A New Semiparametric Conditional Capital Asset Pricing Model with Variable Selection

Zongwu CAI\textsuperscript{a,b} Yu REN\textsuperscript{b}

\textsuperscript{a}University of North Carolina at Charlotte
\textsuperscript{b}WISE, Xiamen University

Present to WISE’s Summer School, July 15, 2011
Outline

1. Motivation and Literature Review
2. Functional-Coefficient Index CAPM
3. Monte Carlo Simulations
4. Empirical Analysis
5. Conclusion and Future Research
Motivation
Capital Asset Pricing Model (CAPM): Sharp (1964), Lintner (1965)

\[ E(R_i) - R_f = \beta_i [E(R_m) - R_f], \] (1)

where

- \( R_i \) is return of the \( i \)th asset; \( R_f \) is return of risk-free asset;
- \( E(R_m) - R_f \): market (factors) risk premium denoted by \( \gamma_1 \);
- \( \beta_i = \frac{Cov(R_i, R_m)}{Var(R_m)} \) is to measure the sensitivity of asset returns to factors.
Usefulness of CAPM

- It can explain the variation in average returns across assets.
- It means that assets with higher betas should get higher average returns.
- It can be used as a tool to measure the risk of an individual stock against the market portfolios or indices.
- In the portfolio management field, a factor model or CAPM can describe successfully market behavior.
Tests of CAPM

Define excess return $r_{i,t} = R_{i,t} - R_{f,t}$, where $R_{f,t}$ is risk free return,

- time series: $r_{i,t} = \alpha_i + \beta_i(R_{m,t} - R_{f,t}) + \varepsilon_{it}$ with $H_0: \alpha_i = 0$ which is testing pricing error. Please note that testing $H_0: \alpha_i = 0$ for each $i$ is much easier than testing $H_0: \alpha_i = 0$ for all $i$.

- cross section: $E[r_{i,t}] = \gamma_0 + \beta_i\gamma_1$ with $H_0: \gamma_0 = 0$. 
The question: How to estimate $\beta_i$ and $\gamma_1$?

- Black, Jensen, and Scholes (1972) proposed estimating the beta of each asset $i$ from the time series regression to obtain the estimate of $\beta_i$, denoted by $\hat{\beta}_i$, and then estimate the risk premium $\gamma_1$ from the cross-sectional regression as

$$\bar{r}_i = \gamma_0 + \gamma_1 \hat{\beta}_i + e_i,$$

where $r_{i,t}$ denotes excess return and $\bar{r}_i = \sum_{t=1}^{T} r_{i,t} / T$. 
Fama and MacBeth (1973) suggested an alternative method to estimate $\gamma_0$ and $\gamma_1$ as follows. For each period $t$, this involves estimating the parameters in the following cross-sectional regression

$$r_{i,t} = \gamma_{0,t} + \gamma_{1,t}\hat{\beta}_i + e_{i,t},$$

to obtain $\hat{\gamma}_{0,t}$ and $\hat{\gamma}_{1,t}$, and then $\hat{\gamma}_0$ and $\hat{\gamma}_1$ are obtained by averaging

$$\hat{\gamma}_0 = \frac{T}{T} \sum_{t=1}^{T} \hat{\gamma}_{0,t} / T,$$

$$\hat{\gamma}_1 = \frac{T}{T} \sum_{t=1}^{T} \hat{\gamma}_{1,t} / T.$$
Fama-French (Fama and French (1992, 1993)) three factor model

\[ R_{it} = \alpha_i + \beta_{i1} R_{mt} + \beta_{i2} \text{SMB}_{it} + \beta_{i3} \text{HML}_{it} + e_{it}, \]

where SMB stands for “small (market capitalization) minus big” and HML for “high (book-to-price ratio) minus low”; they measure the historic excess returns of small caps over big caps and of value stocks over growth stocks. These factors are calculated with combinations of portfolios composed by ranked stocks (BM ranking, Cap ranking) and available historical market data.
Factor Models


3. More factors models · · · ·

4. Factor augmented capital asset pricing models which are similar to the factor-augmented vector autoregressive model proposed by Bernanke, Boivin and Eliasz (2005, QJE). That is a classical three or four factors model plus a $\lambda'_{it} f_t$, where $f_t$ is the factor loadings based on factor analysis or principal component analysis.
Failure of CAPM

Pricing Errors are significant. In addition, CAPM can NOT explain

1. small stocks outperform large stocks;
2. firms with high book-to-market (B/M) ratios outperform those with low B/M ratios;
3. stocks with high returns in the previous year continue to outperform those with low prior returns;

see the papers by Jagannathan and Wang (1996) and Lewellen and Nagel (2006) for details.
Conditional CAPM, proposed by Jagannathan and Wang (1996) is believed to have an ability to capture the time variation from period to period as

\[ E(r_{it} \mid I_{t-1}) = \beta_{i,t-1} \gamma_{1,t-1}, \]  

(2)

where

- \( I_{t-1} \): information set up to time \( t - 1 \)
- \( \gamma_{1,t} \): conditional market risk premium.

Of course, there is a debate in the finance literature on which model ((1) or (2)) should be used in a real application; see Lewellen and Nagel (2006) and Cai, Li and Ren (2011, under investigation).
In the conditional CAPM,

$$\beta_{i,t-1} = \frac{Cov(R_{i,t}, R_{mt}|I_{t-1})}{Var(R_{mt}|I_{t-1})}$$

- depend on the information set $I_{t-1}$
- time-varying

Let me show you some evidences for time-varying risk premium and betas in real examples in the literature.
Empirical Evidence for Time-Varying Risk Premium

Figure: 1. Time-Varying Risk Premium (Ferson and Harvey (1991))
Empirical Evidence for Time-Varying Betas

Figure: 2. Time-Varying Betas in Different Sectors (Braun, Nelson, and Sunier (1995)). Four Sectors: Basic Industries; Capital Goods; Construction; Consumer Durables
Different Research Channels

1. $\beta_{i,t-1}$ is a function of time $t$
   - smooth: Lin and Teräsvirta (1994)

2. $\beta_{i,t-1}$ is a function of financial variables
   - unobservable latent variables: Ang and Chen (2007)

We follow the last approach.
Ferson and Harvey (1999) considered the following model

\[ r_{it} = \alpha_{it} + \beta_{it} r_{mt} + \varepsilon_{i,t}, \]  

(3)

where both coefficient functions \( \alpha_{it} \) and \( \beta_{it} \) change smoothly with \( Z_t \) and

- \( \alpha_{it} = \alpha_i(Z_t) = b_{0i} + b'_{1i} Z_t \)
- \( \beta_{it} = \beta_i(Z_t) = a_{0i} + a'_{1i} Z_t \)
- \( Z_t \): vector of mean zero information variables
- \( b_{0i}, b_{1i}, a_{0i} \) and \( a_{1i} \) are parameters to be estimated
Why do we use smooth structural changes?

1. leading driving forces exhibit evolutionary changes
2. economic agents usually update their beliefs continuously
3. the aggregated economic variables over many individuals may become smooth
4. some empirical evidences support smooth structural changes
The vector of financial variables used in Ferson and Harvey (1999) $Z_t$ contain:

1. The difference between the one-month lagged returns of a three-month and a one-month Treasury bill (Campbell (1987), Harvey (1989), Ferson and Harvey (1991))

2. The dividend yield of S&P 500 index (Fama and French (1988))
3. The default spread between Moody’s Baa and Aaa corporate bond yields (Keim and Stambaugh (1986) or Fama (1990))
4. The term spread between a ten-year and a one-year Treasury bond yields (Fama and French (1989))
5. The lagged value of a one-month Treasury bill yield (Fama and Schwert (1977), Ferson (1989), or Breen, Glosten, and Jagannathan (1989))
Model of linear time-varying betas and alphas used in Ferson and Harvey (1999) is

\[ r_{i,t+1} = \alpha_i(Z_t) + \beta_i(Z_t)r_{m,t+1} + \epsilon_{i,t+1}, \]

where \( r_{i,t} \) is excess returns, \( \alpha_i(Z_t) = a_{0i} + a'_{1i}Z_t \) and \( \beta_i(Z_t) = b_{0i} + b'_{1i}Z_t \).
Main results in Ferson and Harvey (1999) are summarized as:

- reject constant alphas and betas in the Fama-French three-factor model
- reject constant alphas and betas in the four-factor model advocated by Elton, Gruber and Blake (1995)
Problems in Ferson and Harvey (1999):

1. the assumption that the alphas and betas are linear in the financial variables is too strong (Wang (2002, 2003)). Indeed, Wang (2002, 2003) argued that $\beta_i(\cdot)$ is a nonlinear function of $Z_t$: the dividend price ratio (DPR), the default premium (DEF), the one-month Treasury bill rate (RTB), and the excess return on the NYSE equally weighted portfolio (EWR).

The results [see next three slides] in Wang (2003) suggest the presence of nonlinearity in conditional betas, as well as that the beta function differs significantly across test assets.

Figure 1. Plots of the market factor beta. The estimated beta function is estimated nonparametrically. The solid lines are plots of one-dimensional snapshots for the relation between the beta and one conditioning variable, while the other variables are fixed at their means. The plots are for different conditioning variables: Panel A (DPR), Panel B (DEF), Panel C (RTB), and Panel D (EWR).

Figure 2. Plots of the size factor beta. The multivariate beta function is estimated non-parametrically. The solid lines are plots of one-dimensional snapshots for the relation between the beta and one conditioning variable, while the other variables are fixed at their...

Figure 3. Plots of the book-to-market factor beta. The solid lines illustrate the nonparametric estimates. Panel A shows the relation between the book-to-market factor beta and the book-to-market ratio (DPR). Panel B shows the relation with the debt-to-equity ratio (DEF). Panel C shows the relation with the return on assets (RTB). Panel D shows the relation with the earnings to assets ratio (EWR).
Problems in Ferson and Harvey (1999)

2. Inference and estimation based on misspecification can be misleading (Ghysels (1998)). Actually, Ghysels (1998) demonstrated that the specific form of $\beta_i(\cdot)$ should be cautious.

3. Wang (2002, 2003) considered the problem under the mean-covariance efficiency and nonparametric method is used to estimate stochastic discount factor and betas. But, a pure nonparametric method is not reliable because of “curse of dimensionality”.

4. To overcome this problem, one might use an index model as in Aït-Sahalia and Brandt (2001).
Summary:

- conditional CAPM is widely used and cited
- literature does not provide a best way to estimate the time-varying beta
- literature does not provide a good method to select financial variables
- inference based on a wrong model can be misleading
Our Contributions:

- Relax the linearity assumption about the betas
- Select the financial variables by the data; delete the irrelevant financial variables automatically
- Combine the procedure of financial variables selection and model estimation at the same step
- Test the validity of conditional CAPM
Econometric Modeling
Suppose \( \{z_t, x_t, R_t\}_{t=1}^{+\infty} \) is a jointly strictly stationary process.

- \( z_t \): a univariate variable
- \( x_t \): \( p \times 1 \) vector of factors

Define

\[
m(z_0, x_0) = E(R_t | z = z_0, x = x_0),
\]

then

\[
m(z_0, x_0) = \sum_{j=1}^{p} \beta_j(z_0) x_{0j}.
\]

Therefore, the functional coefficient regression model can be written as

\[
R_t = \beta_1(z_t) + \sum_{j=2}^{p} \beta_j(z_t) x_{tj} + \varepsilon_t;
\]

see Cai, Fan and Yao (2000).
The local linear estimator is defined as $\hat{\beta}_j(z_0) = \hat{a}_j$, where $\{\hat{a}_j, \hat{b}_j\}$ minimize the following sum of the locally weighted least squares

$$(\hat{a}, \hat{b}) = \text{argmin}_{a,b} \sum_{t=1}^{n} \left[ R_t - \sum_{j=1}^{p} \{a_j + b_j(z_t - z_0)\} x_{tj} \right]^2 K_h(z_t - z_0),$$

where $K_h(\cdot) = K(\cdot/h)/h$, $K(\cdot)$ is a kernel function on $R^1$ and $h$ is a bandwidth. Note that when the number of predictors is large, the nonparametric method suffers from the so-called “curse of dimensionality”.

Zongwu CAI$^{a,b}$, Yu REN$^b$ A New Semiparametric Conditional Capital Asset Pricing Model with
To overcome this problem, by following Aït-Sahalia and Brandt (2001), assume that $z_t$ is a linear combination of financial variables $\xi_t$, where

- $z_t$: state variable; $\xi_t$: financial variables (the number of financial variables can be large);
- $z_t(c) = c'\xi_t$ with $||c|| = 1$ (identification condition);
- $c$ denotes the relative weights of the financial variables.

Then, the functional coefficient index CAPM model is

$$R_t = \beta_1(c'\xi_t) + \sum_{j=2}^{p} \beta_j(c'\xi_t) x_{tj} + \varepsilon_t. \quad (5)$$

Indeed, model (5) was studied by Fan, Yao and Cai (2003) for statistical standpoint.
Interpretation of Index

1. From a statistical perspective, the index avoids the curse of dimensionality because it allows us to reduce the multivariate problem to one where we can implement the nonparametric approach described above in a univariate setting (since $c'\xi_t$ is univariate).

2. From an economic standpoint, the index offers a convenient univariate summary statistic that describes the current state of the time-varying investment opportunities (various economic indicators).

3. From a normative perspective, our results can help investors with any preference(s) determine which economic variables they should track and, more importantly, in what single combination.
For the index $c' \xi_t$ given, we can estimate the betas $\beta_j(c' \xi_t)$ in (5). We can use the \textbf{local linear approach} so that the sum of the locally weighted least squares is given by

$$\begin{align*}
\begin{pmatrix} \hat{a}(c) \\ \hat{b}(c) \end{pmatrix} &= \arg\min_{a,b} \sum_{t=1}^{n} \left[ R_t - \sum_{j=1}^{p} \{ a_j + b_j(z_t(c) - z_0(c)) \} x_{tj} \right] \\
& \quad \times K_h(z_t(c) - z_0(c)),
\end{align*}$$

and $\hat{\beta}_j(c' \xi_t) = \hat{a}(c)$. 

\textit{Zongwu CAI}^{a,b}, \textit{Yu REN}^{b} A New Semiparametric Conditional Capital Asset Pricing Model with
Then, for given $\hat{\beta}_j(\cdot)$, we can estimate $c$ using the sum of squared residuals

$$Q(c) = \sum_{t=1}^{n} \left[ R_t - \sum_{j=1}^{p} \hat{\beta}_j(c'\xi_t)x_{tj} \right]^2. \quad (7)$$

The question in a real application is how to choose $\xi_{jt}$ which are significant in the model among a big pool of $\{\xi_{jt}\}$. In other worlds, insignificant $\xi_{jt}$ should be removed from the final model. This is a variable selection.
FCR with Variable Selection Method

Least Absolute Shrinkage and Selection Operator (LASSO): By adding Smoothly Clipped Absolute Deviation (SCAD) penalty to the sum of the squared residuals $Q(c)$, it becomes the penalized least squares

$$Q_p(c) = \min_c \left[ Q(c) + n \sum_{j=1}^{k} p_{\lambda,v}(|c_j|) \right], \quad (8)$$

where $p_{\lambda,v}(\cdot)$ is a penalty function and its first order derivative is

$$p'_{\lambda,v}(c_j) = \lambda I(c_j \leq \lambda) + \frac{(v \lambda - c_j)_+}{v - 1} I(c_j > \lambda),$$

where $v > 2$ and $\lambda > 0$. Of course, other types of penalty functions can be applicable.
The resulting estimator for $c_j$ is

$$
\hat{c}_j = \begin{cases} 
\text{sign}(\hat{c}_j^0)(\hat{c}_j^0 - \lambda), & \text{when } |\hat{c}_j^0| < 2\lambda; \\
(v - 1)\hat{c}_j^0 - \text{sign}(\hat{c}_j^0)v\lambda \right) / (v - 2), & \text{when } 2\lambda \leq |\hat{c}_j^0| \leq v\lambda; \\
\hat{c}_j^0, & \text{when } |\hat{c}_j^0| > v\lambda.
\end{cases}
$$

Therefore, we can see that if the true value of $c_j$ is zero, the estimated value should be exactly zero. In such a way, $\xi_{jt}$ should not be selected in $Z_t$. Hence, financial variables selection and model estimation can be done at the same step.
Implementation Algorithm

1. assign an initial value of $c^{(0)}$
2. choose an optimal value of $\lambda^{(0)}$ and $\nu^{(0)}$ based on the (generalized) cross validation method; see Fan and Li (2001) for details
3. regard the $\lambda^{(0)}$ and $\nu^{(0)}$ as given, and estimate $c^{(1)}$ by minimizing $Q(c)$
4. treat $c^{(1)}$ as the initial value, and choose an optimal value of $\lambda^{(1)}$, $\nu^{(1)}$
5. repeat the above procedures until the estimates converge
Consider a more general testing problem as

\[ H_0 : \text{some } \beta_j(Z_t) = \text{constant or zero.} \]

Apply the generalized likelihood ratio (GLR) test of Cai, Fan and Yao (2000) as follows: Compute the sum of squared residuals (SSR) under both null and alternative hypotheses \( H_0 \) and \( H_a \), respectively, denoted by \( \text{SSR}_0 \) and \( \text{SSR}_a \).

Compute the test statistics as

\[ J_n = \frac{[\text{SSR}_0 - \text{SSR}_a]}{\text{SSR}_a}. \]

Apply the wild bootstrap to the \( p \)-value.
The asymptotic theory for both estimation and testing is under investigation; see Cai and Yang (2011) for details.
A Simple Simulation
Purpose

- check the validity of the model
- illustrate the financial variables selection
Data Generating Process: \( R_t = \beta_1(z_t) + \beta_2(z_t)x_t + e_t \), where

- \( \beta_1(z_t) = 0, \beta_2(z_t) = z_t \), and \( z_t = c_1 \xi_{1t} + c_2 \xi_{2t} + c_3 \xi_{3t} \),
- \( \xi_{1t}, \xi_{2t} \) and \( \xi_{3t} \) are generated from \( N(0, 0.38) \),
- \( x_t \) is the factor \( \sim U[0, 1] \),
- \( e_t \) is error term \( \sim N(0, 0.08) \)
- true values for \( c \) are \( c_1 = c_2 = 1/\sqrt{2} \approx 0.7 \), and \( c_3 = 0 \),
- the sample size is \( n = 300, 500, 1000 \).
- the simulation is repeated 1000 times for each sample size.
To measure the performance of the proposed method, we compute the absolute deviation error (ADE) for $\hat{c}_j$ as

$$\text{ADE}_j = |\hat{c}_j - c_j|$$

for $j = 1$ and 2. For $c_3$, it is measured by the shrinkage rate, defined as

$$\text{Shrinkage rate} = \text{number of } \{\hat{c}_3 = 0\} / 1000.$$
## Simulation Result

<table>
<thead>
<tr>
<th>Table: 1. Estimation Results: ADE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$c_1$</th>
<th>$c_2$</th>
<th>$c_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>N=300</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>median</td>
<td>0.0022</td>
<td>0.0022</td>
<td>0.0000</td>
</tr>
<tr>
<td>std</td>
<td>0.0017</td>
<td>0.0017</td>
<td>0.0000</td>
</tr>
<tr>
<td>shrinkage rate</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>N=1000</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>median</td>
<td>0.0014</td>
<td>0.0014</td>
<td>0.0000</td>
</tr>
<tr>
<td>std</td>
<td>0.0010</td>
<td>0.0010</td>
<td>0.0000</td>
</tr>
<tr>
<td>shrinkage rate</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Nonlinear Data Generating Process:

\[ R_t = 0.1 \exp(0.7071u_{1t} + 0.7071u_{2t})x_t + e_t, \]
## Simulation Result

### Table: 2. Estimation Results: ADE

<table>
<thead>
<tr>
<th></th>
<th>$c_1$</th>
<th>$c_2$</th>
<th>$c_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=300</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>median</td>
<td>0.0180</td>
<td>0.0160</td>
<td>0.0000</td>
</tr>
<tr>
<td>std</td>
<td>0.0186</td>
<td>0.0196</td>
<td>0.0753</td>
</tr>
<tr>
<td>shrinkage rate</td>
<td><strong>80%</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N=1000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>median</td>
<td>0.0106</td>
<td>0.0107</td>
<td>0.0000</td>
</tr>
<tr>
<td>std</td>
<td>0.0059</td>
<td>0.0073</td>
<td>0.0000</td>
</tr>
<tr>
<td>shrinkage rate</td>
<td><strong>98%</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Simulation Result: Size of Test

*Table: 3. Size of Test*

We use the linear and nonlinear models described previously to generate the data, and test the null hypothesis that the pricing error is insignificant.

<table>
<thead>
<tr>
<th>Nominal Level (%)</th>
<th>1</th>
<th>5</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>T=100</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear Model (%)</td>
<td>0.3</td>
<td>3.3</td>
<td>7.6</td>
</tr>
<tr>
<td>Nonlinear Model (%)</td>
<td>1.3</td>
<td>3.7</td>
<td>9.2</td>
</tr>
<tr>
<td><strong>T=300</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear Model (%)</td>
<td>0.7</td>
<td>3.9</td>
<td>8.2</td>
</tr>
<tr>
<td>Nonlinear Model (%)</td>
<td>0.9</td>
<td>5.7</td>
<td>10.9</td>
</tr>
<tr>
<td><strong>T=500</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear Model (%)</td>
<td>1.2</td>
<td>4.6</td>
<td>9.8</td>
</tr>
<tr>
<td>Nonlinear Model (%)</td>
<td>1.1</td>
<td>4.8</td>
<td>9.5</td>
</tr>
</tbody>
</table>
In order to check the power of the test, we add a local time-varying intercept to the linear and the nonlinear models described before. For the linear model,

\[ y_t = \frac{c_0}{T^{2/5}} (c_1 z_{1t} + c_2 z_{2t}) + (c_1 z_{1t} + c_2 z_{2t}) x_t + e_t, \]

and for the nonlinear model,

\[ y_t = \frac{c_0}{T^{2/5}} (c_1 z_{1t} + c_2 z_{2t}) + 0.1 \exp(c_1 z_{1t} + c_2 z_{2t}) x_t + e_t. \]

We obtain the rejection frequency based on 1,000 replications.
We use the linear and nonlinear models, which contain a local pricing error, to generate the data, and test the null hypothesis that the pricing error is insignificant at 5%.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$c_0=0.1$</th>
<th>$c_0=0.15$</th>
<th>$c_0=0.2$</th>
<th>$c_0=0.25$</th>
<th>$c_0=0.3$</th>
<th>$c_0=0.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T=100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear (%)</td>
<td>10.5</td>
<td>29</td>
<td>43.5</td>
<td>63.5</td>
<td>79.5</td>
<td>98.0</td>
</tr>
<tr>
<td>Nonlinear (%)</td>
<td>11.0</td>
<td>25</td>
<td>42</td>
<td>62</td>
<td>83.5</td>
<td>97.5</td>
</tr>
<tr>
<td>T=300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear (%)</td>
<td>22.0</td>
<td>45.5</td>
<td>70.5</td>
<td>86.5</td>
<td>95.5</td>
<td>99.5</td>
</tr>
<tr>
<td>Nonlinear (%)</td>
<td>17.0</td>
<td>46.5</td>
<td>70.5</td>
<td>86.5</td>
<td>96.0</td>
<td>100</td>
</tr>
<tr>
<td>T=500</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear (%)</td>
<td>26.5</td>
<td>49</td>
<td>76</td>
<td>91.5</td>
<td>97.5</td>
<td>100</td>
</tr>
<tr>
<td>Nonlinear (%)</td>
<td>25.5</td>
<td>50</td>
<td>76</td>
<td>95.5</td>
<td>99</td>
<td>100</td>
</tr>
</tbody>
</table>
Empirical Analysis
Data

- Excess monthly Fama-French 25 portfolios: July 1963 to December 2009
- Fama-French portfolios formed based on the size and book-to-market ratio of the firms.
- B1: the smallest ratio; B5: the largest ratio; S1: the smallest size; S5: the largest size
Instrument Variables

1. treasury bill yield: spread between the returns of a three-month and a one-month Treasury bill
2. default spread: the yield between Moody’s Baa and Aaa corporate bonds
3. term spread: the yield between a ten-year and one-year Treasury bonds
4. one-month Treasury bill yield
5. other variables such as capital-labor ratio, firm characteristics, liquidity risk, SMB, MOM, HML, · · ·
Data Smoothing

- use our model to estimate the time-varying betas for each portfolio
- collect the residuals of the time series regression for portfolio returns
- report the mean square error
- repeat the estimation by using the model in Ferson and Harvey (1999)
- compare MSE delivered by two models
- Ferson and Harvey (1999) have already rejected the Fama-French model based on the evidence of time-varying betas. So we do not compare our results with the FF model.
## Selective Report on Data Smoothing

<table>
<thead>
<tr>
<th></th>
<th>FCR+SCAD</th>
<th>Ferson and Harvey</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1/B1</td>
<td>22.4948</td>
<td>23.3314</td>
</tr>
<tr>
<td>S1/B3</td>
<td>10.7763</td>
<td>12.048</td>
</tr>
<tr>
<td>S1/B5</td>
<td>12.3372</td>
<td>13.4097</td>
</tr>
<tr>
<td>S2/B3</td>
<td>6.4449</td>
<td>7.3834</td>
</tr>
<tr>
<td>S3/B5</td>
<td>8.1348</td>
<td>9.2337</td>
</tr>
<tr>
<td>S4/B1</td>
<td>4.8292</td>
<td>5.1875</td>
</tr>
<tr>
<td>S4/B3</td>
<td>4.1486</td>
<td>4.6364</td>
</tr>
<tr>
<td>S4/B5</td>
<td>7.5399</td>
<td>8.383</td>
</tr>
<tr>
<td>S5/B1</td>
<td>2.3848</td>
<td>2.7164</td>
</tr>
<tr>
<td>S5/B5</td>
<td>8.4341</td>
<td>9.0278</td>
</tr>
</tbody>
</table>
State Variables

- state variable: the linear combination of the instruments
- try different methods to estimate state variables
- use FCR to do estimation after obtaining the state variables
- compare the goodness-of-fit
Models to Obtain State Variables

1. SCAD

2. OLS: \( r_{it+1} = b_{0i} + b'_{1i}Z_t + \varepsilon_{i,t+1} \)
\[ \Rightarrow \hat{b}_{0i} + \hat{b}'_{1i}Z_t \]

3. Ferson and Harvey:
\( r_{it+1} = (a_{0i} + a'_{1i}Z_t) + (b_{0i} + b'_{1i}Z_t)r_{p,t+1} + \varepsilon_{i,t+1} \)
\[ \Rightarrow \hat{b}_{0i} + \hat{b}'_{1i}Z_t \]
## Table: 6. Mean Square Error

<table>
<thead>
<tr>
<th></th>
<th>SCAD</th>
<th>OLS</th>
<th>Ferson and Harvey</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1/B1</td>
<td>22.4948</td>
<td>23.4331</td>
<td>23.1556</td>
</tr>
<tr>
<td>S1/B3</td>
<td>10.7763</td>
<td>11.7326</td>
<td>11.6373</td>
</tr>
<tr>
<td>S1/B5</td>
<td>12.3372</td>
<td>13.4502</td>
<td>13.3691</td>
</tr>
<tr>
<td>S2/B3</td>
<td>6.4449</td>
<td>7.2664</td>
<td>7.2051</td>
</tr>
<tr>
<td>S2/B5</td>
<td>9.8359</td>
<td>10.6718</td>
<td>10.6696</td>
</tr>
<tr>
<td>S3/B3</td>
<td>4.9300</td>
<td>5.0807</td>
<td>5.0757</td>
</tr>
<tr>
<td>S3/B5</td>
<td>8.1348</td>
<td>9.0188</td>
<td>8.7943</td>
</tr>
<tr>
<td>S4/B2</td>
<td>3.3935</td>
<td>3.5223</td>
<td>3.4804</td>
</tr>
<tr>
<td>S4/B4</td>
<td>4.5217</td>
<td>4.8917</td>
<td>4.9957</td>
</tr>
<tr>
<td>S5/B1</td>
<td>2.3848</td>
<td>2.6724</td>
<td>2.5847</td>
</tr>
<tr>
<td>S5/B3</td>
<td>3.6080</td>
<td>3.9656</td>
<td>3.8562</td>
</tr>
<tr>
<td>S5/B5</td>
<td>8.4341</td>
<td>8.7039</td>
<td>8.6781</td>
</tr>
</tbody>
</table>
The null hypothesis is

\[ H_0 : \alpha_t = a_0 + a'_1 Z_t, \quad \beta_t = b_0 + b'_1 Z_t \] (9)

and the alternative hypothesis is

\[ H_1 : \alpha_t = f_1(c' Z_t), \quad \beta_t = f_2(c' Z_t). \] (10)

This test is about testing whether the model in Ferson and Harvey (1999) is appropriate. Of course, other types of testing hypotheses can be considered in the same fashion.
### Table: 7. $p$–Values for Testing Linearity.

<table>
<thead>
<tr>
<th></th>
<th>B1</th>
<th>B2</th>
<th>B3</th>
<th>B4</th>
<th>B5</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>0.2500</td>
<td>0.1954</td>
<td>0.0002</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>S2</td>
<td>0.2082</td>
<td>0.0125</td>
<td>0.0011</td>
<td>0.0001</td>
<td>0.0088</td>
</tr>
<tr>
<td>S3</td>
<td>0.2344</td>
<td>0.0412</td>
<td>0.0168</td>
<td>0.0248</td>
<td>0.0038</td>
</tr>
<tr>
<td>S4</td>
<td>0.0050</td>
<td>0.0002</td>
<td>0.0136</td>
<td>0.0184</td>
<td>0.0006</td>
</tr>
<tr>
<td>S5</td>
<td>0.2508</td>
<td>0.0360</td>
<td>0.0702</td>
<td>0.2628</td>
<td>0.0806</td>
</tr>
</tbody>
</table>
The null hypothesis is

$$H_0 : \alpha_t = 0$$  \hspace{1cm} (11)

and the alternative hypothesis is

$$H_1 : \alpha_t \neq 0.$$  \hspace{1cm} (12)
Test Pricing Error: Time Series Analysis

**Table:** 8. $p$–Values for Testing Pricing Errors.

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>B1</th>
<th>B2</th>
<th>B3</th>
<th>B4</th>
<th>B5</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>0.6840</td>
<td>0.5480</td>
<td>0.0020</td>
<td>0.0000</td>
<td>0.0002</td>
</tr>
<tr>
<td>S2</td>
<td>0.6520</td>
<td>0.3980</td>
<td>0.0100</td>
<td>0.1500</td>
<td>0.1060</td>
</tr>
<tr>
<td>S3</td>
<td>0.0900</td>
<td>0.2620</td>
<td>0.5300</td>
<td>0.0300</td>
<td>0.1480</td>
</tr>
<tr>
<td>S4</td>
<td>0.0860</td>
<td>0.0800</td>
<td>0.7020</td>
<td>0.0500</td>
<td>0.4880</td>
</tr>
<tr>
<td>S5</td>
<td>0.5400</td>
<td>0.0940</td>
<td>0.6160</td>
<td>0.6660</td>
<td>0.5480</td>
</tr>
</tbody>
</table>
Cross Sectional Regression

\begin{equation}
E[R_{it} | I_{t-1}] = \gamma_{1t-1} \beta_{it-1},
\end{equation}

where $\gamma_{1t-1}$ is the conditional market risk premium.
Following Ferson and Harvey (1999), the cross-sectional regression can be setup as

\[ R_{it} = \gamma_{0,t-1} + \gamma_{1,t-1} \beta_{it-1} + \gamma_{2,t-1} \delta_{it-1}' Z_{t-1} + e_{it}, \]

\( \delta_{it-1}' Z_{t-1} \) is the fitted conditional expected return, where \( \delta_{it-1} \) is estimated by regressing the return on the lagged variable \( Z \), using the data up to time \( t - 1 \).
Cross Sectional Regression

- test the hypothesis that the pricing errors are jointly insignificant
- if our model successfully captures the systematic risks in the market, the pricing errors should be insignificant
- both $\gamma_0$ and $\gamma_2$ should be insignificant
Table: 9. Cross-Sectional Regression

We run the cross-sectional regression of the excess returns on constant, the time-varying beta and fitted conditional expected return. The time-series averages of the cross-sectional regression coefficients are shown in the first row in each panel and the corresponding Fama-MacBeth \( t \)-ratios are reported in the second row with parenthesis.

<table>
<thead>
<tr>
<th></th>
<th>( \gamma_0 )</th>
<th>( \gamma_1 )</th>
<th>( \gamma_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rolling Sample</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.6815</td>
<td>-0.0057</td>
<td>0.8408</td>
</tr>
<tr>
<td></td>
<td>(0.1171)</td>
<td>(-0.0010)</td>
<td>(0.1622)</td>
</tr>
<tr>
<td><strong>Expanding Sample</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.8599</td>
<td>-0.3111</td>
<td>0.9330</td>
</tr>
<tr>
<td></td>
<td>(0.1166)</td>
<td>(-0.0437)</td>
<td>(0.2138)</td>
</tr>
</tbody>
</table>

Zongwu CAI\(^{a,b}\), Yu REN\(^{b}\)
In conclusion,

1. Our model outperforms the other models in literature for estimating the conditional CAPM.
2. We cannot reject the conditional CAPM.
3. Anomalies almost disappear in the proposed conditional CAPM.
Future Research Topics related to this talk:

1. The theoretical econometric issues as mentioned earlier.
2. The model selection idea can be applied to other applications.
3. The testing issues.
4. Our estimation procedure does not take care of the correlations among assets. When we consider a joint test among assets, it is important to consider the correlations among assets.
5. Our study supports using a conditional CAPM. But the question arises is how to classify conditional and unconditional CAPM in a real application. Cai, Li and Ren (2011, under investigation) is working on finding a criterion to identify which model should be used in a real application.
Future Research Topics related to this talk:

1. Unit root issue for financial variables. **WHY?** Let us look at the following graphs.

2. This issue is still open and it is very interesting and worth to explore it. You might use the idea from the paper by Cai, Li and Park (2009, JoE).
Unit Root Evidence for Financial Variables: Aït-Sahalia and Brandt (2001, JoF)

Figure 1. Predictors. This figure shows time-series plots and autocorrelograms of four predictors: the default spread, the log dividend-to-price ratio of the S&P 500 index, the term spread, and the S&P index trend variable. The data is sampled monthly from January 1954 through December 1997. There are 528 observations.
Questions or Comments or Suggestions?
Thank You for Coming!