUE procedure successfully identifies the parameters of CCAPM for the Japanese data. When Yogo (2008) estimates the parameters in an asset pricing model under habit formation, he uses the same procedure and reports its validity.

Following them, in place of 2S-GMM, we attempt some estimators in the GEL family to estimate the parameters of a typical CCAPM and select the appropriate instruments using the empirical Cressie-Read information criterion (ECR-IC) proposed by Hong, Preston, and Shum (2003) and Sueishi (2009). We find that the GEL family estimators, CUE, the empirical likelihood (EL, Owen (1988), Qin and Lawless (1994) and Imbens (1997)) and the exponential tilting (ET, Kitamura and Stutzer (1997) and Imbens, Spady, and Johnson (1998)), perform better than the 2S-GMM estimator as Newey and Smith (2004) suggest. We also find that CUE also belongs to the GMM family of estimators, successfully identifies the model parameters when we apply the CUE to the macroeconomic time series data in Japan.

Section 2 presents the review of a typical CCAPM and the several estimators that we compare: 2S-GMM, ET, EL, and CUE. We also explain the empirical Cressie-Read information criterion (ECR-IC). Section 3 gives the reader information about the data we use. Section 4 shows our empirical results: (1) possibility for overcoming the difficulty in estimating CCAPM, the problem of weak identification by using the GEL family of estimators and the empirical Cressie-Read information criterion and (2) 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 a typical CCAPM (Consumption-based Capital Asset Pricing Model) and empirical methods to estimate the parameters. First we give a brief review of a CCAPM with the CRRA utility function for the readers' convenience. Secondly, we present moment restriction models, which are enabling us to cope generally with statistically procedures: estimating parameters, hypotheses testing, and selecting models. Thirdly, we show methods to estimate the parameters in our CCAPM; the conventional 2S-GMM and methods using criteria such as EL, ET and CUE in a class of GEL. which also belong to a class of the empirical Cressie-Read estimators (ECR). Fourthly, we introduce an information criteria proposed by Andrews (1999) and Sueishi (2009) to choose the appropriate instruments for each estimators. Finally, we summarize alternative tests of statistics appeared in the moment restriction models.

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 consumer's optimization problem is:

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

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

where C_t is real per capita consumption at time t, p_{it} is the price of the i'th asset at time t, d_{it} is the dividend of the i'th asset at time t, A_{it} is the amount of the per capita holdings of the i'th asset at time t, Y_t is real per capita labor income at time t, 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 derive the following Euler equation:

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

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

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

At this stage, we do not suppose any DGP of C_t's and r_t's. In the following part of this section, we show the idea to estimate the parameters in this Euler equation with the conditional expectation operator and to make statistical inference on them.

2.2. *Moment Restriction Model*

We present a framework, called a moment restriction model, letting us to cope generally with statistical models such as the CCAPM, in which we do not specify the distribution of data. Many elaborated estimators such as equation (1) GMM one proposed by Hansen (1982) and GEL ones shown in Newey and Smith (2004) can be discussed as the moment restriction model.

For readers' convenience, we introduce a moment restriction model by transforming the equation (3) into another one without any conditional expectation operator to estimate the parameters in the Euler equation: (1) the GMM estimator proposed by Hansen (1982) and the GEL estimators shown in Newey and Smith (2004). Both types of the estimators are based on a moment restriction model.

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

$$\mathbf{u}_{t+1}(\theta)=\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}\cdot(\mathbf{1}+\mathbf{r}_{it+1})\right]-\mathbf{1},\quad \textit{where}\ \ \mathbf{r}_{it+1}=(r_{1t+1},r_{2t+1},\cdots,r_{Nt+1},)$$

where $\theta=(\beta,\gamma)'$. Let $\mathbf{z}_t$ be an $R\times 1$ vector of instruments known at time t, and define an $N\times R$ vector $m_t(\theta)$ as

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

Then, the Euler equation implies

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

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

Generally, let $y_t,(t=1,\cdots,T)$, denote observations on a finite dimensional process, which is usually supposed to be stationary and strongly mixing (see Smith (2004)). The right-hand side in the equation (5), $m_t(\theta)$ is called a moment indicator, which is a function with respect to the parameters concerened while it also depends on the data, y_t's and probably on the instruments, $\mathbf{z}_t$'s. When we focus our attention on parameters to be estimated, y_t's and $\mathbf{z}_t$'s are usually omitted.

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

2.3. *GMM and GEL*

First we give here a brief review of GMM. If we define $\overline{m_T(\theta)}$ as

$$\overline{m_T(\theta)}:=\frac{1}{T}\sum_{t=1}^{T}m_t(\theta), \tag{7}$$

where T denotes the sample size, then the GMM estimator of θ, $\hat{\theta}_{GMM}$, minimizes the quadratic form:

$$\hat{\theta}_{GMM}=\arg\min_{\theta}\overline{m_T(\theta)\prime}\mathbf{W}_T\overline{m_T(\theta)}, \tag{8}$$

where $\mathbf{W}_T$ is an $R\times R$ weighting matrix, which is supposed to be positive definite for any finite T.

We can obtain the most efficient GMM estimator by choosing the weighting matrix $\mathbf{W}_T = \mathbf{S}^{-1}$, where $\mathbf{S}^{-1}$ is the inverse of the asymptotic covariance matrix of $T^{1/2}\overline{m_T(\theta)}$. However, because we cannot observe the true value of θ, we cannot know $\mathbf{S}^{-1}$ either. Conventionally most econometricians adopt two-step GMM (2S-GMM). In order to estimate $\mathbf{S}$, we use the estimator with the bandwidth parameter proposed by Andrews (1991).

As will shown in Section 2.5, a number of 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 equation (8) is asymptotically distributed as $\chi^2_{N\times R-K}$, where K is the number of parameters.

However, it is widely known that the 2S-GMM estimator has poor small sample properties. Many econometricians have made an effort to improve the small sample properties of GMM and have suggested a number of alternative estimators. These include the EL estimator of Owen (1988), CUE of Hansen, Heaton, and Yaron (1996), and the ET estimator of Kitamura and Stutzer (1997). The EL and ET estimators belong to a class of the GEL estimators; CUE is also a member of this class as shown by Newey and Smith (2004). They point that many estimators in GEL family, for example, ET, EL and CUE, can be represented as the ECR estimators as explained as follows.

The ECR estimator is based on the idea that an estimated distribution from data should be close to the true distribution with respect to an information criterion called the Cressie-Read information criterion (CRIC). Let $g(y)$ be a density function, the Cressie-Read (CR) discrepancy proposed by Cressie and Read (1984) from g to the true density f measures the closedness between g and f:

$$CR(f,g) = \int f(y) h\left(\frac{g(y)}{f(y)}\right) dy \tag{9}$$

where

$$h(x) = \frac{x^{\alpha+1}-1}{\alpha(\alpha+1)}, \quad -\infty < \alpha < \infty \tag{10}$$

Note that the function h is continuous with respect to α even when $\alpha = -1$ or $\alpha = 0$ both of which case can be handled by taking limits. The two cases correspond to the Kullback-Leibler (KL) discrepancy ($h(x) = -\log x$) and the entropy ($h(x) = x \log x$).

In what follows, T observed y's denote $y_1, y_2, \cdots, y_T$ and we assume that they are independent and identically distributed random variables.

We suppose that a statistical model is represented by a set of parametric density functions $\{g(y,\beta) : \beta \in \mathcal{B}\}$, where $\mathcal{B}$ is a parameter space. The CRIC is defined as

$$CRIC(f, g_\beta) = \min_{\beta\in\mathcal{B}} \int f(y) h\left(\frac{g(y,\beta)}{f(y)}\right) dy. \tag{11}$$

The CRIC represents the closedness between the true density function $f(y)$ and its candidate $g(y,\beta)$. We can consider an estimator fo β which minimizes this criterion as

$$\beta^* = \arg\min_{\beta\in\mathcal{B}} \int f(y) h\left(\frac{g(y,\beta)}{f(y)}\right) dy. \tag{12}$$

The reader should note that the traditional maximum likelihood (ML) estimator is the special case of this type of estimator when we suppose the KL discrepancy ($\alpha = -1$) and choose the empirical distribution as $f(y)$.

Following the idea of Newey and Smith (2004), we can apply CRIC to moment restriction models, which covers CCAPM. Now we rewrite the equation (6):

$$E[m(y,\theta_0)] = \mathbf{0}, \tag{13}$$

where $m : \mathbf{R}^d \times \Theta \longrightarrow \mathbf{R}^\ell$ is a known function. The expectation is taken with respect to the true DGP of y. In our case, d is $2 \times k$ where k is the number of instruments and ℓ $2 \times (k+1)$. The reader should note that each component of the vector y corresponds to the product of the error term and the instruments.

CRIC for a moment restriction model space $\mathcal{M}$ which consists of density functions that hold the underlying moment restrictions is defined as $\min_{g\in} \mathcal{M}$. As Sueishi (2009) presents, we characterize CRIC in two steps.

Define $\mathcal{M}_\theta$ as the set $\{g(y) : \int m(y,\theta)g(y)dy = \mathbf{0}\}$. Note that $\mathcal{M} = \cup_\theta \mathcal{M}_\theta$. For fixed $\theta \in \Theta$, we first consider a minimizing problem:

$$\begin{aligned} &\text{Min} \quad \int f(y) h\left(\frac{g(y,\beta)}{f(y)}\right) dy, \\ &\text{s.t.} \quad \int g(y)dy = 1 \text{ and} \int m(y,\theta)g(y)dy = 0. \end{aligned} \tag{14}$$

Let $v(\theta)$ denote the minimand, $\min_{g\in\mathcal{M}_\theta}$. This optimization is infinite dimensional; it is cumbersome to handle the problem directly. Thus, we consider its finite dimensional dual problem.

$$\begin{aligned} &\text{Max} \quad \mu - \int f(y) h^*(\mu + \bar{\lambda}' m(y,\theta))dy, \\ &\text{s.t.} \quad \mu \in \mathbf{R}, \bar{\lambda} \in \mathbf{R}^\ell, \end{aligned} \tag{15}$$

where h^* is the convex conjugate function of h:

$$h^*(x) = \frac{(\alpha x)^{(\alpha+1)/\alpha}}{\alpha+1} - \frac{1}{\alpha(\alpha+1)}. \tag{16}$$

Let $v^*(\theta)$ denote the value at the optimum. Fenchel's duality theorem implies $v(\theta) = v^*(\theta)$. By defining $\lambda = \bar{\lambda}/\mu$ and ignoring a constant, we have

$$CRIC(f, g_\theta) = \min_{\theta\in\Theta} \max_{\lambda\in\Lambda} \int f(y)\rho(\lambda' m(y,\theta))dy, \tag{17}$$

where

$$\rho(\xi) = -\frac{1}{\alpha+1}(1+\alpha\xi)^{(\alpha+1)/\alpha}. \tag{18}$$

and Λ is a feasible set of Lagrangean multipliers. (see Theorem 2.2 in Newey and Smith (2004)). Note that we ignored the constant term in the above definition of CRIC following Sueishi (2009).

We call the value (θ^*, λ^*) attaining $CRIC(f, g_\theta)$ the pseudo-true value. By replacing $f(y)$ by the empirical distribution, we have the ECR estimator:

$$\min_{\theta\in\Theta} \max_{\lambda\in\Lambda} \frac{1}{T}\sum_{i=1}^{T} \rho(\lambda' m(y_i,\theta)), \tag{19}$$

In the rest of this section, we shall sometimes use a notation, $P(\theta, \lambda)$ in place of $T^{-1}\sum_{i=1}^{T} \rho(\lambda' m_{tT}(y_i,\theta))$.

The rest of this subsection summarizes somewhat technical issues. At first, strictly speaking, the ECR family of estimators does not always coincide with GEL. However, they share many estimators that we use here in common. Thus, the reader may identify GEL as ECR in this paper. Second, while GEL estimation does not require to compute explicitly the HAC matrix of the moment conditions, the estimators may not only be inefficient but may also fail to be consistent if we do not take it into account. Smith (2004) proposed a kernel smoothing of the moment indicator (see also Kitamura and Stutzer (1997)). Let m_{tT} be a smoothed moment indicator with a bandwidth parameter S_T and a kernel function, $k(\cdot)$ such as the truncated kernel or the Bartlett kernel (for detail, see Smith (2004)). Thus, the reader should interpret the ECR-IC, as the one made of some kernel smoother.

2.4. *Selection of the Instruments*

It is important to select appropriate instruments to estimate CCAPM by using a moment restriction model. We can deal with the problem by procedure of moment selection that Andrews (1999), Hong, Preston, and Shum (2003) and Sueishi (2009).

Inspired by Andrews (1999), who proposes information criteria like well known, AIC, BIC and HQIC to select appropriate moment conditions Hong, Preston, and Shum (2003) derive the AIC-like criterion for the ECR estimation. This is pararel to AIC, which is the Kullback-Leibler criterion (KLIC) with respect to MLE,

$$\text{AIC} = -2\sum_{i=1}^{T} \log g(y_i, \hat{\beta}) + 2p, \tag{20}$$

where $\hat{\beta}$ is the ML estimate of the vector of parameters in g. Akaike (1970) proposes the procedure for selecting the model which minimize AIC, a good estimate of KLIC. Hong, Preston, and Shum (2003)'s criterion is:

$$\text{ECR-AIC2} = 2\sum_{i=1}^{T} \rho(\hat{\lambda}' m(y_i, \hat{\theta})) - 2(\ell - p), \tag{21}$$

where $\hat{\lambda}$ and $\hat{\theta}$ are the ECR estimates for the data, $(y_1, y_2, \cdots, y_T)$.

In his recent work, Sueishi (2009) presents more sophisticated information criteria inspired by Takeuchi (1976), who derives a general information criteria without assuming correct specification of g in (20). The reader should note that AIC is based on the assumption of correct specification of the statistical model.

Takeuchi (1976)'s information criterion (TIC) is as follows:

$$\text{TIC} = -2\sum_{i=1}^{T} \log g(y_i, \hat{\beta}) + 2\text{tr}(\widehat{Q^{-1}\Omega}), \tag{22}$$

where

$$Q = -E\left[\frac{\partial^2 g(y, \beta^*)}{\partial\beta\partial\beta'}\right], \tag{23}$$

$$\Omega = E\left[\left(\frac{\partial g(y, \beta^*)}{\partial\beta}\right)\left(\frac{\partial g(y, \beta^*)}{\partial\beta'}\right)\right]. \tag{24}$$

Note that TIC coincides with AIC for the model specified correctly since $Q = \Omega$ holds in that case.

Sueishi (2009) applies Takeuchi (1976)'s idea to the ECR estimators of moment restriction models, in which we make no assumption on the underlying distribution. He derives TIC and a different version of AIC than the existing criterion of Hong, Preston, and Shum (2003). Before defining the criteria, we give several notations following Sueishi (2009). Define $\gamma = (\theta', \lambda')'$ and $Q(\gamma) = E[\rho(\lambda' m(y, \theta))]$,

$$\phi(y_i, \gamma) = \frac{\partial}{\partial\gamma'} Q(\gamma). \tag{25}$$

Notice that the ECR estimator $\gamma^* = (\theta^{*\prime}, \lambda^{*\prime})'$ is characterized by the representation of the first order condition:

$$\frac{1}{T}\sum_{i=1}^{T} \phi(y_i, \gamma^*) = \mathbf{0}. \tag{26}$$

In place of Q and Ω in the definition of TIC, Sueishi (2009) introduces the following matrices:

$$H = E\left[\frac{\partial}{\partial\gamma'}\phi(y_i, \gamma)\right], \tag{27}$$

$$S = E\left[\phi(y_i, \gamma)'\phi(y_i, \gamma)\right]. \tag{28}$$

Then, Sueishi (2009) first derives ECR-TIC as follows:

$$\text{ECR-TIC} = 2\sum_{i=1}^{T} \rho(\lambda^{*\prime} m(y_i, \gamma^*)) + \text{tr}(\widehat{H^{-1}S}) + \text{tr}(\widehat{HV}), \tag{29}$$

where $\widehat{H^{-1}S}$ and $\widehat{HV}$ are estimates of $H^{-1}S$ and HV. The matrices V, H and S are defined as

$$\hat{V} = \begin{pmatrix} \hat{V}_{11} & \hat{V}_{11}\hat{A}' \\ \hat{A}\hat{V}_{11} & \hat{A}\hat{V}_{11}\hat{A}' \end{pmatrix}, \tag{30}$$

$$H = E\begin{bmatrix} \rho_{1i}\frac{\partial M_i{}'\lambda^*}{\partial\theta'} + \rho_{2i}M_i{}'\lambda^*\lambda^{*\prime}M_i & \rho_{1i}M_i{}' + \rho_{2i}M_i{}'\lambda^* m(y_i, \theta^*)' \\ \rho_{1i}M_i + \rho_{2i}m(y_i, \theta^*)\lambda^{*\prime}M_i & \rho_{2i}m(y_i, \theta^*)m(y_i, \theta^*)' \end{bmatrix} \tag{31}$$

and

$$S = E\left[\begin{array}{cc} \rho_{1i}^2 M_i{}'\lambda^*\lambda^{*\prime}M_i & \rho_{1i}^2 M_i{}'\lambda^* m(y_i,\theta^*)' \\ \rho_{1i}^2 m(y_i,\theta^*)\lambda^{*\prime}M_i & \rho_{1i}^2 m(y_i,\theta^*)m(y_i,\theta^*)' \end{array}\right] \tag{32}$$

where V_{11} is the asymptotic variance of estimated θ^* and

$$A = -E[\rho_{2i}m(y_i,\theta^*)m(y_i,\theta^*)']^{-1}(E[\rho_{1i}M_i] + E[\rho_{2i}m(y_i,\theta^*)\lambda^{*\prime}M_i]). \tag{33}$$

Notice that we can estimate $H^{-1}S$ by

$$\widehat{H^{-1}V} = \left[\frac{1}{n}\sum_{i=1}^{n}\frac{\partial\phi(y_i,\hat{\gamma}_{ECR})}{\partial\gamma}\right]^{-1}\frac{1}{n}\sum_{i=1}^{n}\phi(y_i,\hat{\gamma}_{ECR})\phi(y_i,\hat{\gamma}_{ECR})'. \tag{34}$$

Sueishi (2009) shows that a different version of AIC for ECR estimators than Hong, Preston, and Shum (2003)'s ECR-AIC2 can be derived as a special case of ECR-TIC. Specifically,

$$\text{ECR-AIC1} = 2\sum_{i=1}^{T}\rho(\lambda^{*\prime}m(y_i,\gamma^*)) - (\ell - 2p), \tag{35}$$

In case of CCAPMs represented by moment restrictions, we do not set a specific distribution of observations, the series of consumption and interest rate; we cannot also specify valid information set. We consider the problem of selection of instruments not for specifying the valid information set, infinitely dimensional, but for improving the forecasting error of the parameters estimated. As we show in the next section, the AICs indicate similar instruments as is appropriate under the assumption that we do correctly specify the model.

2.5. *Statistical Inference*

This subsection presents alternative tests of statistics based on GMM and GEL for confirming the validity of the over identifying restrictions and for checking the significance of parameters concerned. We give a brief review of the statistics: the J statistic, the likelihood ratio (LR) one, the Lagrange multiplier (LM) one, S_{GEL} one. All statistics are asymptotically χ^2 distributed under the several suitable assumptions.

Conventionally, most econometricians have used a type of J test for moment restriction models when they make statistical inference since Hansen (1982) proposed the over identifying test. Such a test is based on the following asymptotic property. Given some regularity conditions, the estimators for the moment restrictions models converge in law to the true ones. Specifically, under several assumptions set on the GMM estimation, the estimator of θ holds $\hat{\theta} \xrightarrow{p} \theta_0$

$$\sqrt{T}(\hat{\theta} - \theta_0) \xrightarrow{L} N(\mathbf{0}, \Sigma) \tag{36}$$

where θ_0 is the true parameter vector,

$$\Sigma = E\left(\frac{\partial m(y,\theta_0)}{\partial\theta}\right)'\Omega(\theta_0)E\left(\frac{\partial m(y,\theta_0)}{\partial\theta}\right) \tag{37}$$

and $\Omega(\theta_0)$ is the variance of $\sqrt{T}\bar{m}(\theta_0)$. For the GEL estimation, the estimators of θ and λ hold $\hat{\theta} \xrightarrow{p} \theta_0$, $\hat{\lambda} \xrightarrow{p} \lambda_0$ and

$$\sqrt{T}\left(\begin{array}{c} \hat{\theta} - \theta_0 \\ \hat{\lambda} - \lambda_0 \end{array}\right) \xrightarrow{L} N(\mathbf{0}, H^{-1}SH^{-1}) \tag{38}$$

where H and S are already defined. We can therefore make inference on the parameter vector posing the assumption that it is asymptotically distributed as $N(\theta_0, T^{-1}\hat{\Sigma})$. The null hypothesis is $H_0:\ E[m(x,\theta)] = 0$; the statistics are

$$\sqrt{T}\bar{m}(\hat{\theta})'\hat{\Omega}(\theta_0)^{-1}\bar{m}(\hat{\theta}) \xrightarrow{L} \chi^2_{\ell-p} \tag{39}$$

for GMM and

$$2\sum_{t=1}^{T}\left(\rho(\hat{\lambda}'m(y_t,\hat{\theta})) - \rho(0)\right) \xrightarrow{L} \chi^2_{\ell-p} \tag{40}$$

for GEL.

In order to test the over identifying restrictions, Smith (2004) proposes three statistics: (1) J statistic, (2) LM statistic and (3) LR statistic. All of them are asymptotically distributed as a $\chi^2_{\ell-p}$. The J statistic is:

$$T\overline{m_{tT}}'\hat{\Omega}(\hat{\theta})^{-1}\overline{m_{tT}}, \tag{41}$$

the second one is a LM statistic:

$$T\hat{\lambda}'\hat{\Omega}(\hat{\theta})^{-1}\hat{\lambda} \tag{42}$$

and the last one is a LR statistic:

$$2\sum_{t=1}^{T}\left(\rho(\hat{\lambda}'m_{tT}(y_t,\hat{\theta}))-\rho(0)\right). \tag{43}$$

Guggenberger and Smith (2005) and Otsu (2006) present robust inference methods for moment restriction models with weakly identified parameters. The latter propose two robust test statistics based on GEL: (1) S statistic, the kernel-smoothed GEL criterion function of Smith (2004) and (2) K statistic, derived from a GEL criterion function which uses transformed moment restrictions of which dimensions is identical to the number of parameters. Otsu (2006) asserts that the limiting distribution of K statistic does not depend on the number of moment restrictions and that the size property of the S statistic is sensitive to the sample size. In this paper, we adopt the S statistic because the number of moment restrictions is relatively small and because of its convenience.

Following Otsu (2006), let S_{GEL} be

$$\max_{\lambda}\hat{P}(\theta,\lambda)=\max_{\lambda}T^{-1}\sum_{i=1}^{T}\rho(\lambda'm_{tT}(y_i,\theta)). \tag{44}$$

Otsu (2006) considers a situation where there might be weakly identified parameters although the other ones are strongly identified. Suppose that α is a parameter vector weakly identified and β a parameter vector strongly identified. Note $\theta=(\alpha,\beta)$. Let $\hat{\theta}_\alpha:=(\alpha,\hat{\beta}(\alpha))$, where $\hat{\beta}(\alpha)$ is the GEL estimator for β_0, the true parameter, given α. The S_{GEL} statistic for testing $H_0:\alpha=\alpha$ is defined as

$$\hat{S}_{GEL}(\alpha):=S_{GEL}(\hat{\theta}_\alpha)=\min_{\beta}S_{GEL}(\alpha,\beta). \tag{45}$$

Otsu (2006) proves that $2T(\hat{S}_{GEL}(\alpha_0)-\rho(0))\times M$ is asymptotically distributed as $\chi^2_{\ell-p_\beta}$, where $M=S_T^{-1}\kappa_2/\kappa_1^2$, $\kappa_1=\int k(x)dx$ and $\kappa_2=\int k(x)^2dx$ for normalization.

We use the above tools to improve the difficulty of the weak identification of CCAPM in the following part of empirical analyses.

3. DATA

We construct two datasets labeled "Dataset 1" and "Dataset 2" in this paper that differ in their choices of consumption. In all datasets, the money market returns (the real call rate) and the stock returns (real rate of returns for stocks in the Tokyo stock exchange (first section)) are treated as assets in the Euler equation, and those total asset return are obtained from Ibbotson Associates. To compute the inflation rate, we use the "Total consumption" and the "Nondurable plus service consumption" deflator published in the *Annual Report on National Accounts.* [1]

In "Dataset 1", quarterly data, for per capita consumption, we adopt "Total consumption (Benchmark year is 2000)" divided by population which is reported in the *Annual Report on National Accounts* in Japan.[2] The per capita consumption data are seasonally adjusted using the X-12 ARIMA procedure. As instruments, we use the lagged values of the real return on assets, the real consumption growth rate, and the growth rate of the deflator. The sample period is from 1980Q2 to 2008Q4. "Dataset 2" differs from "Datasets 1" only in the measure of consumption used. In "Dataset 2", also quarterly data, we adopt "Nondurable goods plus services" as consumption data. Therefore, the real return on assets and real consumption series are computed using the deflator for each consumption data.

[1] "Nondurable plus service consumption" deflator is a weighting inflation rate using "Nondurable goods" and "Service" deflator also published in the *Annual Report on National Accounts.*

[2] We use the labor force as population which is reported in the *Population Estimates* in Japan.

For both the GMM family (2S-GMM and CUE-GMM) and the GEL family (CUE-GEL, EL and ET), 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 I 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.

(Table I 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 puzzle or not in the Japanese financial markets.

4.1. *With GMM Family Estimators*

To confirm the accuracy of our estimates, we first present the GMM estimates for CCAPM. Table II presents the empirical results with GMM family estimators (2S-GMM and CUE-GMM) using "Dataset 1" and "Dataset 2". In 2S-GMM, we employ an appropriate heteroskedasticity and autocorrelation consistent (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 II around here)

Table II shows that the CUE-GMM estimates of β and γ are statistically significant at conventional levels except for using up to fifth lagged instruments.[3] The estimates of β range from 0.9953 to 1.0013; the estimates of γ range from 0.3513 to 1.8123. The p-values for Hansen's J test are large enough that we cannot reject the null that the moment conditions hold. We confirm that the CUE-GMM results are robust to changes of the initial starting values while we omit the detail.

Table II also shows that the 2S-GMM estimates of β range from 0.9851 to 0.9965, which are plausible, but the estimates of γ are negative regardless of the type of the dataset in most cases. Moreover, the p-values for the Hansen's J test are not large enough that we can reject the null that the moment conditions hold in some cases. In contrast to CUE-GMM, we have confirmed that 2S-GMM fails to provide us with robust estimates to changes of these initial values.

Concerning 2S-GMM with $Lag = 1$, we have similar estimates of β and γ to the CUE-GMM ones. At the first glance these results look similar to Hamori (1992). However, we should not conclude the validity of these results because the 2S-GMM estimator has a serious small sample bias in case of many moment conditions. Therefore, we should confirm the robustness of our estimates when we use fewer moment conditions. We estimate the parameters of typical CCAPM in case of one asset. Table III shows that the 2S-GMM and CUE-GMM estimates using call money as the asset.

(Table III around here)

The 2S-GMM estimates are very unstable; the CUE-GMM estimates are stable. These results appear to be same as Stock and Wright (2000) and Noda and Sugiyama (2010). Therefore, we conclude that γ is weakly identified when we employ the 2S-GMM.

4.2. *With GEL Family Estimators*

Table IV through Table VI indicate the empirical results with the GEL family estimators (CUE-GEL, EL and ET) using "Dataset 1"and "Dataset 2". In GEL estimations, we choose the truncated kernel proposed by Kitamura and Stutzer (1997) to smooth the moment function (that is, the equation (19) in our case) because Anatolyev (2005) demonstrates that, in the presence of correlation in the moment function, the smoothed GEL estimator of Kitamura and Stutzer (1997) is efficient. And we employ an appropriate HAC covariance matrix of Andrews (1991) or Newey and West (1994) to reduce estimation biases.[4]

[3]The minimization of CUE-GMM objective function is not converge only when we use up to fifth lagged instruments.
[4]We obtain same empirical results regardless of the type of the HAC covariance matrix.

(Table IV, V and VI around here)

All tables show that the estimates of β and γ are statistically significant at the conventional level when we employ the GEL family estimators and that the estimates of β and γ are also stable.

We summarize the estimates of using the GEL family estimators as follows. (1) we obtain the almost the same estimates of β and γ regardless of the empirical method, (2) Using "Dataset 1", the estimates of β range from 0.9952 to 1.0116 and the estimates of γ from 0.3713 to 5.0519 and (3) Using "Dataset 2", the estimates of β range from 0.9964 to 1.0022 and the estimates of γ from 0.8394 to 2.3741. At this stage, however, we do not know which instruments are appropriate to estimate the parameters of a typical CCAPM using the GEL family estimators. Then we select the appropriate instruments using ECR-IC proposed by Hong, Preston, and Shum (2003) and Sueishi (2009). These tables also show: (1) the ECR-IC does not perform well empirically except for using CUE-GEL and (2) appropriate instruments for CUE-GEL is $Lag = 1$ regardless of the dataset, and then the estimates of β are 0.9952 (using "Dataset 1") and 0.9971 (using "Dataset 2"), the estimates of γ are 0.3785 (using "Dataset 1") and 0.8569 (using "Dataset 2").[5]

Therefore, we obtain the economically realistic parameters of a typical CCAPM when we employ CUE-GEL. These empirical results are consistent with the recent Monte Carlo simulations by Donald, Imbens, and Newey (2009), who show that the CUE-GEL has not only lower higher-order bias but also lower higher-order variance than the other estimators (2S-GMM, EL and ET).

4.3. *Statistical Inference for Weak Identification with CUE-GEL*

In order to check whether there are weakly identified parameters with CUE-GEL or not, we employ the S-statistic proposed by Guggenberger and Smith (2005) and Otsu (2006). In our case, we assume that the estimates of γ is a weakly identified parameter and the estimates of β is a strongly identified parameter. Therefore, we test the null hypothesis H_0 is that $\hat{\gamma} = \gamma_0$, where γ_0 is the true value of γ. Table VII shows that the S-statistic for weak identification with GEL-CUE.

(Table VII around here)

Table VII shows that the null hypothesis is not rejected at the conventional level of significance. Furthermore, we can consider that the estimates of γ is not a weakly identified parameter when we employ the GEL-CUE.

4.4. *A Solution to Risk Free Rate Puzzle in Japan*

Although several earlier studies address to resolve the risk free rate puzzle in the Japanese financial market, there is no consensus about it. For example, the estimates by Hamori (1992) lead to a conclusion that the risk free rate puzzle does not exist, while Nakano and Saito (1998) report quite opposite results: the puzzle exists as well as the equity premium puzzle does. In case of Hamori (1992)'s estimates, they are unreliable as our Tables III and II show; the 2S-GMM brings us highly volatile estimates and fails to let the estimate of γ satisfy the desirable sign condition. When we assign his estimates of β and γ into the equation in Kandel and Stambaugh (1991), explained below, the formula exhibits good fit by accident.

In turn, Nakano and Saito (1998) assert that their 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. Then, Nakano and Saito (1998) suggests the existence of the risk free rate puzzle. However, their conclusion has two drawbacks. First, their estimates are unreliable as Hamori (1992)'s are. Second, the estimates of a single asset CCAPM can not make any contradiction among several financial markets leading us to the puzzle.

In the context of the previous empirical work on risk free rate puzzle, other earlier studies introduce additional assumptions or extremely modify the basic CCAPM to improve the risk free rate puzzle in the Japanese financial market: Bakshi and Naka (1997) use an asset pricing model with habit formation, Maki and Sonoda (2002) consider a trading costs and Basu and Wada (2006) estimate CCAPM considering the international risk sharing between U.S. and Japan.

Therefore, we investigate whether there is the risk free rate puzzle or not in the Japanese financial markets. Under the assumption of joint conditional lognormality and homoskedasticity of asset returns, Hansen and Singleton (1983) deliver a convenient equation:

$$0 = E_t[r_{i,t+1}] + \log\beta - \gamma E_t[\Delta C_{t+1}] + \frac{1}{2}\left(\sigma_i^2 + \gamma^2\sigma_C^2 - 2\gamma\sigma_{ic}\right), \tag{46}$$

[5] We also obtain the similar results when we estimate the parameters of typical CCAPM in case of one asset.

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

$$E[r_t^f] = -\log\beta + \gamma g - \frac{\gamma^2\sigma_c^2}{2}, \tag{47}$$

where $E[r_t^f]$ is the unconditional mean of risk free interest rate, g the mean growth rate of real consumption and σ_c^2 the variance of g.[6] We substitute our estimates of β and γ on CCAPM into this equation We employ the $\hat{\beta}_{CUE-GEL}$ and $\hat{\gamma}_{CUE-GEL}$ using $Lag = 1$ instruments for each datasets, which are the valid estimates of β and γ on CCAPM selected by ECR-IC. Then we derive the $E[r_t^f]$ are 0.0061 (using "Dataset1") and 0.0052 (using "Dataset 2"), these are almost equal to 0.0061 and 0.0050, which are the sample mean of the returns on the risk-free asset (see Table I for details). Thus we conclude that the risk free rate puzzle does not exist in Japan when we adopt the appropriate empirical method (CUE-GEL).

5. CONCLUDING REMARKS

Following Noda and Sugiyama (2010), who use CUE-GMM in place of 2S-GMM to estimate the parameters of CCAPM and find that CUE successfully identifies the parameter, the degree of risk aversion, of CCAPM for the Japanese data, we have more accurate estimates of the parameters in CCAPM by using alternative estimators in the GEL family and by selecting the appropriate instruments using ECR-IC proposed by Hong, Preston, and Shum (2003) and Sueishi (2009) than by using 2S-GMM.

We can summarize the estimates of CCAPM as follow. First, the estimators of GEL family (CUE-GEL, EL and ET) perform better than the 2S-GMM estimator as Newey and Smith (2004) suggest. Our estimates are robust over the two datasets and over the number of assets. Second, we find that CUE-GMM proposed by Hansen, Heaton, and Yaron (1996), which also belongs to the GMM family of estimators, successfully identifies the model parameters for the macro economics data in Japan. Third, CUE-GEL provides us with the most stable estimates for our goal, in the sense that all estimates for various Lag's satisfy the desirable sign condition and give relatively stable results on ECR-AIC1. 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 by using CUE-GEL to the well known equation by which we can check the resolution of the risk free rate puzzle, we conclude that there is no puzzle in the Japanese financial markets, while it is still a puzzle for the US data. It should be noted that even the GEL estimators fail to resolve the problem of weak identification, which brings us unstable estimates of parameters of CCAPM or its alternative models for the US data. The difficulty relates to the puzzles in the US. What difference in the financial markets in the US and Japan is left for the further research.

REFERENCES

Akaike, H. (1970): "Statistical Predictor Identification," *Annals of the Institute of Statistical Mathematics*, 22(1), 203–217.
Anatolyev, S. (2005): "GMM, GEL, Serial Correlation, and Asymptotic Bias," *Econometrica*, 73(3), 983–1002.
Andrews, D. (1991): "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59(3), 817–858.
——— (1999): "Consistent Moment Selection Procedures for Generalized Method of Moments Estimation," *Econometrica*, 67(3), 543–564.
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(1), 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(3), 436–442.
Breeden, D. (1986): "Consumption, Production, Inflation and Interest Rates: A Synthesis," *Journal of Financial Economics*, 16(1), 3–39.
Cressie, N., and T. Read (1984): "Multinomial Goodness-of-Fit Tests," *Journal of the Royal Statistical Society. Series B*, 46(3), 440–464.
Dickey, D., and W. Fuller (1981): "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root," *Econometrica*, 49(4), 1057–1072.
Donald, S., G. Imbens, and W. Newey (2009): "Choosing Instrumental Variables in Conditional Moment Restriction Models," *Journal of Econometrics*, 152(1), 28–36.
Guggenberger, P., and R. Smith (2005): "Generalized Empirical Likelihood Estimators and Tests under Partial, Weak, and Strong Identification," *Econometric Theory*, 21(4), 667–709.
Hamori, S. (1992): "Test of C-CAPM for Japan: 1980-1988," *Economics Letters*, 38(1), 67–72.
Hansen, L. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50(4), 1029–1054.

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

Hansen, L., J. Heaton, and A. Yaron (1996): “Finite-Sample Properties of Some Alternative GMM Estimators,” *Journal of Business and Economic Statistics*, 14(3), 262–280.

Hansen, L., and K. Singleton (1982): “Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models,” *Econometrica*, 50(5), 1269–1286.

——— (1983): “Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns,” *Journal of Political Economy*, 91(2), 249–265.

Hong, H., B. Preston, and M. Shum (2003): “Generalized Empirical Likelihood-Based Model Selection Criteria for Moment Condition Models,” *Econometric Theory*, 19(6), 923–943.

Imbens, G. (1997): “One-Step Estimators for Over-Identified Generalized Method of Moments Models,” *Review of Economic Studies*, 64(3), 359–383.

Imbens, G., R. Spady, and P. Johnson (1998): “Information Theoretic Approaches to Inference in Moment Condition Models,” *Econometrica*, 66(2), 333–357.

Kandel, S., and R. Stambaugh (1991): “Asset Returns and Intertemporal Preferences,” *Journal of Monetary Economics*, 27(1), 39–71.

Kitamura, Y., and M. Stutzer (1997): “An Information-theoretic Alternative to Generalized Method of Moments Estimation,” *Econometrica*, 65(4), 861–874.

Kocherlakota, N. (1996): “The Equity Premium: It’s Still a Puzzle,” *Journal of Economic Literature*, 34(1), 42–71.

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(8), 601–612.

Mehra, R., and E. Prescott (1985): “The Equity Risk Premium: A Puzzle,” *Journal of Monetary Economics*, 15(2), 145–161.

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

Nakano, K., and M. Saito (1998): “Asset Pricing in Japan,” *Journal of the Japanese and International Economies*, 12(2), 151–166.

Newey, W., and R. Smith (2004): “Higher Order Properties of GMM and Generalized Empirical Likelihood Estimators,” *Econometrica*, 72(1), 219–255.

Newey, W., and K. West (1994): “Automatic Lag Selection in Covariance Matrix Estimation,” *Review of Economic Studies*, 61(4), 631–653.

Noda, A., and S. Sugiyama (2010): “Measuring the Intertemporal Elasticity of Substitution for Consumption: Some Evidence from Japan,” *Economics Bulletin*, 30(1), 524–533.

Otsu, T. (2006): “Generalized Empirical Likelihood Inference for Nonlinear and Time Series Models under Weak Identification,” *Econometric Theory*, 22(3), 513–527.

Owen, A. (1988): “Empirical Likelihood Ratio Confidence Intervals for a Single Functional,” *Biometrika*, 75(2), 237–249.

Qin, J., and J. Lawless (1994): “Empirical Likelihood and General Estimating Equations,” *Annals of Statistics*, 22(1), 300–325.

Smith, R. (1997): “Alternative Semi-Parametric Likelihood Approaches to Generalised Method of Moments Estimation,” *Economic Journal*, 107(441), 503–519.

——— (2004): “GEL Criteria for Moment Condition Models,” mimeo.

Stock, J., and J. Wright (2000): “GMM with Weak Identification,” *Econometrica*, 68(5), 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 and Economic Statistics*, 20(4), 518–529.

Sueishi, N. (2009): “Information Criteria for Moment Restriction Models,” mimeo.

Takeuchi, K. (1976): “Distribution of Information Statistics and Criteria for Adequacy of Models,” *Mathematical Science*, 153, 12–18, in Japanese.

Weil, P. (1989): “The Equity Premium Puzzle and the Risk-rate Puzzle,” *Journal of Monetary Economics*, 24(3), 401–421.

Yogo, M. (2008): “Asset Prices under Habit Formation and Reference-Dependent Preferences,” *Journal of Business and Economic Statistics*, 26(2), 131–143.

TABLE I

Descriptive Statistics and Unit Root Tests

D/S	Variables	Mean	SD	Min	Max	ADF	$\mathcal{N}$
1	CG_t	1.0033	0.0107	0.9569	1.0321	-10.8501	115
	r_t^f	0.0061	0.0065	-0.0185	0.0205	-5.5675	
	r_t	0.0123	0.1049	-0.3355	0.2396	-6.9726	
	π_t	1.0018	0.0062	0.9916	1.0301	-7.1070	
2	CG_t	1.0027	0.0093	0.9759	1.0293	-11.5397	115
	r_t^f	0.0050	0.0065	-0.0141	0.0228	-6.1035	
	r_t	0.0112	0.1051	-0.3354	0.2427	-6.9901	
	π_t	1.0029	0.0058	0.9907	1.0333	-7.4078	

"D/S" denotes the dataset, "CG_t" denotes the gross real per capita consumption growth, "r_t^f" denotes the real return on risk-free asset (the real call rate), "r_t" denotes the real return on risky asset (the real rate of return for stocks in Tokyo stock exchange (first section)), "π_t" denotes the inflation rate, "SD" denotes the standard deviation, "ADF" denotes the Augmented Dickey-Fuller (ADF) test statistics, and "$\mathcal{N}$" denotes the number of observations. In computing the ADF test, we assume a model with a time trend and a constant. 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 rejected at the 1% significance level.

TABLE II

Empirical Results with 2S-GMM and CUE-GMM: case of two assets

E/M	D/S	Lag	$\hat{\beta}$	$SE(\hat{\beta})$	$\hat{\gamma}$	$SE(\hat{\gamma})$	p_J	DF
2S-GMM	1	1	0.9941	0.0006	0.2014	0.1117	0.5988	8
		2	0.9899	0.0003	-0.5623	0.0869	0.0149	16
		3	0.9909	0.0001	-0.1536	0.0271	0.0545	24
		4	0.9891	0.0001	-0.5683	0.0196	0.2333	32
		5	0.9862	0.0000	-1.4086	0.0043	0.3935	40
	2	1	0.9965	0.0008	0.7366	0.2108	0.5552	8
		2	0.9924	0.0002	-0.1044	0.0647	0.0111	16
		3	0.9903	0.0001	-0.7257	0.0286	0.0577	24
		4	0.9904	0.0001	-0.7005	0.0226	0.1770	32
		5	0.9851	0.0000	-2.4997	0.0109	0.3549	40
CUE-GMM	1	1	0.9953	0.0007	0.3513	0.1403	0.6382	8
		2	1.0013	0.0003	1.8123	0.1079	0.2381	16
		3	1.0010	0.0002	1.5292	0.0765	0.4535	24
		4	0.9999	0.0001	1.2779	0.0316	0.6341	32
		5	-	-	-	-	-	-
	2	1	0.9971	0.0009	0.8221	0.2280	0.5573	8
		2	0.9988	0.0002	1.4822	0.0954	0.2400	16
		3	0.9989	0.0001	1.5474	0.0456	0.2921	24
		4	0.9996	0.0001	1.5986	0.0303	0.5195	32
		5	-	-	-	-	-	-

"E/M" denotes the empirical method used, "D/S" denotes the dataset used, "Lag" denotes the number of lags of the instruments used, "$\hat{\beta}$" denotes the estimate of the subjective discount rate, "$\hat{\gamma}$" denotes the estimate of the relative risk aversion, "$SE(\cdot)$" denotes the Andrews adjusted standard error of "$\hat{\beta}$" or "$\hat{\gamma}$", respectively, "p_J" denotes the p-value for Hansen's J test statistics and "DF" denotes the degrees of freedom for the Hansen's J test. To compute the estimates, R version 2.10.1 was used. The starting values of the parameters are set equal to $\beta = 1$, $\gamma = 1$.

TABLE III

EMPIRICAL RESULTS WITH 2S-GMM AND CUE-GMM: CASE OF ONE ASSET

E/M	D/S	Lag	$\hat{\beta}$	$SE(\hat{\beta})$	$\hat{\gamma}$	$SE(\hat{\gamma})$	p_J	DF
2S-GMM	1	1	0.9924	0.0013	-0.2958	0.1150	0.0113	2
		2	0.9919	0.0007	-0.6572	0.0577	0.0063	5
		3	0.9923	0.0007	-0.5007	0.0437	0.0344	8
		4	0.9912	0.0001	-0.6196	0.0161	0.0638	11
		5	0.9912	0.0001	-0.5527	0.0137	0.4448	14
	2	1	0.9941	0.0014	-0.3834	0.1726	0.0022	2
		2	0.9942	0.0009	-0.2464	0.0709	0.0471	5
		3	0.9954	0.0005	0.0807	0.0462	0.1209	8
		4	0.9973	0.0009	0.6777	0.1900	0.0333	11
		5	0.9977	0.0009	0.8658	0.1601	0.1402	14
CUE-GMM	1	1	0.9953	0.0007	0.3191	0.1397	0.3864	2
		2	1.0089	0.0041	3.8794	0.6720	0.1750	5
		3	1.0046	0.0024	2.6237	0.4177	0.2255	8
		4	1.0058	0.0025	2.7623	0.3999	0.2867	11
		5	-	-	-	-	-	-
	2	1	0.9976	0.0010	0.9231	0.2717	0.3334	2
		2	1.0018	0.0018	2.6422	0.2297	0.2130	5
		3	1.0040	0.0017	3.0577	0.2686	0.2833	8
		4	1.0036	0.0008	2.7855	0.0818	0.4728	11
		5	-	-	-	-	-	-

As for Table III.

TABLE IV

EMPIRICAL RESULTS WITH CUE-GEL: CASE OF TWO ASSETS

D/S	Lag	$\hat{\beta}$	$SE(\hat{\beta})$	$\hat{\gamma}$	$SE(\hat{\gamma})$	p_J	p_{LM}	p_{LR}	ECR-AIC1	ECR-AIC2	DF
1	1	0.9952	0.0008	0.3785	0.1340	0.8099	0.8099	0.8099	-114.94	-124.94	10
	2	1.0116	0.0004	3.5207	0.1236	0.1566	0.1566	0.1566	-105.05	-123.05	18
	3	1.0063	0.0003	2.2521	0.0990	0.1461	0.1461	0.1461	-103.43	-129.43	26
	4	1.0006	0.0001	1.1984	0.0388	0.1551	0.1551	0.1551	-102.69	-136.69	34
	5	-	-	-	-	-	-	-	-	-	-
2	1	0.9971	0.0011	0.8569	0.2292	0.7721	0.7721	0.7721	-114.50	-124.50	10
	2	0.9995	0.0003	1.7686	0.1046	0.1213	0.1213	0.1213	-103.87	-121.87	18
	3	0.9976	0.0001	2.0720	0.0437	0.1152	0.1152	0.1152	-102.16	-128.16	26
	4	1.0022	0.0002	2.3741	0.0625	0.1127	0.1127	0.1127	-100.78	-134.78	34
	5	-	-	-	-	-	-	-	-	-	-

As for Table III except for some over-identifying tests and the empirical Cressie-Read information criteria (ECR-IC) for the GEL estimation. "p_{LM}" denotes the p-value for Lagrange multiplier test statistics, "p_{LR}" denotes the p-value for likelihood ratio test statistics which are proposed by Smith (2004). "ECR-AIC1" and "ECR-AIC2" denote the ECR-AIC proposed by Sueishi (2009).

TABLE V

EMPIRICAL RESULTS WITH EL: CASE OF TWO ASSETS

D/S	Lag	$\hat{\beta}$	$SE(\hat{\beta})$	$\hat{\gamma}$	$SE(\hat{\gamma})$	p_J	p_{LM}	p_{LR}	ECR-AIC1	ECR-AIC2	DF
1	1	0.9951	0.0008	0.3713	0.1323	0.8093	0.0000	0.4945	3.40	-6.60	10
	2	1.0116	0.0004	5.0519	0.1487	0.1404	0.0000	0.0000	53.42	35.42	18
	3	1.0030	0.0002	1.0153	0.0699	0.0310	0.0000	0.0000	-	-	26
	4	1.0031	0.0001	1.0063	0.0387	0.0587	0.0000	0.0000	-	-	34
	5	-	-	-	-	-	-	-	-	-	-
2	1	0.9969	0.0010	0.8060	0.2192	0.7639	0.0000	0.4799	3.56	-6.44	10
	2	0.9979	0.0003	1.1774	0.1042	0.0884	0.0000	0.0000	-	-	18
	3	0.9964	0.0001	1.9130	0.0436	0.1080	0.0000	0.0000	71.25	45.25	26
	4	-	-	-	-	-	-	-	-	-	34
	5	-	-	-	-	-	-	-	-	-	-

As for Table IV.

TABLE VI

Empirical Results with ET: case of two assets

D/S	Lag	$\hat{\beta}$	$SE(\hat{\beta})$	$\hat{\gamma}$	$SE(\hat{\gamma})$	p_J	p_{LM}	p_{LR}	ECR-AIC1	ECR-AIC2	DF
1	1	0.9952	0.0008	0.3759	0.1336	0.8098	0.1458	0.7083	-228.82	-238.82	10
	2	1.0041	0.0004	2.5796	0.1225	0.1229	0.0000	0.0002	-196.14	-214.14	18
	3	1.0020	0.0002	0.9334	0.0665	0.0324	0.0000	0.0000	-	-	26
	4	1.0006	0.0002	2.3630	0.0491	0.0896	0.0000	0.0000	-	-	34
	5	-	-	-	-	-	-	-	-	-	-
2	1	0.9970	0.0011	0.8394	0.2255	0.7708	0.1582	0.6705	-228.43	-238.43	10
	2	0.9983	0.0003	1.3118	0.1049	0.1038	0.0000	0.0004	-198.81	-216.81	18
	3	0.9969	0.0001	1.9472	0.0434	0.1101	0.0000	0.0001	-188.94	-214.94	26
	4	0.9975	0.0001	2.0484	0.0310	0.0797	0.0000	0.0000	-	-	34
	5	-	-	-	-	-	-	-	-	-	-

As for Table IV.

TABLE VII

S statistics for the GEL-CUE estimates

D/S	Lag	Case of one asset		Case of two assets	
1	1	0.0000	[1.0000]	0.0000	[1.0000]
	2	0.0000	[1.0000]	0.0000	[1.0000]
	3	0.0000	[1.0000]	0.0000	[0.9998]
	4	0.0000	[1.0000]	0.0000	[0.9948]
	5	0.0000	[1.0000]	-	-
2	1	0.0000	[1.0000]	0.0000	[1.0000]
	2	0.0000	[1.0000]	0.0000	[1.0000]
	3	0.0000	[1.0000]	0.0000	[0.9998]
	4	0.0000	[1.0000]	0.0000	[0.9948]
	5	0.0000	[1.0000]	-	-

"D/S" denotes the dataset, "*Lag*" denotes the number of lags of the instruments used for CUE-GEL. p-value for S statistics are in brackets. To compute each statistics, R version 2.10.1 was used.

KEIO ECONOMIC SOCIETY
DISCUSSION PAPER SERIES

No. 08-1 (2008) Hideo Akabayashi and Michio Naoi "Does the Public Sector Crowd Out the Private Sector in the Higher Education Market?: Theory and Evidence from Japan", 20 pages.

No. 08-2 (2008) 寺出道雄「比例と均衡―「経済表　範式」再考―」、20 pages.

No. 09-1 (2009) Michio Naoi, Miki Seko and Kazuto Sumita "Community Rating, Cross Subsidies and Underinsurance: Shy So Many Households in Japan Do Not Purchase Earthquake Insurance", 25 pages.

No. 09-2 (2009) Michio Naoi, Miki Seko and Kazuto Sumita "Earthquake Risk and Housing Prices in Japan: Evidence Before and After Massive Earthquakes", 30 pages.

No. 09-3 (2009) Miki Seko, Kazuto Sumita and Michio Naoi, "Residential Mobility Decision in Japan: Identifying the Effects of Housing Equity Constraints and Income Shocks under the Recourse Loan System", 24 pages.

No. 09-4 (2009) Takuji Arai, "Convex Risk Measures on Orlicz Spaces: Convolution and Shortfall", 18 pages.

No. 10-1 (2010) Mikio Ito and Akihiko Noda, "Information Criteria for Moment Restriction Models: An Application of Empirical Cressie-Read Estimator for CCAPM", 17 pages.