Time-Varying Beta and the Value Premium

Hui Guo, Chaojiang Wu, Yan Yu*

First Draft: October 2013

This Version: September 2016

Abstract

We model conditional market beta and alpha as flexible functions of state variables identified via a formal variable selection procedure. In the post-1963 sample, beta of the value premium comoves strongly with unemployment, inflation, and price-earnings ratio in a countercyclical manner. We also uncover a novel nonlinear dependence of alpha on business conditions: It falls sharply and becomes even negative during severe economic downturns but is positive and flat otherwise. Conditional CAPM performs better than unconditional CAPM but does not fully explain the value premium. Our findings are consistent with a conditional CAPM with rare disasters.

Key Words: Conditional CAPM; Penalized Splines; Single-Index Models; Value Premium; Variable Selection.

---

*Hui Guo is Briggs Swift Cunningham Professor of the Department of Finance and Real Estate at the University of Cincinnati, PO BOX 210130, Cincinnati, OH, 45221, email: Hui.Guo@uc.edu. Chaojiang Wu is Assistant Professor of the Decision Sciences Department at Drexel University, Philadelphia PA 19104, email: Chaojiang.Wu@drexel.edu. Yan Yu is the Joseph S. Stern Professor of Business Analytics at the University of Cincinnati, PO BOX 210130, Cincinnati, OH, 45221, email: Yan.Yu@uc.edu. We thank the editor, Hank Bessembinder, an anonymous referee, and Stefan Nagel for helpful comments.

1 Introduction

In the post-1963 sample, value stocks have smaller market beta but higher average returns than do growth stocks. An important risk-based explanation is that the value premium, the return difference between value and growth stocks, is riskier during business recessions when conditional equity premium is high than during business expansions when conditional equity premium is low (e.g., Lettau and Ludvigson (2001) and Zhang (2005)). Existing empirical evidence that the value premium has a countercyclical market beta is, however, inconclusive (e.g., Ang and Kristensen (2012)). We shed new light on this issue using a novel statistical model that allows us to tackle two possible misspecifications in previous studies.

First, earlier authors assume that the conditional beta and alpha are a *linear* function of state variables, although such a relation may be more complex, as Bai, Hou, Kung, and Zhang (2015) highlight in their investment-based model with rare disasters. Extant studies provide little empirical evidence on this potentially restrictive assumption due to the lack of an appropriate statistical method. We fill the gap by proposing a single-index varying-coefficient setup. Our empirical model allows for a flexible data-determined functional dependence of the conditional beta and alpha on an index that is a linear function of some state variables. The single-index varying-coefficient model nests previous linear models as a special case.

Second, only time series variation in market beta that comoves with conditional equity premium and variance drives a wedge between the unconditional and conditional CAPM, and earlier authors use standard stock market return and variance predictors as state variables. This approach is potentially problematic because Welch and Goyal (2008) and Paye (2012) show that many of these variables have rather weak predictive power for stock market returns and stock market variance, respectively. In addition, Ghysels (1998), Harvey (2001), and Cooper and Gubellini (2011) have cautioned that the estimation of conditional factor models can be sensitive to the choice of state variables. To address these issues, we identify the most relevant state variables from a comprehensive pool via a formal exhaustive variable selection procedure. Investment-based models, e.g., Zhang (2005), stress a direct link between the conditional beta and business fluctuations. To investigate this conjecture, we include closely watched gauges of aggregate economic activity as candidate state variables, in addition to commonly used stock market return and variance predictors.

The single-index modeling is a popular nonparametric approach that alleviates the curse of dimensionality (e.g., Carroll, Fan, Gijbels, and Wand (1997), Yu and Ruppert (2002)). It is a natural fit of our empirical specification. We can think of the single index as a composite measure of business conditions that are the main driver of conditional equity premium and variance in risk-based asset pricing models. When the conditional beta is a monotonic function of the single index, the interpretation of the single-index coefficients is similar to that for linear models. That is, we can test the hypothesis that the conditional beta changes countercyclically across time by simply inspecting the signs of the single-index coefficients. We make several novel extensions to the existing single-index varying-coefficient model (Xia and Li (1999), and Fan, Yao, and Cai (2003)) when using it to estimate and test the conditional CAPM. First, to account for heteroskedasticity in stock returns, we model the conditional volatility as a flexible function of the single index. Second, to facilitate the exhaustive variable selection procedure, we develop a computationally expedient iterative fixed point algorithm using penalized splines for estimation. Last, we derive asymptotic distributions for test statistics of the conditional CAPM.

We document strong countercyclical variation in the value premium's market beta over the July 1963 to December 2012 period. The unemployment rate and the inflation rate, the two most closely monitored macroeconomic variables by the Federal Reserve, are selected as the state variables for the conditional beta. Moreover, the conditional beta correlates negatively with the procyclical inflation rate and correlates positively with the countercyclical unemployment rate. To the best of our knowledge, the finding of a close relation between the value premium's conditional beta and key aggregate economic activity measures is novel.

Among commonly used stock market return predictors, the price-earnings ratio is identified as a significant state variable. Again, the conditional beta correlates negatively with the procyclical price-earnings ratio. In Figure 1, we plot the fitted conditional beta, with shaded areas indicating business recessions dated by the National Bureau of Economic Research (NBER). Consistent with risk-based explanations, it increases sharply during business recessions and decreases during business expansions. The countercyclical beta helps explain the value premium. The alpha is 5.6% per annum for the unconditional CAPM, while it decreases by more than 20% to 4.4% per annum for conditional CAPM. For comparison, the alpha is 5.3% per annum when using standard predictors, i.e., the dividend yield, the default premium, the term premium, and the risk-free rate, as the conditioning variables. The conditional CAPM, however, does not fully account for the value premium either; its alpha is statistically significant at the 1% level.

The conditional CAPM fails to fully explain the value premium possibly because we omit some important state variables. To address partially this underconditioning bias, we follow the advice by Boguth, Carlson, Fisher, and Simutin (2011) and include the value premium's lagged realized market beta as a candidate state variable. The variable selection procedure identifies the lagged realized beta as a significant conditioning variable in the conditional CAPM model for the value premium. The realized beta, however, does not subsume the information content of financial and macroeconomic variables that we consider. Specifically, the inflation rate and the price-earnings ratio remain significant conditioning variables. Again, we find that (1) the value premium's conditional beta changes countercyclically across time and (2) the conditional CAPM does not fully explain the value premium.

Our main results are qualitatively similar for the full sample spanning the January 1927 to December 2012 period. Figure 2 shows that the value premium's conditional beta increases during business recessions and decreases during business expansions. Campbell and Vuolteenaho (2004) and others find that the unconditional CAPM explains the value premium over the January 1927 to December 1963 period. Figure 2 shows that the value premium's market beta increases drastically during the Great Depression. This result suggests that the CAPM explains the value premium in the pre-1963 period possibly because, as Bai, Hou, Kung, and Zhang (2015) point out, the rare disaster risk that accounts for a substantial portion of the value premium materializes during the Great Depression.

Wang (2003) adopts a fully nonparametric model and uncovers a nonlinear dependence of the conditional beta on prespecified state variables. In contrast, using a single-index model that is less susceptible to the curse of dimensionality and allows for variable selection and time-varying volatility, we fail to reject the null hypothesis that the value premium's conditional beta is a linear function of the selected state variables. Interestingly, there is, however, novel evidence that the conditional alpha depends nonlinearly on the single index. Figure 3 shows that it falls sharply and becomes even negative during severe business downturns but is significantly positive and essentially flat otherwise. Ang and Kristensen (2012) use daily data to estimate the conditional CAPM in which the conditional beta is a nonparametric function of calendar time. In contrast with our study, Ang and Kristensen do not find countercyclical variation in the value premium's conditional beta.

Our novel evidence of a strong correlation of the conditional beta with key aggregate economic activity gauges suggests that investment-based models remain a viable explanation of the value premium. Bai, Hou, Kung, and Zhang (2015) propose such a model in which the conditional CAPM fails to explain the value premium in small samples because investors face rare disaster risk. The observed nonlinear dependence of the conditional alpha on business cycles is also a salient feature of their model.

The remainder of the paper proceeds as follows. We first explain the statistical methodology in Section 2. We then discuss data in Section 3. We present empirical findings in Section 4. We offer some concluding remarks in Section 5. Some technical materials are relegated to the Appendices.

2 Statistical Methodology

We propose a novel single-index varying-coefficient model for the conditional CAPM:

$$R_{t+1} = \alpha(\mathbf{z}_t\boldsymbol{\gamma}) + \beta(\mathbf{z}_t\boldsymbol{\gamma})R_{m,t+1} + \sigma(\mathbf{z}_t\boldsymbol{\gamma})\epsilon_{t+1}, t = 1, \dots, n, \tag{1}$$

where $\{R_{t+1}\}_{t=1}^n$ and $\{R_{m,t+1}\}_{t=1}^n$ denote respectively the series of the value premium and the excess stock market return observed at n time points. The conditional alpha $\alpha(\cdot)$ and the conditional beta $\beta(\cdot)$ are flexible functions of the so called single index $\mathbf{z}_t\boldsymbol{\gamma}$, which is a linear function of d-dimensional state variables $\mathbf{z}_t$, $\boldsymbol{\gamma}$ being the vector of single-index parameters. For identifiability, $\|\boldsymbol{\gamma}\| = 1$ and its first nonzero element is positive. ϵ_{t+1} is the error term with zero mean and unit variance. The conditional volatility $\sigma(\mathbf{z}_t\boldsymbol{\gamma})$ is also a flexible function of the single index. ϵ_{t+1} is independent of $\mathbf{z}_s$ and $R_{m,s+1}$ for $s \leq t$, and ϵ_{t+1} is independent of $\epsilon_{t'+1}$ for $t \neq t'$. The single-index varying-coefficient model nests the previous linear models, e.g., Petkova and Zhang (2005), as special cases, and we develop new statistical test for the functional form of alpha and beta, e.g., whether they are constant, linear, or nonlinear.

Compared with the fully nonparametric specification in which $\alpha(\cdot)$, $\beta(\cdot)$, and $\sigma(\cdot)$ are fully nonparametric functions of d-dimensional state variable vector $\mathbf{z}_t$, the model in equation (1) has two distinct advantages. First, single-index modeling overcomes the "curse of dimensionality" via projecting the d-dimensional state space to the one-dimensional single index. This advantage is especially important for our empirical investigation because we consider a moderately large number of state variables for a relatively short sample. Second, the single-index specification is consistent with our maintained hypothesis that the value premium's market beta comoves with business cycles. Specifically, we can think of the single index as a composite measure of business conditions, which, as Zhang (2005) and others show in their theoretical models, are the main driver of the value premium's conditional beta.

We develop a new estimation method by combining (log) penalized splines or P-splines and a fixed point algorithm. P-splines (e.g., Ruppert, Wand, and Carroll (2003), and Jarrow, Ruppert, and Yu (2004)) are used to estimate the unknown coefficient functions $\alpha(\cdot)$ and $\beta(\cdot)$, while log P-splines are used to estimate the non-negative volatility function $\sigma(\cdot)$. Fixed point algorithm (Cui, Härdle, and Zhu (2011)) is adapted to estimate the single-index parameters. In our experience, the fixed point algorithm based on P-splines is computationally fast and statistically robust in the estimation of the unknown flexible functions and single-index parameters. These properties are especially desirable when we conduct an exhaustive search from a comprehensive pool to identify the most relevant state variables for modeling the conditional beta.[1]

2.1 P-spline Estimation

Both varying coefficients $\alpha(u)$ and $\beta(u)$ are univariate flexible functions of the single index $u = \mathbf{z}\boldsymbol{\gamma}$ and can be estimated using P-splines. To impose the constraint of positive volatility, we estimate the volatility function $\sigma(u)$ using log P-splines, i.e., the logarithmic transformation of the volatility function is modeled using P-splines (e.g., Yu, Yu, Wang, and Li (2009)). We illustrate P-splines using truncated power basis for its easy interpretation. Other bases such as B-splines can be similarly adopted. The truncated power basis of degree p, with knots at $v_1, \ldots, v_k$, is $\mathbf{B}(u) = (1, u, u^2, \ldots, u^p, (u - v_1)_+^p, \ldots, (u - v_k)_+^p)$, where $(u - v_k)_+^p$ equals $(u - v_k)^p$ if $u \geq v_k$ and equals 0 otherwise. Any function with $p - 1$ continuous derivatives can be approximated by

$$\delta_0 + \delta_1 u + \delta_2 u^2 + \ldots + \delta_p u^p + \delta_{p+1}(u - v_1)_+^p + \ldots + \delta_{p+k}(u - v_k)_+^p = \mathbf{B}\boldsymbol{\delta},$$

where $\boldsymbol{\delta} = (\delta_0, \delta_1, \ldots, \delta_{p+k})^\mathsf{T}$ is the spline coefficient vector.

Denote the bases $\mathbf{B}_a$ and $\mathbf{B}_b$ and the spline coefficient vectors $\boldsymbol{\delta}_a$ and $\boldsymbol{\delta}_b$ for conditional alpha and beta respectively. Then at u_t the conditional alpha is $\alpha(u_t) = \mathbf{B}_{a,t}\boldsymbol{\delta}_a$ and the conditional beta is $\beta(u_t) = \mathbf{B}_{b,t}\boldsymbol{\delta}_b$. Combining the scalar 1 and $R_{m,t+1}$ with spline bases,

[1]For d candidate state variables, the number of all possible combinations is 2^d, given that the constant term is also a possible choice. Therefore, the exhaustive search can be computationally demanding when d is relatively large.

we can write the t-th row of "design" matrix $\mathbf{X}$ as $\mathbf{X}_t = (\mathbf{B}_{a,t}, \mathbf{B}_{b,t} R_{m,t+1})$. Define spline coefficient parameters $\boldsymbol{\delta}_1 = (\boldsymbol{\delta}_a^\top, \boldsymbol{\delta}_b^\top)^\top$, the mean function of R_{t+1} $(t = 1, 2, \dots, n)$ can be written as

$$m_t \equiv E(R_{t+1}) = \mathbf{B}_{a,t} \boldsymbol{\delta}_a + \mathbf{B}_{b,t} R_{m,t+1} \boldsymbol{\delta}_b = \mathbf{X}_t \boldsymbol{\delta}_1.$$

Similarly, we approximate the volatility function using $log\sigma(u_t) = \mathbf{B}_2(u_t)\boldsymbol{\delta}_2$, or equivalently, $\sigma(u_t) = exp\{\mathbf{B}_2(u_t)\boldsymbol{\delta}_2\}$. Then the penalized negative log-likelihood function is

$$\sum_{t=1}^{n} \Big(exp\{-2\mathbf{B}_2(\mathbf{z}_t\boldsymbol{\gamma})\boldsymbol{\delta}_2\}\{R_{t+1} - \mathbf{B}_{a,t}(\mathbf{z}_t\boldsymbol{\gamma})\boldsymbol{\delta}_a - \mathbf{B}_{b,t}(\mathbf{z}_t\boldsymbol{\gamma})\boldsymbol{\delta}_b R_{m,t+1}\}^2 + 2\mathbf{B}_2(\mathbf{z}_t\boldsymbol{\gamma})\boldsymbol{\delta}_2 \Big) + \frac{n}{2}\lambda_a \boldsymbol{\delta}_a^\top \mathbf{D}_a \boldsymbol{\delta}_a + \frac{n}{2}\lambda_b \boldsymbol{\delta}_b^\top \mathbf{D}_b \boldsymbol{\delta}_b + \frac{n}{2}\lambda_2 \boldsymbol{\delta}_2^\top \mathbf{D}_2 \boldsymbol{\delta}_2, \tag{2}$$

where $\lambda_a > 0$, $\lambda_b > 0$, and $\lambda_2 > 0$ are penalty or smoothing parameters for varying coefficient functions and volatility function respectively; $\mathbf{D}_\xi$ $(\xi = a, b, 2)$ is a 0-1 diagonal matrix such that $\boldsymbol{\delta}_\xi^\top \mathbf{D}_\xi \boldsymbol{\delta}_\xi = \sum_{j=1}^{k_\xi} \delta_{p_\xi + j}^2$, where p_ξ and $k_\xi + j$ are the polynomial degree and number of knots of the splines, respectively.[2] For notational convenience, we reserve subscript 1 for mean and 2 for volatility unless otherwise indicated.

An appealing feature of P-splines is that the number and location of knots are no longer crucial and the smoothness can easily be controlled by the smoothing parameter λ through a roughness penalty term (e.g., Ruppert (2002), and Ruppert, Wand, and Carroll (2003)). We use quadratic spline with knots placed at the equidistant quantiles of the index value. The above P-spline estimation also allows different smoothness by using separate smoothing parameters.

One could optimize the penalized log likelihood function (2) with respect to the model parameters $\boldsymbol{\theta} = (\boldsymbol{\gamma}^\top, {\boldsymbol{\delta}_a}^\top, {\boldsymbol{\delta}_b}^\top, \boldsymbol{\delta}_2^\top)^\top$ in one step. However, the number of parameters could be large and the estimation algorithm may not be efficient. Instead, we propose a two-step algorithm which reweighs the mean function using the estimates of volatility function, in a

[2] As noted in Carroll, Fan, Gijbels, and Wand (1997), this specification may also apply to quasi-likelihood models, where only the relationship between the model mean and the model variance is specified, by replacing the conditional log-likelihood function by a quasi-likelihood function.

fashion similar to weighted least squares. We first estimate the mean function and then use it to calculate the volatility. Parameters in the mean function are thereafter recalculated using weighted least squares where the weights are the inverse of the estimated volatility. The two steps can be iterated a few times. We find two or three iterations are usually adequate.[3] The two-step estimation procedure is described as follows.

Step 1: Mean Estimation.

The mean function can be estimated by minimizing

$$\sum_{t=1}^{n}\{R_{t+1} - \mathbf{B}_{a,t}(\mathbf{z}_t\boldsymbol{\gamma})\boldsymbol{\delta}_a - \mathbf{B}_{b,t}(\mathbf{z}_t\boldsymbol{\gamma})\boldsymbol{\delta}_b R_{m,t+1}\}^2 + \frac{n}{2}\boldsymbol{\lambda}_1\boldsymbol{\delta}_1^\mathsf{T}\mathbf{D}_1\boldsymbol{\delta}_1. \tag{3}$$

Given single index $u = \mathbf{z}\boldsymbol{\gamma}$ and the "design" matrix $\mathbf{X}$ defined above, the spline coefficients can be obtained by a linear shrinkage estimator $\widehat{\boldsymbol{\delta}}_1 = (\mathbf{X}^\mathsf{T}\mathbf{X} + n\boldsymbol{\lambda}_1\mathbf{D}_1)^{-1}\mathbf{X}^\mathsf{T}\mathbf{R}$, where $\boldsymbol{\lambda}_1 = blockdiag(\boldsymbol{\lambda}_a, \boldsymbol{\lambda}_b) = diag(\lambda_a, \ldots, \lambda_a, \lambda_b, \ldots, \lambda_b)$ is penalty parameter matrix, $\mathbf{D}_1 = blockdiag\{\mathbf{D}_a, \mathbf{D}_b\}$, and $\mathbf{R} = (R_2, \ldots, R_{n+1})^\mathsf{T}$ is the vector of value premiums. Estimated value premium is $\widehat{\mathbf{R}} = \mathbf{X}\widehat{\boldsymbol{\delta}}_1 = \mathbf{H}\mathbf{R}$, where $\mathbf{H} = \mathbf{X}(\mathbf{X}^\mathsf{T}\mathbf{X} + n\boldsymbol{\lambda}_1\mathbf{D}_1)^{-1}\mathbf{X}^\mathsf{T}$ is the smoothing matrix.

Step 2: Volatility Estimation and Reweighting.

The volatility coefficient parameter $\boldsymbol{\delta}_2$ can be estimated by minimizing the penalized negative log-likelihood

$$\sum_{t=1}^{n}\Big(e^2(\mathbf{z}_t\boldsymbol{\gamma})exp\{-2\mathbf{B}_2(\mathbf{z}_t\boldsymbol{\gamma})\boldsymbol{\delta}_2\} + 2\mathbf{B}_2(\mathbf{z}_t\boldsymbol{\gamma})\boldsymbol{\delta}_2\Big) + \frac{n}{2}\boldsymbol{\lambda}_2\boldsymbol{\delta}_2^\mathsf{T}\mathbf{D}_2\boldsymbol{\delta}_2, \tag{4}$$

where $e(\mathbf{z}_t\boldsymbol{\gamma}) = R_{t+1} - \mathbf{X}_t\widehat{\boldsymbol{\delta}}_1 \approx \sigma(\mathbf{z}_t\boldsymbol{\gamma})\epsilon_{t+1}$ is the residual from the estimation of mean function. The volatility function is approximated using

$$\widehat{\sigma}(\mathbf{z}_t\boldsymbol{\gamma}) = exp\{\mathbf{B}_2(\mathbf{z}_t\boldsymbol{\gamma})\widehat{\boldsymbol{\delta}}_2\}. \tag{5}$$

The new estimator $\widehat{\boldsymbol{\delta}}_1$ of the mean can be obtained through (penalized) weighted least squares where the weight is the inverse of the estimated squared volatility. More specifically,

[3] Similar treatments of iteratively estimating mean and volatility can be found in Carroll, Wu, and Ruppert (1988) and Yu, Yu, Wang, and Li (2009).

the new estimator $\widehat{\boldsymbol{\delta}}_1$ minimizes the penalized weighted least squares

$$\sum_{t=1}^{n} \frac{1}{\widehat{\sigma}^2(\mathbf{z}_t\boldsymbol{\gamma})}\Big(R_{t+1} - \mathbf{X}_t\boldsymbol{\delta}_1\Big)^2 + \frac{n}{2}\boldsymbol{\lambda}_1\boldsymbol{\delta}_1^{\mathsf{T}}\mathbf{D}_1\boldsymbol{\delta}_1.$$

It has an analytical solution $\widehat{\boldsymbol{\delta}}_1 = (\mathbf{X}^{\mathsf{T}}\mathbf{W}\mathbf{X}+n\boldsymbol{\lambda}_1\mathbf{D})^{-1}\mathbf{X}^{\mathsf{T}}\mathbf{W}\mathbf{R}$, where $\mathbf{W} = diag\{1/\widehat{\sigma}^2(\mathbf{z}_t\boldsymbol{\gamma})\}$ is a diagonal weight matrix with each diagonal element calculated by (5). In the case of homoscedasticity, the weighting procedure is unnecessary. That is, the second step and the reweighting can be omitted altogether.

2.2 Smoothing Parameter Selection

The selection of smoothing parameters is crucial in any nonparametric estimation. The most common criteria used to select smoothing parameters include Cross Validation (CV), Generalized Cross Validation (GCV), Mallow's C_p, Akaike's Information Criterion (AIC), and some other variations. We use GCV, which is particularly popular in the P-splines approach. In this paper, the smoothing parameter $\boldsymbol{\lambda}_1$ of mean minimizes the GCV criterion

$$GCV(\boldsymbol{\lambda}_1) = \frac{\sum_{t=1}^{n}(R_{t+1} - \mathbf{X}_t\boldsymbol{\delta}_1)^2}{n\{1 - n^{-1}tr(\mathbf{H})\}^2},$$

where $tr(\mathbf{H})$ is the trace of the smoothing matrix and is often called the "effective" degrees of freedom of the fit (see Hastie and Tibshirani 1990). Similarly, the smoothing parameter λ_2 of volatility minimizes the GCV criterion

$$GCV(\lambda_2) = \frac{Deviance(\lambda_2)}{n\{1 - n^{-1}tr(\mathbf{B}_2(\mathbf{B}_2^{\mathsf{T}}\mathbf{B}_2 + n\lambda_2\mathbf{D}_2)^{-1}\mathbf{B}_2^{\mathsf{T}})\}^2},$$

where the numerator is the deviance of the model for a given value of λ_2. The optimal smoothing parameter is often selected through grid search. For the two-dimensional $\boldsymbol{\lambda}_1$, to effectively search the optimal amount of smoothing, we recommend a straightforward grid searching scheme. First we constraint the two smoothing parameters to be the same and conduct a one-dimensional grid search to find the optimal smoothing parameter. Then we fix one smoothing parameter at the value obtained from the previous grid search and conduct grid search for the other parameter. We have studied two-dimensional separate smoothing parameters and found no significant improvement in the fit using separate smoothing parameters. Therefore, we present our results using a single smoothing parameter in this paper when estimating the mean functions.

2.3 A Fixed Point Algorithm

The objective function in (2) or (3) is nonlinear in the single-index parameters $\boldsymbol{\gamma}$, and we need to estimate them using numerical methods. We do not use standard nonlinear optimization routines, e.g, *lsqnonlin* and *fminunc* from Matlab and *nls* from S-PLUS, because for high dimensional $\boldsymbol{\gamma}$, the convergence of the estimates may be slow and the resulting estimates may be very sensitive to starting values. Instead, we develop a fixed point algorithm. Cui, Härdle, and Zhu (2011) propose the fixed point algorithm for local linear estimation for single-index models using an estimating function approach because it works well for high dimensional single-index models. In this paper, we adapt the algorithm to P-spline estimation for our single-index varying-coefficient models. In our implementation, the fixed point algorithm enables us to obtain efficient single-index parameter estimates $\boldsymbol{\gamma}$ in equation (3) by using a series of equations of its first derivatives through an iterative procedure. Compared with the aforementioned nonlinear optimization routines, our fixed point algorithm is less sensitive to starting values and works better for high dimensional $\boldsymbol{\gamma}$. The details of the fixed point algorithm as well as the estimation algorithm are relegated to Appendix A.

2.4 Variable Selection Procedure

Few studies have addressed variable selection for single-index varying-coefficient models. In a noticeable exception, Fan, Yao, and Cai (2003) use backward stepwise deletion in combination with the modified t-statistic and the Akaike Information Criterion (AIC) for variable selection. This heuristic procedure, however, does not guarantee a global solution of the best subset variables. To address this concern, we use some criterion to choose the best model from all possible combinations of candidate state variables. The exhaustive search approach can be computationally intensive and time-consuming. This drawback is not a serious concern for our study. While we consider a quite comprehensive set of candidate state variables, the number of all possible combinations is still manageable. Moreover, as we discuss above, P-splines and fixed point algorithm allow for expedient computation.

Akaike's Information Criterion (AIC, Akaike 1974) and Bayesian Information Criterion (BIC, Schwartz 1978) are the most popular likelihood-based information criteria for variable selection. Both criteria attempt to balance goodness of fit and simplicity of the model. We use BIC, which tends to select more parsimonious models. BIC in our model is defined as

$$\sum_{t=1}^{n}\Big(exp\{-2\mathbf{B}_2(\mathbf{z}_t\boldsymbol{\gamma})\boldsymbol{\delta}_2\}\{R_{t+1}-\mathbf{B}_{a,t}(\mathbf{z}_t\boldsymbol{\gamma})\boldsymbol{\delta}_a-\mathbf{B}_{b,t}(\mathbf{z}_t\boldsymbol{\gamma})\boldsymbol{\delta}_b R_{m,t+1}\}^2+2\mathbf{B}_2(\mathbf{z}_t\boldsymbol{\gamma})\boldsymbol{\delta}_2\Big)+ln(n)df,$$

where the first term is the negative two log-likelihood function of the model, and df is the model effective degrees of freedom.[4]

2.5 Tests of Alpha and Beta

To test whether conditional alpha or beta is constant, linear, or nonlinear, we first establish consistency and asymptotic normality, and then construct the Wald statistic for inferences. There are two types of asymptotics for P-splines: Fixed knots and increasing number of knots. We use fixed-knot asymptotics because they are most practical and relevant in empirical applications (Yu and Ruppert (2002), and Jarrow, Ruppert, and Yu (2004)). In addition, the bias due to fixed-knot approximation of P-splines of smooth function is negligible compared with the standard deviation of the function estimates and the bias due to penalty. The asymptotics introduced here are in spirit similar to Theorem 2 in Yu and Ruppert (2002) and to Theorem 2 in Wu, Lin, and Yu (2011).

[4] AIC assigns smaller weight to the simplicity of the model, where $ln(n)$ is replaced by 2. As a consequence, AIC tends to select more complex models. This is confirmed in our empirical study. For example, for the monthly data spanning the July 1963 to December 2012 period, in addition to the unemployment rate, the inflation rate, and the price-earnings ratio selected by BIC, the dividend-price ratio is also selected by AIC as a conditioning variable in the single index. Nevertheless, qualitatively similar results remain.

Theorem 1. *Under mild regularity conditions, if the smoothing parameter* $\lambda_n \sim o(n^{-1/2})$*, then a sequence of estimators* $\widehat{\boldsymbol{\theta}} = (\widehat{\boldsymbol{\gamma}}^{\mathsf{T}}, \widehat{\boldsymbol{\delta}}_a^{\mathsf{T}}, \widehat{\boldsymbol{\delta}}_b^{\mathsf{T}}, \widehat{\boldsymbol{\delta}}_2^{\mathsf{T}})^{\mathsf{T}}$ *is root-n consistent and converges to a normal distribution,*

$$\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{D} N(0, \mathbf{I}^{-1}(\boldsymbol{\theta})), \tag{6}$$

where $\mathbf{I}(\boldsymbol{\theta})$ *is the usual Fisher Information.*

The proof of the above theorem is similar to that in Yu, Yu, Wang, and Li (2009). The penalty parameter is assumed to vanish fast enough as n goes to infinity to ensure the result given in (6) involves no penalty parameter. For finite sample inference, asymptotic results with fixed penalty parameter would be preferred to avoid overestimating the variance of $\boldsymbol{\theta}$ as in (6). Using the delta method, the sandwich formula for the covariance matrix will be given by $\boldsymbol{\Gamma}_n^{-1}\big(\boldsymbol{\theta}(\boldsymbol{\lambda})\big)\boldsymbol{\Lambda}_n\big(\boldsymbol{\theta}(\boldsymbol{\lambda})\big)\boldsymbol{\Gamma}_n^{-T}\big(\boldsymbol{\theta}(\boldsymbol{\lambda})\big)$, where $\boldsymbol{\Gamma}_n\big(\boldsymbol{\theta}(\boldsymbol{\lambda})\big) = \sum_{t=1}^{n} (\partial/\partial\boldsymbol{\theta}^{\mathsf{T}})\Phi_{\mathbf{z}_t\boldsymbol{\gamma}}(\boldsymbol{\theta})$, $\boldsymbol{\Lambda}_n\big(\boldsymbol{\theta}\big) = \sum_{t=1}^{n} \Phi_{\mathbf{z}_t\boldsymbol{\gamma}}(\boldsymbol{\theta})\Phi_{\mathbf{z}_t\boldsymbol{\gamma}}^{\mathsf{T}}(\boldsymbol{\theta})$, $\Phi_{\mathbf{z}_t\boldsymbol{\gamma}}(\boldsymbol{\theta}) = -(\partial/\partial\boldsymbol{\theta}^{\mathsf{T}})l_n(\boldsymbol{\theta}; \mathbf{z}_t\boldsymbol{\gamma}) + \boldsymbol{\lambda}\mathbf{D}\boldsymbol{\theta}$, and $l_n(\boldsymbol{\theta}; \mathbf{z}_t\boldsymbol{\gamma})$ is the log-likelihood function (see Yu and Ruppert (2002), and Yu, Yu, Wang, and Li (2009)). The sandwich estimator of the covariance matrix is

$$\boldsymbol{\Omega}_n(\widehat{\boldsymbol{\theta}}(\boldsymbol{\lambda})) = \boldsymbol{\Gamma}_n^{-1}\big(\widehat{\boldsymbol{\theta}}(\boldsymbol{\lambda})\big)\boldsymbol{\Lambda}_n\big(\widehat{\boldsymbol{\theta}}(\boldsymbol{\lambda})\big)\boldsymbol{\Gamma}_n^{-T}\big(\widehat{\boldsymbol{\theta}}(\boldsymbol{\lambda})\big). \tag{7}$$

The asymptotic properties in (7) can be conveniently used for joint inferences concerning the spline coefficients. In general, if we are testing the null hypothesis $H_0 : \mathbf{L}\boldsymbol{\theta} = \mathbf{c}$, where $\mathbf{L}$ is a r by $dim(\boldsymbol{\theta})$ matrix of full row rank, we construct the Wald statistic

$$W = (\mathbf{L}\widehat{\boldsymbol{\theta}} - \mathbf{c})^{\mathsf{T}}\big(\mathbf{L}\boldsymbol{\Omega}_n(\widehat{\boldsymbol{\theta}}(\boldsymbol{\lambda}))\mathbf{L}\big)^{-1}(\mathbf{L}\widehat{\boldsymbol{\theta}} - \mathbf{c}). \tag{8}$$

Under H_0, the Wald statistic W has a limiting χ^2 distribution with r degrees of freedom. In particular, hypotheses on testing conditional alpha and beta can be formulated using spline coefficients. For example, testing $H_0 : \alpha \equiv 0$ is equivalent to testing all its spline coefficients are zeros, i.e., $H_0 : \boldsymbol{\delta}_a \equiv 0$. The null hypothesis H_0 : β *is linear in single index* is equivalent to test all nonlinear spline coefficients are simultaneously zero.

The covariance matrix $\mathbf{\Omega}_n(\widehat{\boldsymbol{\theta}})$ in the Wald statistic in (8) can be difficult to calculate using the sandwich formula for finite samples. Instead we use bootstrapping resamples to construct the covariance matrices in (8). Because of heteroskedasticity, we use a specialized wild bootstrap procedure. Wild bootstrap procedure, originally proposed by Wu (1986) and extended by Mammen (1993) and others, is particularly useful in the presence of heteroskedasticity and small sample sizes. Unlike the usual residual bootstrap where random residuals are drawn and added to the estimates to construct resamples, the wild bootstrap accounts for heteroskedasticity by creating weighted residuals where the weight is a random variable with zero mean and unit variance. The details of the wild bootstrap procedure are relegated to Appendix B.

Following Ang and Kristensen (2012), we define the long-run alpha as the average of conditional alphas:

$$\alpha_{LR} \equiv \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \alpha(\mathbf{z}_t \boldsymbol{\gamma}). \tag{9}$$

The long-run alpha measures the part of the value premium that is unaccounted by the conditional CAPM. We calculate the long-run alpha and its standard errors from the conditional alpha estimates and the wild bootstrap resamples. Note that the hypothesis of zero long-run alpha, i.e., $H_0 : \alpha_{LR} = 0$, is not the same as the hypothesis that conditional alpha is zero over any values of single index. Zero conditional alpha implies zero long-run alpha; the opposite, however, is not true.

3 Data

We obtain the monthly risk-free rate, excess stock market return, and value premium data from Ken French at Dartmouth College. Following earlier studies, e.g., Ferson and Harvey (1999) and Petkova and Zhang (2005), we use standard stock market return predictors as conditioning variables. The default premium (DEF) is the yield spread between Baa- and Aaa-rated corporate bonds. The term premium (TERM) is the yield spread between 10-

year Treasury bonds and 3-month Treasury bills. The stochastically de-trended risk-free rate (RREL) is the difference between the risk-free rate and its average in the previous 12 months. The dividend-price ratio (DP) is the dividend paid in the most recent 12 months divided by the end-of-month stock market price. The price-earnings ratio (PE) is the end-of-month stock market price divided by the earnings in the most recent 10 years. The realized stock market volatility (VOL) is the sum of squared daily excess stock market returns in a month. We obtain these data up to 2008 from Amit Goyal at Emory University and update them to 2012 using the data obtained from Ken French at Dartmouth College, Robert Shiller at Yale University, and the St. Louis Fed.

We also consider several major economic indicators as additional candidate conditioning variables. The unemployment (UE) and inflation (INF) are arguably the most important gauges of aggregate economic activity because by legal mandates, the Federal Reserve is required to maintain full employment and price stability.[5] We obtain the civilian unemployment rate and the consumer price index for all urban consumers (all items) from the St. Louis Fed. We use the year-over-year log changes in the consumer price index to measure the inflation rate and use the year-over-year log changes in the unemployment rate to measure the job market condition.[6] Chen, Roll, and Ross (1986) and others find that industrial production (IP) is a priced state variable in the cross-section of stock returns. We include its year-over-year log change obtained from the St. Louis Fed as a conditioning variable.

[5] Several recent studies, e.g., Chen and Zhang (2011), Petrosky-Nadeau, Zhang, and Kuehn (2013), and Belo, Lin, and Bazdresch (2014), have noted the important effect of labor market conditions on asset prices.

[6] In the Taylor (1993) rule, the federal funds rate adjusts with the inflation rate and the deviation of the unemployment rate from its natural rate. Because there is substantial uncertainty about the natural rate of unemployment, Staiger, Stock, and Watson (1997) and Orphanides and Williams (2002), however, argue that the Fed should target changes in the unemployment rate instead of the level. In addition, Berger, Dew-Becker, Schmidt, and Takahashi (2015) suggest that the Fed targets changes in the unemployment rate because layoffs lead to substantial social welfare losses. Consistent with these arguments, Berger, Dew-Becker, Schmidt, and Takahashi find that the federal funds rate correlates more closely with changes in the unemployment rate than with the level of the unemployment rate. That is, changes in the unemployment rate appear to be a better measure of business conditions than does the level of the unemployment rate. Similarly, when we include both changes and level of the unemployment rate as candidate conditioning variables, changes are always selected but the level is not.

Lewellen and Nagel (2006) use the realized beta, $BETA_R$, as a proxy for the conditional beta. Boguth, Carlson, Fisher, and Simutin (2011) illustrate via simulation that under some mild conditions, the lagged realized beta is a good instrumental variable for the conditional beta and can substantially reduce the underconditioning bias. As a robustness check, we include the lagged realized beta as a candidate conditioning variable. Following Lewellen and Nagel (2006), we estimate the realized beta of a month by regressing the daily value premium on a constant and the daily excess market return in that month using the ordinary least squares (OLS) regression.

Table 1 provides summary statistics of the state variables. Consistent with the conjecture that the value premium's conditional beta is countercyclical, we find that the realized beta, $BETA_R$, a proxy for the conditional beta, correlates positively with UE and correlates negatively with IP and INF. On the other hand, except for PE and TERM, the correlation of $BETA_R$ with stock market return predictors is rather weak. Similarly, correlations between macroeconomic variables and stock market return predictors are not particularly strong. These results highlight the importance of including macroeconomic variables as candidate conditioning variables because they may provide additional information about the conditional beta beyond financial variables.

4 Empirical Findings

Campbell and Vuolteenaho (2004), Fama and French (2006), Ang and Chen (2007), and others find that the value premium poses a challenge to the unconditional CAPM in the post-1963 sample but not in the pre-1963 sample. Following Ang and Kristensen (2012) and others, we focus mainly on the post-1963 sample in our empirical test of the conditional CAPM. Nevertheless, we show that the value premium's conditional beta changes countercyclically in both post- and pre-1963 samples.

4.1 Conditional Beta as a Function of Macroeconomic and Financial Variables

For the July 1963 to December 2012 period, we select the model with PE, INF, and UE as the conditioning variables via an exhaustive variable selection procedure. The scaled stock price such as PE plays a prominent role in the stock market return predictability literature because of its mechanical relation with the conditional equity premium (e.g., Campbell and Shiller (1988)), as highlighted in leading asset pricing models by Campbell and Cochrane (1999) and Bansal and Yaron (2004). Lettau and Ludvigson (2001) emphasize using the scaled stock price (i.e., the consumption-wealth ratio) in estimating the conditional CAPM for the value premium, and Zhang (2005) formalizes the argument using an equilibrium model in which the value premium's conditional market beta comoves closely with the scaled stock price. By identifying PE as a significant state variable, we confirm that the scaled stock price indeed has important effects on the dynamic of the value premium's conditional market beta. Moreover, because the unemployment and inflation are the two most closely watched gauges of aggregate economic activity by the Federal Reserve, our results confirm a key implication of the aforementioned investment-based asset pricing models that the value premium's conditional market beta changes with business conditions. To the best of our knowledge, this important empirical evidence for investment-based models is novel. [7]

Unlike the scaled stock price, the existing literature does not provide a fully fledged theoretical explanation for the relation between the other financial variables and the conditional equity premium. Financial economists, e.g., Fama and French (1989), usually justify the predictive power of these financial variables for stock market returns with the arguments that (1) the conditional equity premium is countercyclical and (2) these financial variables have strong cyclical variation across time. For example, Stock and Watson (2003) and others show

[7] In Zhang's (2005) model, the value premium's conditional market beta depends only on the scaled stock price because of the simplifying assumption that productivity shocks correlate perfectly with pricing kernel shocks. In general cases where the assumption is relaxed, the conditional market beta should comove with both the aggregate productivity and the scaled stock price, as we document in this paper.

that TERM, a commonly used stock market return predictor, has strong predictive power for aggregate output and inflation. Interestingly, while TERM is a standard conditioning variable in the extant empirical studies, it is not selected into the conditional CAPM model. This is because, as we show later, UE subsumes the information content of TERM about the value premium's conditional beta: The selected model has PE, INF, and TERM as the conditioning variables when we exclude UE as a candidate state variable. Our variable selection results suggest that consistent with investment-based models, macroeconomic variables can be more informative than can financial variables in estimating the conditional CAPM.

We report the main estimation results of the selected single-index model in Table 2. Panel A shows that the single index correlates positively with the procyclical variables INF and PE and correlates negatively with the countercyclical variable UE. Because there is a monotonically negative relation between the conditional beta and the single index (untabulated), our results indicate that the conditional beta of the value premium increases with UE and decreases with INF and PE and thus changes countercyclically across time. Figure 1 illustrates this point visually. We observe a sharp spike in the value premium's conditional beta for each of seven business recessions over the July 1963 to December 2012 period. Moreover, the conditional beta tends to decrease during business expansions. Our results reveal a strong countercyclical variation in the conditional beta of the value premium.[8]

Table 3 shows that the selected conditional variables forecast excess market returns. One-month-ahead excess market returns correlate positively with UE and negatively with INF and PE; both INF and PE are statistically significant at the 5% level. Similarly, for three-month-ahead excess market returns, their positive correlation with UE is significant at the 10% level, while their negative correlations with INF and PE are significant at the 1% level

[8]In Zhang's (2005) model, value stocks are riskier than growth stocks during business recessions. Consistent with this theoretical prediction, we find that betas of value stocks correlate negatively with INF and PE and correlate positively with UE. By contrast, betas of growth stocks correlate positively with INF and PE and correlate negatively UE. As a result, betas of value stocks increase sharply during business recessions, while betas of growth stocks decrease during economic downturns. For brevity, we do not report these results here but they are available on request.

(row 5). Moreover, the estimated single index (SI) that is a linear function of UE, INF, and PE has significant predictive power for one-month-ahead market returns at 5% level (row 2) and for three-month-ahead market returns at 1% level (row 7). The single index correlates closely with the first principle component of the nine candidate conditioning variables, PC1, with a correlation coefficient of 84% (untabulated); nevertheless, the predictive power for market returns is noticeably stronger for the former (rows 4 and 8). Our results indicate that the value premium's market beta correlates closely with conditional equity premium.

In panel B of Table 2, we test the hypotheses about the variation in the conditional beta across time. We reject the null hypothesis that the beta is zero at the 1% significance level. Moreover, we find strong evidence against the null hypothesis that the conditional beta is constant. We fail to reject the null hypothesis that the conditional beta is a linear function of the state variables INF, PE, and UE at the conventional significance level, and untabulated results show that the relationship between the conditional beta and the single index is essentially linear. This novel finding, which is in contrast with that reported by Wang (2003) who uses a fully nonparametric model, provides first empirical justification for the linear specification of the conditional beta adopted in the extant studies.

Panel C of Table 2 investigates cyclical variation in the conditional alpha. We overwhelmingly reject the null hypothesis that the conditional alpha is zero. Interestingly, our evidence reveals a significant nonlinear dependence of the conditional alpha on the single index: We reject the null hypothesis that the conditional alpha is a linear function of the single index at the 10% significance level. Specifically, untabulated results indicate that the conditional alpha is more sensitive to the single index during business recessions (when the single index is low) than during business expansions (when the single index is high). As a result, Figure 3 shows that the conditional alpha falls sharply and becomes negative during severe business downturns such as the Great Recession but is significantly positive and essentially flat otherwise. The novel evidence from Figure 3 lends empirical support to the rare disaster risk model analyzed by Bai, Hou, Kung, and Zhang (2015). These authors show using simulated data that the conditional alpha is positive and flat in normal times but becomes insignificant or negative when disaster risk materializes. Their model also implies that disaster risk leads to a spike in the conditional beta, as we document in Figure 1.

The value premium has a significantly positive alpha of 0.47% per month at the 1% significance level when we use the unconditional CAPM (untabulated). Panel D of Table 2 shows that the alpha decreases noticeably to 0.38% per month in the conditional CAPM. Untabulated results show that the 20% reduction in the alpha is also statistically significant. These findings are consistent with the notion that the conditional CAPM helps explain the value premium because its risk exposure is larger during business recessions when conditional equity premium is high than during business expansions when conditional equity premium is low. Nevertheless, we find that the alpha remains significantly positive at the 1% significance level even when we control for time-varying conditional beta. Again, this finding is consistent with the rare disaster model by Bai, Hou, Kung, and Zhang (2015). Like the peso problem, we observe a positive alpha in finite samples because it is a compensation for expected but unrealized rare disasters. That is, as we document in Figure 3, the conditional alpha is significantly positive during normal times when the economy is spared from catastrophic shocks. Moreover, consistent with their model implication that value stocks are more vulnerable to rare disasters than are growth stocks because of their higher investment adjustment costs, untabulated results show that the conditional alpha is positive for value stocks but is negligible for growth stocks.

4.2 Realized Beta as a Conditioning Variable

In the previous section, we find that the conditional CAPM fails to explain fully the value premium. One possible explanation is that our specification has an underconditioning bias because we omit some important state variables. Boguth, Carlson, Fisher, and Simutin

(2011) suggest that the lagged realized beta, $BETA_R$, is a good instrumental variable for the conditional beta. Nevertheless, because it is not necessarily an efficient measure of the conditional beta, the macroeconomic and financial variables may provide additional information about the conditional beta beyond $BETA_R$. To address these issues, we include $BETA_R$ as an additional candidate conditioning variable in the variable selection procedure. Over the July 1963 to December 2012 period, we select the model with $BETA_R$, INF, and PE as the conditioning variables.

Table 4 reports the main estimation results of the single-index varying-coefficient model with $BETA_R$, INF, and PE as the conditioning variables. Panel A shows that the single index correlates positively with $BETA_R$. Because the conditional beta increases monotonically with the single index (untabulated), our results indicate that the lagged realized beta correlates positively with the conditional beta. The single index also correlates negatively with INF and PE, again supporting the notion that the conditional beta changes countercyclically across time. Panel B shows that we reject at the 1% significance level the null hypothesis that the conditional beta is zero and the null hypothesis that the conditional beta is constant. However, we fail to reject the null hypothesis at the 10% significance level that the conditional beta is a linear function of the conditioning variables. Panel C shows that we reject the null hypothesis that the conditional alpha is constant at the 5% significance level, and reject the null hypothesis that the conditional alpha is linear in the single index at the 15% significance level. Overall, panel D shows that the long-run alpha remains significantly positive at the 1% level. Therefore, including the realized beta in the conditional CAPM does not change our main findings in any qualitative manner.

4.3 Ang and Kristensen's (2012) Time-Varying Beta

Ang and Kristensen (2012) use a nonparametric approach to model the value premium's conditional beta as a flexible function of the calendar time. In contrast with our findings, these authors document a rather weak relation between their beta estimates and business cycles. For comparison, in this section, we adopt their specification to estimate the time-varying beta and then use it as an additional candidate conditioning variable in our variable selection procedure. Ang and Kristensen (2012) use daily return data and the local polynomial smoothing algorithm in their estimation, and then construct the monthly beta using some bandwidth controlling criterion. In this paper, we estimate the monthly beta using monthly return data, and find qualitatively similar results using both the penalized spline smoothing algorithm and the local polynomial smoothing algorithm. For brevity, we report the results only for the former; and the results for the latter are available up request.

Figure 4 plots the estimated value premium's time-varying beta based on Ang and Kristensen's (2012) model over the post-1963 sample period. Interestingly, it appears to increase during business recessions and to decrease during business expansions. To investigate formally whether the estimated time-varying beta changes countercyclically across time, we regress it on the three most important conditioning variables (i.e., PE, INF, and UE) that we identify via the exhaustive variable selection procedure. Panel E of Table 5 shows that the estimated time-varying beta ($BETA_{AK}$) correlates negatively and significantly with PE and INF at the 1% level. It also correlates positively with UE albeit the correlation is statistically insignificant at the 10% level. Therefore, we again find countercyclical variation in the alternative measure of the value premium's conditional beta. Our results differ from those in Ang and Kristensen (2012) possibly because of bandwidths that these authors choose when constructing monthly conditional beta from nonparametric estimation based on the daily data. Moreover, we extend Ang and Kristensen's (2012) sample from December 2006 to December 2012 and thus include the 2008 subprime loan crisis, during which the value premium's conditional beta has a big spike (Figure 4).

Like the realized beta, Ang and Kristensen's (2012) time-varying beta may provide important information about the conditional beta. We address this issue in Table 5. Panel

A shows that we select the time-varying beta estimated using Ang and Kristensen's (2012) specification, $BETA_{AK}$, along with DEF, TERM, and INF as the conditioning variables. Because the conditional beta increases monotonically with the single index (untabulated), the results reported in panel A suggest that $BETA_{AK}$ correlates positively with the conditional beta. Moreover, we find that the conditional beta correlates positively with DEF and correlates negatively with TERM and INF, revealing its countercyclical variation. In panel B, we reject the null hypothesis that the conditional beta is zero or constant but fail to reject the null hypothesis that the conditional beta is a linear function of the selected state variables. In panel C, we reject the null hypothesis that the conditional alpha is zero but do not reject the null hypothesis that the conditional alpha is constant. Overall, panel D shows that the conditional CAPM fails to explain fully the value premium, despite countercyclical variation in the value premium's conditional beta. Therefore, controlling for Ang and Kristensen's (2012) conditional beta does not change our main findings qualitatively either.

4.4 Different Samples

While many authors, e.g., Ang and Kristensen (2012), focus on the post-1963 sample, as a robustness check, we estimate the conditional CAPM using a longer sample spanning the January 1949 to December 2012 period over which we have observations for all the variables. Results are qualitatively similar to those obtained from the post-1963 sample. We identify INF, PE, and UE as the most relevant state variables. The conditional beta is countercyclical because it correlates positively with INF and PE and correlates negatively with UE. We reject the null hypothesis that the conditional beta is zero or constant but fail to reject the null hypothesis that the conditional beta is a linear function of the selected state variables. In addition, the conditional alpha drops steeply toward zero during severe recessions but is significantly positive and essentially flat otherwise. Overall, the conditional CAPM has a long-run alpha of about 0.33% per month, a noticeable reduction from the 0.46% for the unconditional CAPM. For brevity, we do not report these results here but they are available on request.

As we mention above, several studies show that the value premium poses a challenge to the unconditional CAPM for the post-1963 sample but not for the pre-1963 sample. To address this issue, we estimate the conditional CAPM using the full sample (January 1927 to December 2012) and two subsamples (January 1927 to June 1963 and July 1963 to December 2012). Because the unemployment rate is available only from January 1949, for direct comparison of all three samples, we exclude UE from our set of candidate state variables. We summarize our main results in Table 6. For the full sample, we find that both macroeconomic variables (INF and IP) and financial variables (DEF, PE, and TERM) are significant conditioning variables for the value premium's conditional beta. The results are qualitatively similar when we include $BETA_R$ as an additional conditioning variable. More importantly, similar to the results obtained from the post-1963 sample, Figure 2 shows that the value premium's conditional beta exhibits strong countercyclical variation in the full sample as well. Nevertheless, the conditional CAPM fails to explain the value premium over the full sample. The variable selection results are similar for the pre-1963 sample. Both the macroeconomic variable (IP) and the financial variables (DEF, PE, and TERM) are significant conditioning variables for the value premium's conditional CAPM, although we confirm that the unconditional CAPM explains the value premium for the January 1927 to June 1963 period. For the post-1963 sample, when we exclude UE as a conditioning variable, TERM becomes a significant conditioning variable along with INF and PE. To summarize, the result that the value premium's conditional beta changes countercyclically across time is robust to the full sample and the pre-1963 sample.

5 Conclusion

We revisit whether the conditional CAPM helps explain the value premium using a general framework that has two novel features. First, it allows for a nonlinear dependence of the conditional alpha, beta, and volatility on the state variables and nest the linear specification as a special case. Second, we adopt an exhaustive variable selection method to choose the most relevant conditioning variables from a large set of candidate state variables. Furthermore, we consider key aggregate economic activity gauges as candidate conditioning variables, in addition to the stock market return and variance predictors and realized beta measures that are commonly used in previous studies.

We find strong countercyclical variation in the value premium's market beta. We also uncover a novel nonlinear dependence of the conditional alpha on business conditions: It decreases sharply and becomes even negative during severe economic downturns and is significantly positive and essentially flat otherwise. The conditional CAPM performs better than the unconditional CAPM; however, it does not fully explain the value premium either. Moreover, the value premium puzzle comes mainly from value stocks: The unconditional alpha is significantly positive for value stocks but is negligible for growth stocks. These findings are broadly consistent with the investment-based model proposed by Bai, Hou, Kung, and Zhang (2015), in which rare disasters generate a nonlinear dependence of alpha on business cycles and are responsible for the failure of CAPM in finite samples.

In this paper, we focus on the value premium to illustrate empirical applications of our novel statistical specification for three reasons. First, it is arguably the most prominent asset pricing anomaly. Second, previous empirical studies have investigated intensively whether it is explained by the conditional CAPM. Last, various theoretical explanations have been proposed for it, and they provide important guidance for our empirical analysis. In future studies, it will be interesting to apply the single-index model systematically to other important CAPM-related anomalies to shed light on their theoretical explanations.

Appendices

A Reparameterization and Fixed Point Algorithm

In this Appendix, we provide the technical details for the reparameterization of the single-index parameters and the estimation using the fixed point algorithm.

In statistical literature on single-index and related models, to ensure the identifiability of single-index parameters, we commonly impose the constraints of unit norm and positive first component, i.e., $\|\boldsymbol{\gamma}\| = 1$ and $\gamma_1 > 0$ (see e.g. Carroll, Fan, Gijbels, and Wand (1997), Yu and Ruppert (2002)). Let $\gamma_1 = \sqrt{1 - (\gamma_2^2 + \ldots + \gamma_d^2)}$, where $\sum_{j=2}^{d} \gamma_j^2 < 1$, thus the unconstrained single-index parameters are $\boldsymbol{\varphi} = (\gamma_2, \gamma_3, \ldots, \gamma_d)^\intercal$.[9] Denote $\boldsymbol{\theta} = (\boldsymbol{\gamma}^\intercal, \boldsymbol{\delta}_a^\intercal, \boldsymbol{\delta}_b^\intercal, \boldsymbol{\delta}_2^\intercal)^\intercal$ and $\tilde{\boldsymbol{\theta}} = (\boldsymbol{\varphi}^\intercal, \boldsymbol{\delta}_a^\intercal, \boldsymbol{\delta}_b^\intercal, \boldsymbol{\delta}_2^\intercal)^\intercal$ as the parameter vector before and after parameterization. The Jacobian matrix of this transformation is

$$\mathbf{J}(\tilde{\boldsymbol{\theta}}) = \begin{bmatrix} -(1 - \|\boldsymbol{\varphi}\|^2)^{-1/2}\boldsymbol{\varphi}^T & \mathbf{0} \\ \mathbf{I}_{d-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{dim(\boldsymbol{\delta}_a)+dim(\boldsymbol{\delta}_b)+dim(\boldsymbol{\delta}_2)} \end{bmatrix}. \tag{A.1}$$

All inference results in Section 2.5 are for the reparameterized $\tilde{\boldsymbol{\theta}}$. We can formulate test statistics concerning the original parameters $\boldsymbol{\theta}$ by the inverse transformation using the Jacobian matrix. For example, the Fisher Information in Theorem 1 will become $\mathbf{J}(\tilde{\boldsymbol{\theta}})\mathbf{I}^{-1}(\tilde{\boldsymbol{\theta}})\mathbf{J}(\tilde{\boldsymbol{\theta}})^\intercal$, and sandwich estimator of the covariance matrix will become $\mathbf{J}(\tilde{\boldsymbol{\theta}})\boldsymbol{\Omega}_n(\tilde{\boldsymbol{\theta}}(\boldsymbol{\lambda}))\mathbf{J}(\tilde{\boldsymbol{\theta}})^\intercal$.

We adapt the fixed point algorithm (Cui, Härdle, and Zhu (2011)) to penalized spline estimation in our single-index varying-coefficient models and find computationally promising results. Specifically, given $\boldsymbol{\delta}_1$, to obtain the single-index parameter estimates $\boldsymbol{\gamma}$ in equation (3), we let the first derivatives of the squared loss with respect to γ_s be zeros. Note that the roughness penalty terms vanish as the penalty is only on the spline coefficients. Define

$$M_s = \sum_{t=1}^{n} (R_{t+1} - \mathbf{B}_{a,t}\boldsymbol{\delta}_a - \mathbf{B}_{b,t}\boldsymbol{\delta}_b R_{m,t+1})(\mathbf{B}'_{a,t}\boldsymbol{\delta}_a + \mathbf{B}'_{b,t}\boldsymbol{\delta}_b R_{m,t+1}) z_{s,t}, \;\; s = 1, \ldots, d, \tag{A.2}$$

[9] An alternative way of reparameterization is letting $\gamma_1 = 1$, and the new parameter $\boldsymbol{\varphi} = (1, \gamma_2, \gamma_3, \ldots, \gamma_d)/\sqrt{1 + \sum_{j=2}^{d} \gamma_j^2}$. However, this reparametrization may not be suitable for the fixed point algorithm proposed here.

where $\mathbf{B}'_{\xi,t}(\xi = a, b)$ is the first derivative of the spline basis evaluated at $\mathbf{z}_t$, and define $\mathbf{M} = (M_1, \ldots, M_d)$. Setting $\mathbf{M} = 0$ and by the chain rule, we obtain $d-1$ equations

$$\begin{cases} \gamma_2 M_1 & = & \gamma_1 M_2 \\ & \vdots & \\ \gamma_d M_1 & = & \gamma_1 M_d, \end{cases}$$

which have the solution

$$\begin{cases} \gamma_1 & = & \frac{||M_1||}{||\mathbf{M}||} \\ \gamma_s^2 & = & \frac{||M_s||^2}{||\mathbf{M}||^2}, 2 \leq s \leq d \\ sign(\gamma_s M_1) & = & sign(M_s), 2 \leq s \leq d. \end{cases}$$

The above system equals $\boldsymbol{\gamma} \frac{M_1}{||\mathbf{M}||} = \frac{|M_1|}{||\mathbf{M}||} \cdot \frac{\mathbf{M}}{||\mathbf{M}||}$, which automatically handles the constraints $||\boldsymbol{\gamma}|| = 1$ and $\gamma_1 > 0$. For fast convergence and robustness of the fixed point algorithm, we add some constant C to $||\mathbf{M}||$ to avoid dividing by a small value. To achieve this, add $C\boldsymbol{\gamma}$ to both sides and after transformation we obtain

$$\boldsymbol{\gamma} = \frac{C}{M_1/||\mathbf{M}|| + C} \boldsymbol{\gamma} + \frac{M_1/||\mathbf{M}||^2}{M_1/||\mathbf{M}|| + C} \mathbf{M}. \tag{A.3}$$

We choose the constant C properly to avoid dividing by zero and refer further discussions to Cui, Härdle, and Zhu (2011).

We describe the two-phase iterative algorithm for iteratively estimating the mean and volatility functions as the following.

Step 0: Initialization. Initialize single-index parameter $\boldsymbol{\gamma}$. For example, we can use the usual ordinary linear squares (OLS) estimates for linear regression $R_{t+1} = (\mathbf{z}_t + \mathbf{z}_t R_{m,t+1})\boldsymbol{\gamma} + \epsilon_{t+1}$. Alternatively, we can use random initial points from unit sphere. Normalize $\boldsymbol{\gamma}$ such that $\boldsymbol{\gamma} = sign(\gamma_1)\boldsymbol{\gamma}/||\boldsymbol{\gamma}||$.

Step 1: Mean Estimation.

1a) Given $\widehat{\boldsymbol{\gamma}}$, compute spline coefficients of $\alpha(\cdot)$ and $\beta(\cdot)$, or $\widehat{\boldsymbol{\delta}}_1$ and the estimated value premium $\widehat{R}_{t+1}$. We calculate spline coefficients using $\widehat{\boldsymbol{\delta}}_1 = (\mathbf{X}^\intercal\mathbf{X} + n\boldsymbol{\lambda}_1\mathbf{D}_1)^{-1}\mathbf{X}^\intercal\mathbf{R}$ and estimate value premium using $\widehat{\mathbf{R}} = \mathbf{X}\widehat{\boldsymbol{\delta}}_1$, where $\mathbf{X}$ is the design matrix.

1b) Fixed Point Algorithm. Given $\widehat{\boldsymbol{\delta}}_1$ and $\widehat{\mathbf{R}}$, calculate the first derivatives by (A.2).

Use the fixed point algorithm to update single-index parameters $\boldsymbol{\gamma}$ by (A.3) and normalize $\boldsymbol{\gamma}$ such that $\boldsymbol{\gamma} = sign(\gamma_1)\boldsymbol{\gamma}/||\boldsymbol{\gamma}||$.

Repeat 1a) & 1b) until $max_{1\le s\le d}|\boldsymbol{\gamma}_{s,new} - \boldsymbol{\gamma}_{s,old}| \le tol$, where tol is the prescribed tolerance level.

Step 2: Volatility Estimation and Reweighting.

Given $\widehat{\boldsymbol{\gamma}}$ and $\widehat{\boldsymbol{\delta}}_1$, estimate $\boldsymbol{\delta}_2$ by minimizing the penalized negative log-likelihood function in (4). Use the updated squared volatility estimates $\widehat{\sigma}^2(\mathbf{z}_t\widehat{\boldsymbol{\gamma}})$ to re-estimate the mean spline coefficients by the (penalized) weighted least square estimator $\widehat{\boldsymbol{\delta}}_1 = (\mathbf{X}^\intercal\mathbf{W}\mathbf{X} + n\boldsymbol{\lambda}_1\mathbf{D})^{-1}\mathbf{X}^\intercal\mathbf{W}\mathbf{R}$, where $\mathbf{W} = diag\{1/\widehat{\sigma}^2(\mathbf{z}_t\widehat{\boldsymbol{\gamma}})\}$ is the diagonal weight matrix.

In our implementation, we find that two or three iterations of Step 1 and Step 2 are sufficient as Carroll, Wu, and Ruppert (1988).

B Wild Bootstrap Procedure

In this Appendix, we describe in detail the wild bootstrap procedure for obtaining the covariance matrix.

(1) Fit the single-index varying-coefficient model, obtain $\widehat{R}_{t+1} = \mathbf{B}_a(\mathbf{z}_t\widehat{\boldsymbol{\gamma}})\widehat{\boldsymbol{\delta}}_a + \mathbf{B}_b(\mathbf{z}_t\widehat{\boldsymbol{\gamma}})\widehat{\boldsymbol{\delta}}_b R_{m,t+1}$, and residual $\widehat{\epsilon}_{t+1} = R_{t+1} - \widehat{R}_{t+1}$ $(t = 1, \ldots, n)$.

(2) Center the residual $\epsilon^*_{t+1} = \widehat{\epsilon}_{t+1} - \bar{\widehat{\epsilon}}_{t+1}$ $(t = 1, \ldots, n)$, where $\bar{\widehat{\epsilon}}_{t+1}$ is the mean of residuals.

(3) Create the bootstrap sample using $\widehat{R}_{t+1} + v_{t+1}\epsilon^*_{t+1}$ $(t = 1, \ldots, n)$, where v_{t+1} is a random variable of standard normal distribution.[10]

Repeat (3) to create N resamples. Sample covariance matrix and standard errors of long-run alpha can be calculated based on the N estimates.

[10] We have also implemented the wild bootstrap weighting scheme of Mammen (1993) where the random weight takes the form:

$$v_{t+1} = \begin{cases} -(\sqrt{5}-1)/2 & \text{with probability } (\sqrt{5}+1)/(2\sqrt{5}) \\ (\sqrt{5}+1)/2 & \text{with probability } (\sqrt{5}-1)/(2\sqrt{5}). \end{cases}$$

The testing results agree with that of using standard normal weight as in Wu (1986).

References

Akaike, H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(1974), 716-723.

Ang, A., J. Chen. CAPM over the long run: 1926-2001. *Journal of Empirical Finance*, 14(2007), 1-40.

Ang, A., D. Kristensen. Testing conditional factor models. *Journal of Financial Economics*, 106(2012), 132-156.

Bai, H., K. Hou, H. Kung, and L. Zhang. The CAPM Strikes Back? An Investment Model with Disasters. *Unpublished Working Paper*, Ohio State University (2015).

Bansal, R., A. Yaron. Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance*, 59(2004), 1481-1509.

Belo, F., X. Lin, S. Bazdresch. Labor hiring, investment and stock return predictability in the cross section. *Journal of Political Economy*, 122(2014), 129-177.

Berger, D., I. Dew-Becker, L. Schmidt, and Y. Takahashi. Layoff risk, the welfare cost of business cycles, and monetary policy. *Unpublished Working Paper*, Northwestern University (2015).

Boguth, O., M. Carlson, A. Fisher, M. Simutin. Conditional risk and performance evaluation: Volatility timing, overconditioning, and new estimates of momentum alphas, *Journal of Financial Economics*, 102(2011), 363-389.

Campbell, J. Y., J. H. Cochrane. By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy*, 107(1999), 205-251.

Campbell, J. Y., R. J. Shiller. The dividendprice ratio and expectations of future dividends and discount factors. *Review of Financial Studies*, 1(1988), 195-228.

Campbell, J. Y., T. Vuolteenaho. Inflation illusion and stock prices. *American Economic Review*, 94(2004), 19-23.

Carroll, R. J., J. Fan, I. Gijbels, M.P. Wand. Generalized partially linear single-index models. *Journal of the American Statistical Association*, 92(1997), 477-489.

Carroll, R. J., C. J. F. Wu, D. Ruppert. The effect of estimating weights in weighted least squares. *Journal of the American Statistical Association*, 83(1988), 1045-1054.

Chen, N.-F., R. Roll, S. A. Ross. Economic forces and the stock market. *Journal of Business*, 59(1986), 383-403.

Chen, L., L. Zhang. Do time-varying risk premiums explain labor market performance? *Journal of Financial Economics*, 99(2011), 385-399.

Cooper M., S. Gubellini. The critical role of conditioning information in determining if value is really riskier than growth. *Journal of Empirical Finance*, 18(2011), 289-305.

Cui, X., W. Härdle, L. Zhu. The EFM approach for single-index models. *Annals of Statistics*, 39(2011), 1658-1688.

Fama, E. F., K. R. French. Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics*, 25(1989), 23-49.

Fama, E. F., K. R. French. The value premium and the CAPM. *Journal of Finance*, 61(2006), 2163-2185.

Fan, J., Q. Yao, Z. Cai. 2003. Adaptive varying-coefficient linear models. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*. 65(1) 57-80.

Ferson, W. E., C. R. Harvey. Conditioning variables and the cross section of stock returns. *Journal of Finance*, 54(1999), 1325-1360.

Ghysels, E. On stable factor structures in the pricing of risk: Do time-varying betas help or hurt? *Journal of Finance*, 53(1998), 549-573.

Hastie, T., Tibshirani, R. Generalized Additive Models, Chapman and Hall (1990).

Harvey, C. The specification of conditional expectations. *Journal of Empirical Finance*, 8(2001), 573-638.

Jarrow, R., D. Ruppert, Y. Yu. Estimating the term structure of corporate debt with a semiparametric penalized spline model, *Journal of the American Statistical Association*, 99(2004), 57-66.

Lettau M., S. Ludvigson. Resurrecting the (c)capm: A cross-sectional test when risk premia are time-varying. *Journal of Political Economy*, 109(2001), 1238-1287.

Lewellen J., S. Nagel. The conditional CAPM does not explain asset-pricing anomalies. *Journal of Financial Economics*, 82(2006), 289-314.

Mammen, E. Bootstrap and wild bootstrap for high dimensional linear models. *Annals of Statistics*, 21(1993), 255-285.

Orphanides, A., and J. Williams. Robust monetary policy rules with unknown natural rates, *Brookings Papers on Economic Activity*, 2(2002), 63-145.

Paye, B. 'Deja vol': Predictive regressions for aggregate stock market volatility using macroeconomic variables. *Journal Of Financial Economics*, 106(2012), 527-546.

Petrosky-Nadeau, N., L. Zhang, L. Kuehn. Endogenous economic disasters and asset prices. Unpublished Working Paper, Ohio State University (2013).

Petkova, R., L. Zhang. Is value riskier than growth? *Journal of Financial Economics*, 78(2005), 187-202.

Ruppert, D. Selecting the number of knots for penalized splines. *Journal of Computational and Graphical Statistics*, 11(2002), 735-757.

Ruppert, D., M. P. Wand, R. J. Carroll. Semiparametric regression. New York: Cambridge University Press (2003).

Schwarz, G. Estimating the dimension of a model. *Annals of Statistics*, 6(1978), 461-464.

Staiger, D., J. Stock, and M. Watson. How precise are estimates of the natural rate of unemployment? In Reducing Inflation: Motivation and Strategy, edited by C. Romer and D. Romer, University of Chicago Press (1997).

Stock, J., M. Watson. Forecasting output and ination: the role of asset prices. *Journal of Economic Literature*, 41(2003), 788-829.

Taylor, J., Discretion versus policy rules in practice, *Carnegie-Rochester Conference Series on Public Policy*, 39 (1993), 195-214.

Wang, K. Q. Asset pricing with conditioning information: A new test. *Journal of Finance*, 58(2003), 161-196.

Welch, I., A. Goyal. A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, 21(2008), 1455-1508.

Wu, C.F.J. Jackknife, bootstrap and other resampling methods in regression analysis (with discussions). *Annals of Statistics*, 14(1986), 1261-1350.

Wu, T. Z., H. Lin, Y. Yu. Single index coefficient models for nonlinear time series. *Journal of Nonparametric Statistics*, 23(2011), 37-58.

Xia Y., W. K. Li. On single-index coefficient regression models. *Journal of the American Statistical Association*, 94(1999), 1275-1285.

Yu, Y., D. Ruppert. Penalized spline estimation for partially linear single-index models. *Journal of the American Statistical Association*, 97(2002), 1042-1054.

Yu, Y., K. Yu, H. Wang, M. Li. Semiparametric estimation for a class of time-inhomogeneous diffusion processes. *Statistica Sinica*, 19(2009), 843-867.

Zhang, L. The value premium. *Journal of Finance*, 60(2005), 67-103.

Table 1: Summary Statistics of Conditioning Variables

Table 1 reports descriptive statistics of conditioning variables (Panel A) and their correlation matrix (Panel B). $BETA_R$ is the value premium's realized beta. DEF is the default premium. TERM is the term premium. RREL is the stochastically detrended risk-free rate. DP is the dividend yield. IP is the industrial production. INF is inflation. VOL is the realized stock market volatility. PE is the price-earnings ratio. UE is unemployment. For comparison, we also include the NBER business recession indicator, CYCLE. All variables are in percent points except for $BETA_R$, PE, and CYCLE. PE has been divided by 100. The monthly sample spans the July 1963 to December 2012 period.

Panel A: Descriptive Statistics of Conditioning Variables.

Variable	$BETA_R$	DEF	TERM	RREL	DP	IP	INF	VOL	PE	UE	CYCLE
Mean	-0.19	1.05	1.80	0.00	2.98	2.66	4.08	0.23	0.20	0.72	0.15
Median	-0.21	0.93	1.82	0.00	2.92	3.21	3.36	0.13	0.20	-4.45	0.00
Standard Deviation	0.26	0.47	1.54	0.09	1.11	4.78	2.72	0.46	0.08	16.63	0.36
Kurtosis	1.27	3.63	-0.38	2.52	-0.61	2.18	1.93	144.29	0.61	1.31	1.80
Skewness	0.50	1.60	-0.29	-0.11	0.37	-1.19	1.39	10.59	0.77	1.24	1.95
Minimum	-1.10	0.32	-3.65	-0.35	1.07	-16.28	-2.13	0.01	0.07	-33.85	0.00
Maximum	0.88	3.38	4.55	0.38	6.01	11.82	13.80	7.50	0.44	58.78	1.00

Panel B: Correlation Matrix

	$BETA_R$	DEF	TERM	RREL	DP	IP	INF	VOL	PE	UE	CYCLE
$BETA_R$	1.00										
DEF	0.07	1.00									
TERM	0.23	0.23	1.00								
RREL	-0.05	-0.37	-0.53	1.00							
DP	-0.08	0.43	-0.22	0.05	1.00						
IP	-0.11	-0.63	-0.19	0.48	-0.15	1.00					
INF	-0.24	0.28	-0.42	0.19	0.71	-0.17	1.00				
VOL	0.02	0.25	0.07	-0.09	0.02	-0.19	0.03	1.00			
PE	-0.10	-0.51	-0.02	0.00	-0.90	0.18	-0.58	0.00	1.00		
UE	0.14	0.57	0.17	-0.51	0.16	-0.88	0.19	0.19	-0.19	1.00	
CYCLE	-0.05	0.43	-0.08	-0.33	0.30	-0.52	0.37	0.24	-0.23	0.54	1.00

Table 2: Estimation of the Single-Index Varying-Coefficient Model

Table 2 reports the estimation results of the single-index varying-coefficient model in which we identify the conditioning variables via the exhaustive variable selection from standard macroeconomic and financial variables. Panel A reports the single-index parameter estimates of the selected conditioning variables. INF is inflation. PE is the price-earnings ratio. UE is unemployment. Panels B to D report statistical tests on conditional beta, conditional alpha, and long-run alpha, respectively. Tests on conditional betas and alphas are based on the Wald statistics in equation (8). Tests on long-run alpha are based on equation (9). All covariance matrices and standard errors are calculated using the wild bootstrap. The monthly sample spans the July 1963 to December 2012 period.

Panel A: Single-Index Parameter Estimates		
		Estimate(s.e.)
$\gamma_1(INF)$		0.9574(0.0994)
$\gamma_2(PE)$		0.2524(0.1305)
$\gamma_3(UE)$		-0.1402(0.0872)
Panel B: Tests on Beta		
Hypothesis	Wald Statistic	p-Value
$H_o : \beta = 0$	89.12	0.0000
$H_o : \beta = Constant$	70.10	0.0000
$H_o : \beta = Linear$	7.31	0.2926
Panel C: Tests on Alpha		
Hypothesis	Wald Statistic	p-Value
$H_o : \alpha = zero$	27.22	0.0006
$H_o : \alpha = Constant$	12.76	0.0781
$H_o : \alpha = Linear$	11.09	0.0857
Panel D: Test on Long-Run Alpha		
Hypothesis		Estimate(s.e.)
$H_o : \alpha_{LR} = 0$		0.0038(0.0011)

Table 3: Forecasting Excess Market Returns

Table 3 reports the OLS regression results of forecasting excess stock market returns. INF is inflation. PE is the price-earnings ratio. UE is unemployment. SI is the single index estimated using the post-1963 sample. PC1 is the first principle component of the nine candidate macroeconomic and financial variables. The monthly sample spans the August 1963 to December 2012 period. Heteroskedasticity-corrected t-value is reported in parentheses. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

	UE	INF	PE	SI	PC1	Adjusted R^2
Panel A: One-Month ahead Excess Market Returns						
(1)	0.015	-0.252**	-0.066**			0.014
	(1.235)	(-2.463)	(-2.175)			
(2)				-0.159**		0.011
				(-2.388)		
(3)					-0.002	0.003
					(-1.451)	
(4)				-0.156*	-0.001	0.009
				(-1.922)	(-0.071)	
Panel B: Three-Month ahead Excess Market Returns						
(5)	0.044*	-0.675***	-0.190***			0.039
	(1.868)	(-3.991)	(-3.804)			
(6)				-0.490***		0.034
				(-4.530)		
(7)					-0.005***	0.014
					(-2.668)	
(8)				-0.460***	-0.009	0.033
				(-3.492)	(-0.358)	

Table 4: Controlling for Realized Beta

Table 4 reports the estimation results of the single-index varying-coefficient model with the value premium's realized beta, $BETA_R$, as a candidate conditioning variable, in addition to standard macroeconomic and financial variables. Panel A reports the single-index parameter estimates of the selected conditioning variables via the exhaustive variable selection. INF is inflation. PE is the price-earnings ratio. Panels B to D report statistical tests on conditional beta, conditional alpha, and long-run alpha, respectively. Tests on conditional betas and alphas are based on the Wald statistics in equation (8). Tests on long-run alpha are based on equation (9). All covariance matrices and standard errors are calculated using the wild bootstrap. The monthly sample spans the July 1963 to December 2012 period.

Panel A: Single-Index Parameter Estimates		
		Estimate(s.e.)
$\gamma_1(BETA\_R)$		0.3352(0.1583)
$\gamma_2(INF)$		-0.8533(0.2866)
$\gamma_3(PE)$		-0.3994(0.1424)
Panel B: Tests on Beta		
Hypothesis	Wald Statistic	p-Value
$H_o : \beta = 0$	140.30	0.0000
$H_o : \beta = Constant$	85.69	0.0000
$H_o : \beta = Linear$	10.64	0.4165
Panel C: Tests on Alpha		
Hypothesis	Wald Statistic	p-Value
$H_o : \alpha = zero$	28.83	0.0003
$H_o : \alpha = Constant$	16.18	0.0235
$H_o : \alpha = Linear$	9.58	0.1436
Panel D: Tests on Long-Run Alpha		
Hypothesis		Estimate(s.e.)
$H_o : \alpha_{LR} = 0$		0.0041(0.0011)

Table 5: Controlling for Ang and Kristensen's (2012) Conditional Beta

Table 5 reports the estimation results of the single-index varying-coefficient model with Ang and Kristensen's (2012) conditional beta of the value premium, $BETA_{AK}$, as a candidate conditioning variable, in addition to standard macroeconomic and financial variables. Panel A reports the single-index parameter estimates of the selected conditioning variables via the exhaustive variable selection. DEF is the default premium. TERM is the term premium. INF is inflation. Panels B to D report statistical tests on conditional beta, conditional alpha, and long-run alpha, respectively. Tests on conditional betas and alphas are based on the Wald statistics in equation (8). Tests on long-run alpha are based on equation (9). All covariance matrices and standard errors are calculated using the wild bootstrap. Panel E reports the OLS regression results of regressing the estimated time-varying beta based on Ang and Kristensen (2012) onto the three most important conditioning variables (i.e., PE, INF, and UE). The monthly sample spans the July 1963 to December 2012 period.

Panel A: Single-Index Parameter Estimates

	Estimate(s.e.)
$\gamma_1(BETA_{AK})$	0.1506(0.0896)
$\gamma_2(DEF)$	0.8490(0.3425)
$\gamma_3(TERM)$	-0.4281(0.1682)
$\gamma_4(INF)$	-0.2707(0.1445)

Panel B: Tests on Beta

Hypothesis	Wald Statistic	p-Value
$H_o : \beta = 0$	167.28	0.0000
$H_o : \beta = Constant$	93.60	0.0000
$H_o : \beta = Linear$	0.47	0.9981

Panel C: Tests on Alpha

Hypothesis	Wald Statistic	p-Value
$H_o : \alpha = zero$	29.82	0.0002
$H_o : \alpha = Constant$	6.97	0.4325
$H_o : \alpha = Linear$	0.35	0.9992

Panel D: Tests on Long-Run Alpha

Hypothesis	Estimate(s.e.)
$H_o : \alpha_{LR} = 0$	0.0042(0.0010)

Panel E: OLS Regression of $BETA_{AK}$ onto INF, PE and UE

	Estimate(s.e.)	t Value	p-Value
Intercept	0.17(0.05)	3.57	0.0004
INF	-3.04(0.52)	-5.88	0.0000
PE	-1.47(0.15)	-9.50	0.0000
UE	0.05(0.07)	0.65	0.5149

Table 6: Variable Selection Results and Long-Run Alpha Estimates in Different Samples

Table 6 summarizes the variable selection results and long-run alpha estimates in different samples. Full sample spans the January 1927 to December 2012 period. Pre-1963 sample spans the January 1927 to June 1963 period. Post-1963 sample spans the July 1963 to December 2012 period. α_{LR} is the long-run alpha defined in (9). $BETA_R$ is the value premium's realized beta. DEF is the default premium. INF is inflation. IP is industrial production. PE is the price-earnings ratio. TERM is the term premium. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively.

Sample	Variable Pool	Best Variable Set	α_{LR}
Full Sample	No $BETA_R$	DEF, INF, IP, PE, TERM	0.0031***
	Include $BETA_R$	DEF, INF, IP, PE, $BETA_R$	0.0025***
Pre-1963 Sample	No $BETA_R$	DEF, IP, PE, TERM	0.0014
	Include $BETA_R$	DEF, IP, PE, TERM	0.0014
Post-1963 Sample	No $BETA_R$	INF, PE, TERM	0.0043***
	Include $BETA_R$	INF, PE, $BETA_R$	0.0040***

Figure 1: Conditional Beta over the July 1963 to December 2012 Period

Figure 1 plots monthly estimates of the value premium's conditional beta in solid line along with its 95% confidence bands in dashed lines. The conditioning variables are inflation (INF), the price-earning ratio (PE), and unemployment (UE) that we identify via the exhaustive variable selection. Shaded areas indicate the NBER recession periods.

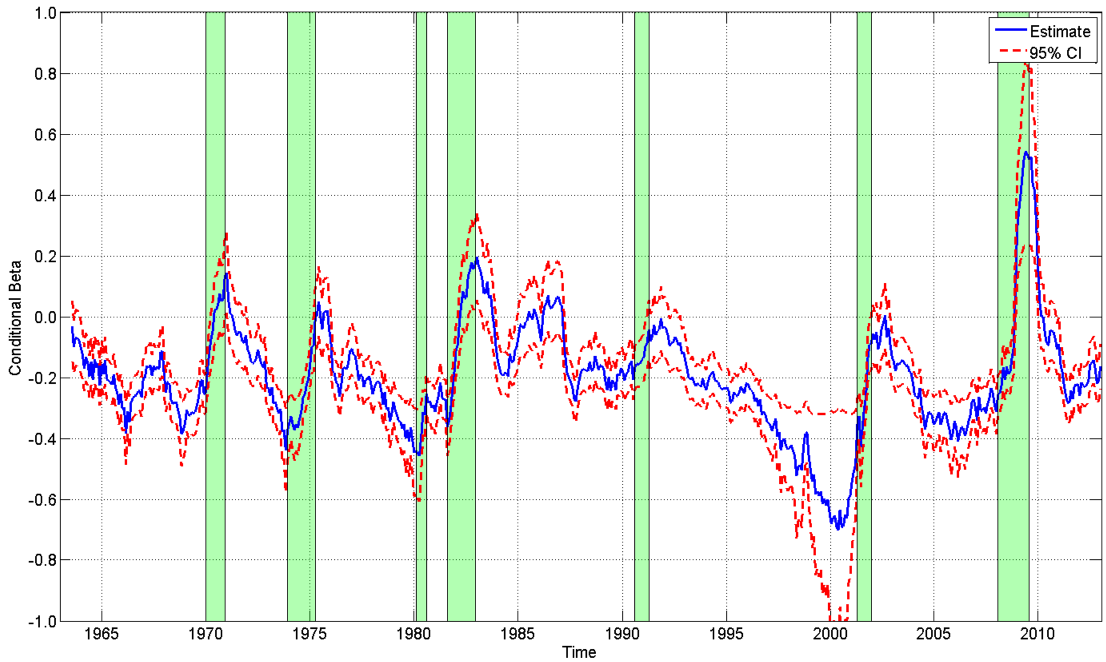

Figure 2: Conditional Beta over the January 1927 to December 2012 Period

Figure 2 plots monthly estimates of the value premium's conditional beta in solid line along with its 95% confidence bands in dashed lines. The conditioning variables are the default premium (DEF), industrial production (IP), inflation (INF), the price-earning ratio (PE), and the term premium (TERM) that we identify via the exhaustive variable selection. Shaded areas indicate the NBER recession periods.

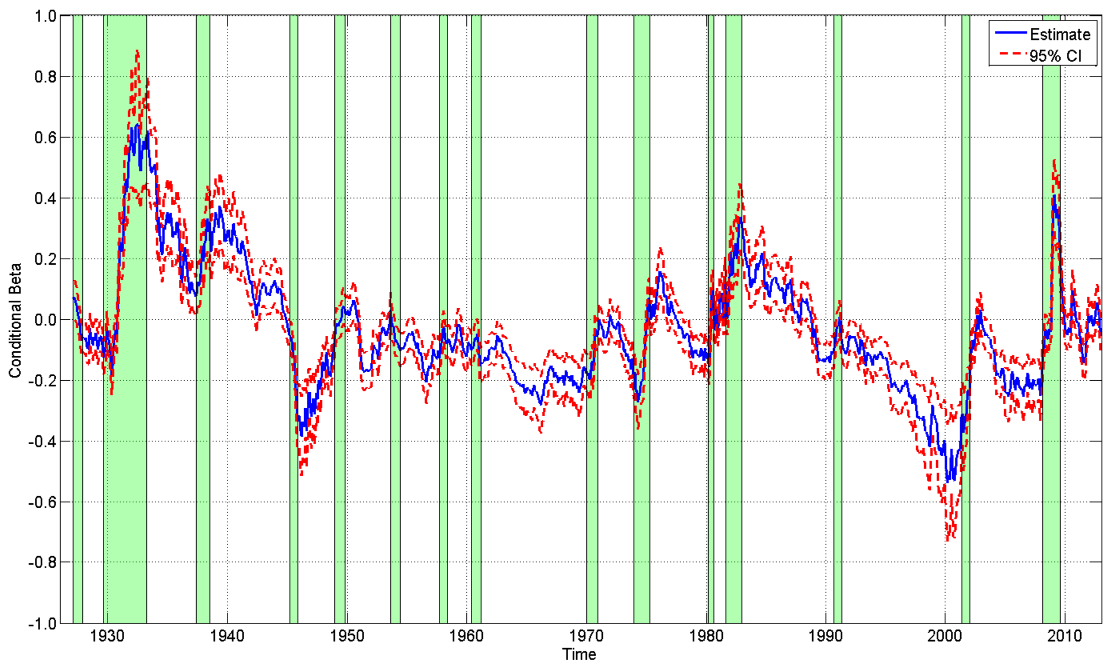

Figure 3: Conditional Alpha over the July 1963 to December 2012 Period

Figure 3 plots monthly estimates of the value premium's conditional alpha in solid line along with its 95% confidence bands in dashed lines. The conditioning variables are inflation (INF), the price-earning ratio (PE), and unemployment (UE) that we identify via the exhaustive variable selection. Shaded areas indicate the NBER recession periods.

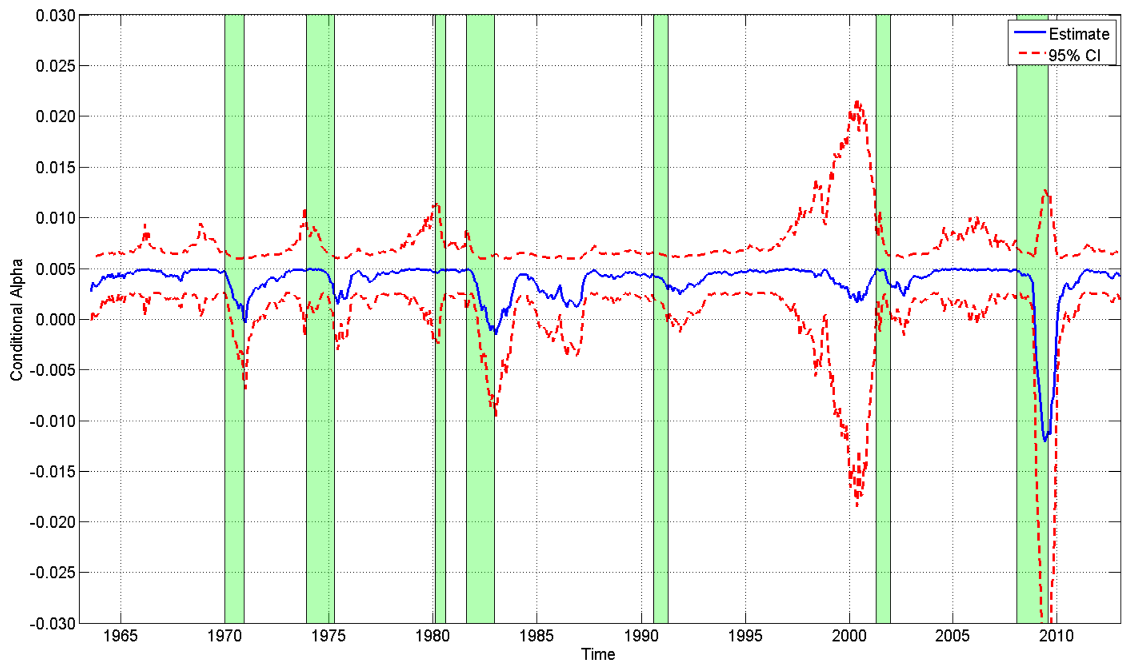

Figure 4: Ang and Kristensen's (2012) Conditional Beta

Figure 4 plots monthly estimates of Ang and Kristensen's (2012) conditional beta of the value premium, $BETA_{AK}$, in solid line along with its 95% confidence bands in dashed lines. The sample spans the July 1963 to December 2012 period. Shaded areas indicate the NBER recession periods.