Nonparametric Estimation and Testing of Stochastic Discount Factor

Ying Fang * Yu Ren † Yufei Yuan ‡

WISE, Xiamen University, China.

February 10, 2011

Abstract

This paper attempts to estimate stochastic discount factor (SDF) proxies non-parametrically using the conditional Hansen-Jagannathan distance. Nonparametric estimation can not only avoid misspecification when dealing with nonlinearity in the model but also provide more precise information about the local properties of the estimators. Empirical studies show that our method performs better than the alternative parametric polynomial models, and furthermore, we find that the return on aggregate wealth can sufficiently explain the SDF proxies when one deals with nonlinearity appropriately.

JEL classification: C13; C52; G12

Keywords: Stochastic Discount Factor; Nonparametric Estimation; HJ Distance

*WISE, Xiamen University, P.R.China, 361005. Fang’s research is partially supported by the National Natural Science Foundation of China 70971113 and by the Fundamental Research Funds for the Central Universities 2010221092.

†Corresponding Author. WISE, Xiamen University, P.R.China, 361005. E-mail: renxmu@gmail.com.

‡WISE, Xiamen University, P.R.China, 361005.
1 Introduction

In finance theory, an asset pricing model usually implies a stochastic discount factor (SDF). SDFs display which prices are reasonable given the potential returns of risky assets. Asset prices can be represented as inner products of payoffs and SDFs. If an asset pricing model represents the true data generating process of the return, SDF can price the risky assets perfectly.

In reality, SDFs cannot be observed. Therefore, we need to find good proxies for them. Because of the close relationship between asset pricing models and SDFs, researchers usually adopt factors that are frequently used in asset pricing models to estimate SDF proxies. Hence, the estimation of SDF proxies encounters problems similar to those in the asset pricing model, such as the problem of factor selection and the problem of model misspecification.

Since one can obtain different SDF proxies in different setups, it is important to compare and evaluate them, which requires a measure of pricing errors produced by different proxies. When there is only one asset, a natural measure is the pricing error associated with that asset, which is the difference between the price of that asset and the hypothetical price assigned by a candidate proxy. In this case, it is straightforward to compare the relative performance of any two SDF proxies. When there are a number of assets, there is a vector of pricing errors associated with each proxy. Comparing the specification error in two different proxies becomes not very straightforward. For this purpose, Hansen and Jagannathan (1997) developed a measure of the degree of misspecification of an asset pricing model. This measure, called the Hansen-Jagannathan (HJ) distance, is defined as the least squares distance between the stochastic discount factor associated with an asset pricing model and the family of stochastic discount factors that price all assets correctly.

When the SDF proxy is assumed to be a parametric model of several factors, it can be estimated by minimizing the sample analogue of the HJ distance. Parametric modeling is simple and easy to estimate. However, it often suffers from the problem of model
misspecification. For example, Dittmar (2002) suggested that the SDF proxy should be a nonlinear function of the return on aggregate wealth. A Taylor series expansion was employed to linearize the model before estimation. Yet there are some limits in Dittmar’s method. Firstly, the signs of the coefficients in the expansion are determined by theory, which may not be always consistent with the data. Secondly, instruments were added to facilitate the time-varying features of the coefficients but the selection of those instruments are quite arbitrary. Both drawbacks motivate the use of a nonparametric method for estimating stochastic discount factors. A nonparametric estimation allows a flexible functional form which is no need to impose the sign restriction onto the coefficients. In addition, the nonparametric method evaluates estimators using local information which automatically adapts to the time-varying features.

In this paper, we propose a local generalized method of moments (GMM) estimator (Lewbel, 2007) of stochastic discount factors. To the best of our knowledge, this is the first paper which tries to estimate SDF proxies using local GMM. Unlike the traditional GMM, the weighting matrix adopted in our model is not the inverse of the variance of the pricing errors. Instead, we use the conditional version of the HJ distance and employ the conditional second moment of asset returns as the weighting matrix. Compared to the unconditional HJ distance, the conditional version can measure the pricing errors of SDF proxies qualified by local information. The proxy estimators are obtained by minimizing the squares of the conditional HJ distance. Moreover, we extend the HJ distance test to the case of a nonparametric estimation and the limiting distribution is also provided.

We apply our estimation and testing method to analyze the Fama-French 25 portfolios by regarding the return on aggregate wealth as a single factor. Our results reveal some interesting findings. Firstly, there is structure changes in SDF proxies. Specifically, as economy conditions change, the SDF proxies may vary from a linear model to a nonlinear model. Secondly, nonlinearity is more likely to occur when an economy is in a period of booming or shrinking. Last but not least, the return on aggregate wealth can explain the
SDF proxies well enough as long as the nonlinearity is dealt with appropriately.

The rest of this paper is organized as follows: Section 2 introduces our econometric model and demonstrates the estimation and testing methods; Section 3 reports our empirical analysis; and Section 4 concludes.

2 Econometric Model

2.1 Unconditional and Conditional Hansen-Jagannathan Distance

Consider a portfolio of $N$ primitive assets, and let $R_t$ denote the $t$-th period gross returns of these assets. $R_t$ is a $1 \times N$ row vector. A valid stochastic discount factor, $m_t$, which is a $1 \times 1$ scalar, satisfies $E(m_t R'_t) = 1_N$, where $1_N$ is a $N \times 1$ vector of ones. If an asset pricing model implies a stochastic discount factor $m_t(\delta)$, where $\delta$ is a $d \times 1$ unknown parameter vector, then the HJ-distance corresponding to this asset pricing model is given by

$$HJ(\delta) = \sqrt{E[w_t(\delta)]'G^{-1}E[w_t(\delta)]},$$  \hspace{1cm} (1)

where $w_t(\delta) = R'_t m_t(\delta) - 1_N$ denotes pricing errors and $G = E(R'_t R_t)$.

Linear factor pricing models assume the SDF to be a linear form $m_t(\delta) = \tilde{X}_t \delta$, where $\tilde{X}_t = [1 \ X_t]$ is a row vector of factors including a constant 1; see Hansen and Jagannathan (1997). The HJ-distance can be estimated by its sample analogue

$$HJ_T(\delta) = \sqrt{w_T(\delta)'G_T^{-1}w_T(\delta)},$$  \hspace{1cm} (2)

where $w_T(\delta) = T^{-1} \sum_{t=1}^{T} w_t(\delta)$ and $G_T = T^{-1} \sum_{t=1}^{T} R'_t R_t$. When $N$ is larger than $d$, following Jagannathan and Wang (1996), the parameter $\delta$ can be estimated by minimizing the sample HJ-distance $HJ_T(\delta)$. Note that the resulting estimator of $\delta_T$ is a GMM
estimator using the moment condition $E[w_t(\delta)] = 0$ and the weighting matrix $G_T^{-1}$.

Nagel and Singleton (2008) proposed a conditional HJ distance in order to incorporate conditional information. Specifically, the conditional HJ distance is defined by

$$
\overline{HJ} = \sqrt{E[w_t|I]'E(R_t'R_t|I)^{-1}E[w_t|I]}, \tag{3}
$$

where the expectations are taken under the information set $I$. The conditional HJ distance evaluates SDF proxies based on an information set. In the case that the two different SDF proxies may statistically generate the same unconditional HJ distance, the conditional measure makes it possible to distinguish them. In addition, knowing the relative performance of SDF proxies for different economic situations, one can possibly construct a SDF proxy that outperforms all the existing ones. In our paper, the conditional HJ distance will serve as the criterion function for our local estimation of the SDF.

### 2.2 Local GMM Estimation of SDFs

Suppose that stochastic discount factor $m_t$ is an unknown function of a single factor $z_t$, i.e. $m_t = m(z_t)$. The pricing error is defined as

$$
w_t = 1_N - R_t'm_t = 1_N - R_t'm(z_t). \tag{4}
$$

Our target is to estimate $m(z_t)$ nonparametrically. When the model is correctly specified, we have

$$
E[w_t|z_t = z_0] = E[1_N - R_t'm(z_0)] = 0. \tag{5}
$$

When $z_t$ is continuously distributed, following Lewbel (2007), Equation (5) holds if and only if

$$
E[w_t(m(z_0))f(z_0)] = 0, \tag{6}
$$

\footnote{In general, $z_t$ can be a $d \times 1$ vector of factors.}
which can be served as a moment condition to estimate \( m(z_0) \). Note that \( f(z_0) \) is the density function of \( z_t \) evaluated at \( z_0 \) and \( E[w_t(m(z_0))f(z_0)] = \int w_t(m(z_0))f(z_0)f(R_t|z_0)dR_t \).

Define

\[
S_T[m(z_0)] = \frac{1}{Th} \sum_{t=1}^{T} [1_N - R'_t m(z_0)]K\left(\frac{z_t - z_0}{h}\right),
\]

(7)

where \( K \left( \frac{z_t - z_0}{h} \right) \) is a Kernel function satisfying regular conditions and \( h \) is the bandwidth. \( S_T(m(z_0)) \) is a localized empirical analogue of the moment condition defined in (6). Let \( G_T(z_0) \) denote the estimator of \( E(R'_tR_t|z_t = z_0) \). Note that a typical element in \( G_T(z_0) \) can be estimated by

\[
(G_T(z_0))_{ij} = \frac{\sum_t R_{it}R_{jt}K\left(\frac{z_t - z_0}{h}\right)}{\sum_t K\left(\frac{z_t - z_0}{h}\right)}.
\]

(8)

Next, the conditional HJ distance can be adopted as a criterion to evaluate pricing errors, which implies that one can minimize

\[
\widehat{HJ}[m(z_0)] = \sqrt{E[w_t|z_t = z_0]'E(R'_tR_t|z_t = z_0)^{-1}E[w_t|z_t = z_0]}
\]

(9)

to obtain the estimator of \( m(z_0) \). The sample analogue of \( \widehat{HJ} \) is

\[
\widehat{HJ}_T[\hat{m}(z_0)] = \sqrt{S_T[\hat{m}(z_0)]'G_T(z_0)^{-1}S_T[\hat{m}(z_0)]}.
\]

(10)

Then we can choose \( \hat{m}(z_0) \) to minimize

\[
\widehat{HJ}_T^2[\hat{m}(z_0)] = S_T[\hat{m}(z_0)]'G_T(z_0)^{-1}S_T[\hat{m}(z_0)].
\]

(11)

The estimator of \( m(z_0) \), denoted by \( \hat{m}(z_0) \), is given by

\[
\hat{m}(z_0) = \arg\min S_T[\hat{m}(z_0)]'G_T(z_0)^{-1}S_T[\hat{m}(z_0)].
\]

(12)

When the bandwidth \( h \) satisfies that \( Th^5 \rightarrow 0 \) and \( Th \rightarrow \infty \), under mild conditions, we
show in the appendix that \( \hat{m}(z_0) \) is a consistent estimator and with a limiting distribution

\[
\sqrt{Th} ( \hat{m}(z_0) - m(z_0) ) \rightarrow^d N[0, (D'G^{-1}D)^{-1} D'G^{-1}V G^{-1}D (D'G^{-1}D)^{-1}] \quad (13)
\]

where \( D = E[R_t' f(z_0)] \), \( G = E(R_t' R_t | z_t = z_0) \) and \( V = E[w_t w_t' f(z_0) \int K(u)^2 du] \). Note that \( G \) can be consistently estimated by \( G_T(z_0) \) and \( V \) can be consistently estimated by \( V_T = S_T(m(z_0)) S_T'(m(z_0)) \).

In the above result, we require a undersmooth such that the bandwidth is chosen by \( h = o(T^{-1/5}) \), which converges to zero faster than the means quare optimal rate \( h = O(T^{-1/5}) \), to avoid the higher order bias due to the nonparametric estimation.\(^2\) As Lewbel (2007) claimed, the higher order bias involves a complicated form and is usually ignored in practice.

### 2.3 Nonparametric Hansen-Jagannathan Distance Test

Jagannathan and Wang (1996) derived the asymptotic distribution of the HJ distance test when the SDF proxy is linear in factors. Here, by following their procedures, we extend the HJ distance test to a nonparametric case.

Note that \( S_T(\hat{m}(z_0)) \) can be approximated around its true value \( m(z_0) \) by using Taylor series expansion,

\[
S_T(\hat{m}(z_0)) = S_T(m(z_0)) + D_T(\hat{m}(z_0) - m(z_0)) \equiv (14)
\]

where \( D_T = \frac{1}{Th} \sum_{t=1}^{T} R_t' K \left( \frac{z_t - m}{h} \right) \).\(^3\) Multiplying both sides by \( D_T' G_T^{-1} \), rearranging the above equation and applying the first order condition of Equation (12), we have

\[
\hat{m}(z_0) - m(z_0) = -(D_T' G_T^{-1} D_T)^{-1} D_T' G_T^{-1} S_T(m(z_0)). \quad (15)
\]

\(^2\) The higher order bias here is \( \frac{h^2}{T} (D'G^{-1}D)^{-1} D'G^{-1} [ \int u^2 K(u) du ] \int w_t f''(z_0, R_t) dR_t \).

\(^3\) Note that \( D_T \) is a consistent estimator of \( D \).
Plugging the above expression into Equation (14), we have

$$S_T(\hat{m}(z_0)) = (I_N - D_T(D'_T G_T^{-1} D_T)^{-1} D'_T G_T^{-1}) S_T(m(z_0)), \quad (16)$$

where $I_N$ is a $N \times N$ identity matrix. Thus the square HJ distance is given by

$$\tilde{H} J_T^2 = S_T(m(z_0))^T (G_T^{-1} - G_T^{-1} D_T (D'_T G_T^{-1} D_T)^{-1} D'_T G_T^{-1}) S_T(m(z_0)). \quad (17)$$

Under regular conditions, we show in the appendix that

$$\sqrt{T} h S_T(m(z_0)) \rightarrow^d N(0, \Sigma) \quad (18)$$

where $\Sigma = V = E[w_t w'_t f(z_0) \int K(u)^2 du]$ which is defined in the previous subsection. Then,

$$Th(\tilde{H} J_T^2) \rightarrow^d \Phi'(G^{-1} - G^{-1} D(D'G^{-1}D)^{-1} D'G^{-1}) \Phi, \quad (19)$$

where $\Phi$ is a $N$-dimensional random vector of the normal distribution with zero mean and covariance matrix $\Sigma$. Define $v = (\Sigma^{-1/2})' \Phi$, then $v$ is a standard normal random variable. Thus, we have

$$Th(\tilde{H} J_T^2) \rightarrow v' \Sigma^{1/2} (G^{-1} - G^{-1} D(D'G^{-1}D)^{-1} D'G^{-1}) \Sigma^{1/2} v. \quad (20)$$

Let $A = \Sigma^{1/2}(G^{-1} - G^{-1} D(D'G^{-1}D)^{-1} D'G^{-1}) \Sigma^{1/2}$. It’s easy to show that the rank of $A$ is $N - 2$. By following Jagannathan and Wang (1996), we have

$$Th(\tilde{H} J_T^2) \rightarrow^d \sum_{j=1}^{N-2} \lambda_j v_j^2, \quad (21)$$

where $v_j \sim N(0, 1)$ and $\lambda_j$ is nonzero eigenvalues of $A$.

The above distribution can be employed to test the hypothesis whether or not the
conditional HJ distance delivered by the estimator is zero. The p-values of the test are calculated by the algorithm proposed by Jagannathan and Wang (1996). First, draw $M \times (N - 2)$ independent random variables from $\chi^2(1)$ distribution. Next, calculate $u_j = \sum_{i=1}^{N-2} \lambda_i \nu_{ij}$ ($j = 1, \ldots, M$). Then the empirical p-value of the HJ-distance test is

$$p = M^{-1} \sum_{j=1}^{M} I[u_j \geq Th(HJ_T^2)],$$

where $I(\cdot)$ is an indicator function that equals one if the expression in the brackets is true and zero otherwise. In our simulations, we set $M = 5,000$.

3 Empirical Analysis

We collect monthly returns on the Fama-French 25 portfolios and the return on aggregate wealth from July 1963 to December 2007. Following Dittmar (2002), the return on aggregate wealth, $R_w$, is defined as the linear combination of the labor return $R_l$ and the market return $R_m$,

$$R_w = aR_m + (1 - a)R_l. \quad (22)$$

In the following analysis, we adopt two ways to choose the weight $a$: fixed value and varying values.

3.1 Fixed $a$

We assume $a = 0.558$, which is the estimation result in Dittmar (2002), and treat $R_w$ as the single factor $z_t$ in our model. Our target is to obtain the estimator of $m(z_0)$, denoted by $\hat{m}(z_0)$, where $z_0$ runs over the whole sample. We report the corresponding conditional HJ distances, and then test the hypothesis whether the distance is zero. At the same time, we conduct a similar analysis by assuming SDF proxies to be a linear, a quadratic or a cubic function of the return on aggregate wealth. The results are summarized by
Figure 1 and Figure 2

Figure 1 shows the conditional HJ distance delivered by four models respectively. We can divide the values of the factor into three groups based on the values of HJ distances. When the factor values are less than 0, a nonparametric model can produce pricing errors smaller than those in others. When the values of the factor locate between 0 and 1, all models perform similarly. When the values of the factor increase to above 1, the nonparametric model dominates others. The parametric polynomial models work well only when the return on aggregate wealth is positive and less than 1 percent. In other cases, the parametric polynomial models cannot adequately describe the SDF proxies. This finding implies the existence of structure breaks of the SDF proxies for different economy situations, which demonstrates the merits of the use of local estimation on SDF proxies.

Figure 2 summarizes the curve of $p$-values for the nonparametric conditional HJ distance tests at all sample values of the single factor. The conditional HJ distance in the nonparametric model cannot be rejected if the return on aggregate wealth explain the SDF proxies very well locally. However, the figure clearly shows that the rejection occurs when the factor value is within the range of $[-2, 1]$, which implies that we may need more factors to estimate the SDF proxies for that period. For other periods, the single factor, the return on aggregate wealth, seems to be sufficient.

If we combine the findings in the above two figures, we can conclude that the poor performance of a linear (quadratic or cubic) model is largely due to the unknown form of nonlinearity in the model. In the asset pricing literature, researchers are keen on finding more factors to capture systematic risks. However, the reason for the failure of existing models is possibly due to missing the nonlinearity rather than missing factors. Although
finding more factors is important, avoiding such misspecification is also crucial to fairly judging SDF proxies we already have.

3.2 Variation of $a$

In Dittmar (2002), the linear combination of the labor return, $R_l$, and the market return, $R_m$, is defined as the the wealth return. The value of $a$ is estimated to be around 0.558. This result depends heavily on the utility function assumed in that model. In order to check the robustness of our claims, we try different values of $a$.

We set $a$ to be grid values between 0 and 1 with an increment of 0.01. We repeat the empirical analysis here as we did for the fixed $a$. That is to say we try different linear combinations of $R_l$ and $R_m$ as the factor in the model. Figure 3 plots the $p$-value over $R_l$ and $a$. The results show that the $p$-value of the square conditional HJ distance is decreasing with $a$. Figure 4 is obtained when we project the results on the $p$-value and $a$ plane. The figure demonstrates that the appropriate values of $a$ should be less than 0.2 to avoid rejecting the hypothesis that the corresponding conditional HJ distance is equal to zero for the whole sample of $R_l$. This is quite consistent with the evidence that equity may represent as little as one-ninth of aggregate wealth, a small proportion of total wealth relative to human capital.

[Figure 3 around here]

[Figure 4 around here]

4 Conclusion

This paper proposes a local GMM to estimate the SDF proxy by minimizing the squares of the conditional HJ distance. A HJ distance test based on a nonparametric estimator is derived to test whether or not the pricing error driven by the nonparametric estimator is zero. Our method can alleviate biases due to missing the nonlinearity in the model.
Empirical studies demonstrate that the return on aggregate wealth can adequately explain the SDF proxies in most cases as long as the nonlinearity is dealt with appropriately. Moreover, we find that the preferred functional form of SDF proxies varies with the status of an economy. Nonlinearity is more likely to occur when an economy is booming or shrinking.

References


Appendix

Proof of Equation 18

It suffices to show that $E[S_T(m(z_0))] = 0$ and $Th\Var[S_T(m(z_0))] = V$ as $Th \to \infty$ and $h \to 0$, and then the asymptotic normality follows directly from the standard kernel regression limiting distribution theory.

Note that $E[S_T(m(z_0))] = 1$.

\begin{align*}
E[S_T(m(z_0))] &= \frac{1}{Th} \sum_{t=1}^{T} E[w_t(m(z_0))K\left(\frac{z_t - z_0}{h}\right)] \\
&= \frac{1}{Th} \sum_{t=1}^{T} \int \int w_t(m(z_0))K\left(\frac{z_t - z_0}{h}\right)f(z_t, R_t)dz_t dR_t \\
&= \frac{1}{T} \sum_{t=1}^{T} \int \int w_t(m(z_0))K(u)f(z_0 + uh, R_t)dudR_t
\end{align*}

where $u = \frac{z_t - z_0}{h}$. Expanding $f(\cdot, \cdot)$ around $z_0$, we have $f(z_0 + uh, R_t) = f(z_0, R_t) + uhf'(z_0, R_t) + O_p(h^2)$. Then,

\begin{align*}
E[S_T(m(z_0))]
&= \int w_t f(z_0, R_t)\int K(u)du dR_t + \int hw_t f'(z_0, R_t)\int uK(u)du dR_t + O_p(h^2) \\
&= \int w_t f(z_0) f(R_t|z_0)dR_t + O_p(h^2) \\
&= E[w_t(m(z_0))f(z_0)] + O_p(h^2) \\
&= 0
\end{align*}

The last equality follows from the moment condition in (6). Note that the bias term in $O_p(h^2)$ is $\frac{h^2}{2} \int \int u^2 K(u)du \int w_t f''(z_0, R_t)dR_t$.

Next, we consider the variance term. Taking advantage of the IID assumption and \footnote{The IID assumption can be relaxed to $\alpha$-mixing and the result still holds; see Cai (2007).}
the zero mean, the variance can be expressed as

\[
Th\text{Var}[S_T(m(z_0))] = h\text{Var}[w_t(m(z_0))K(\frac{z_t - z_0}{h})]
\]
(A.9)

\[
= h \int \int w_t w_t' K^2(\frac{z_t - z_0}{h}) f(z_t, R_t) dz_t dR_t
\]
(A.10)

\[
= \int \int w_t w_t' K^2(u) f(z_0 + uh, R) du dR
\]
(A.11)

\[
\rightarrow \int w_t w_t' f(z_0, R_t) \int K^2(u) du dR_t
\]
(A.12)

\[
= \int w_t w_t' f(z_0) f(R_t|z_0) dR_t \int K^2(u) du
\]
(A.13)

\[
= E[w_t w_t' f(z_0) \int K^2(u) du]
\]
(A.14)

Proof of Equation 13

Define

\[
Q_T[\hat{m}(z_0)] = S_T[\hat{m}(z_0)]^' G_T(z_0)^{-1} S_T[\hat{m}(z_0)],
\]
(A.15)

\[
S_0[(m(z_0)] = E[w_t f(z_0)],
\]
(A.16)

and

\[
Q_0[m(z_0)] = E[w_t f(z_0)]^' G(z_0)^{-1} E[w_t f(z_0)].
\]
(A.17)

Suppose that \(w_t(m(z_0))\) is twice differentiable in \(m(z)\) for all \(m(z)\) in a compact set \(\Theta(z)\) and there exists a unique \(m(z_0) \subseteq \Theta(z)\) such that \(E[w_t f(z_0)] = 0\), then we can have \(S_T[(\hat{m}(z_0)]\) and \(Q_T[\hat{m}(z_0)]\) uniformly converges to \(S_0[(m(z_0)]\) and \(Q_0[m(z_0)]\) respectively over the compact set \(\Theta(z)\). Then the consistency of \(\hat{m}(z_0)\) can be obtained by following Lewbel (2007).

For the asymptotic normality, expanding the first order condition of Equation (12)
around \( m(z_0) \), we can have

\[
S_T(\hat{m}(z_0))' G_T^{-1}(z_0) [S_T(m(z_0)) + \frac{\partial S_T(\hat{m}(z_0))}{\partial m} (\hat{m}(z_0) - m(z_0))] = 0 \tag{A.18}
\]

where \( \tilde{m}(z_0) \) is a value between \( \hat{m}(z_0) \) and \( m(z_0) \). Note that

\[
\frac{\partial S_T(m(z_0))}{\partial m} = \frac{1}{Th} \sum_{t=1}^{T} R_t' K(\frac{z_t - z_0}{h}) \tag{A.19}
\]

\[
\rightarrow E[R_t' f(z_0)] \tag{A.20}
\]

as \( Th \to \infty \) and \( h \to 0 \). Thus, solving for \( (\hat{m}(z_0) - m(z_0)) \) in (A.14), we obtain

\[
\sqrt{Th}(\hat{m}(z_0) - m(z_0)) \to (D'G^{-1}D)^{-1} D'G^{-1} \sqrt{Th}S_T(m(z_0))(1 + o_p(\sqrt{Th})) \tag{A.21}
\]

The asymptotic normality now follows from Equation (18). Thus, as \( Th \to \infty \) and \( h \to 0 \), we can have

\[
\sqrt{Th} \left( \hat{m}(z_0) - m(z_0) - \frac{h^2}{2} \Delta \right) \to^d N[0, (D'G^{-1}D)^{-1} D'G^{-1} V G^{-1} D (D'G^{-1}D)^{-1}] \tag{A.22}
\]

where \( \Delta = (D'G^{-1}D)^{-1} D'G^{-1} [\int u^2 K(u)du \int w_t f''(z_0, R_t) dR_t] \). When \( Th^5 \to 0 \) and \( Th \to \infty \), the higher order bias \( \Delta \) can be ignored.
Figure 1: This figure reports the conditional HJ distances. The blue line is for the nonparametric model. The red, the green and the black lines are for the linear, the quadratic and the cubic models respectively.
Figure 2: This figure illustrates the $p-$values of the conditional HJ distances for the nonparametric model. The horizontal line denotes 5% level.
Figure 3: This figure illustrates the $p-$values of the conditional HJ distances over $R_t$ and $a$. 
Figure 4: This figure shows the projection of Figure 3 on $p$–value and $a$ plane.