The estimation of time-varying risks in asset pricing modelling using B-Spline method

Nurjannah, Solimun and Adji Rinaldo
Department of Statistics, Brawijaya University, Malang, Indonesia

E-mail: nj_ann@ub.ac.id

Abstract. Asset pricing modelling has been extensively studied in the past few decades to explore the risk-return relationship. The asset pricing literature typically assumed a static risk-return relationship. However, several studies found few anomalies in the asset pricing modelling which captured the presence of the risk instability. The dynamic model is proposed to offer a better model. The main problem highlighted in the dynamic model literature is that the set of conditioning information is unobservable and therefore some assumptions have to be made. Hence, the estimation requires additional assumptions about the dynamics of risk. To overcome this problem, the nonparametric estimators can also be used as an alternative for estimating risk. The flexibility of the nonparametric setting avoids the problem of misspecification derived from selecting a functional form. This paper investigates the estimation of time-varying asset pricing model using B-Spline, as one of nonparametric approach. The advantages of spline method is its computational speed and simplicity, as well as the clarity of controlling curvature directly. The three popular asset pricing models will be investigated namely CAPM (Capital Asset Pricing Model), Fama-French 3-factors model and Carhart 4-factors model. The results suggest that the estimated risks are time-varying and not stable overtime which confirms the risk instability anomaly. The results is more pronounced in Carhart’s 4-factors model.

1. Introduction
The main problem in the dynamic asset pricing model literature is that the set of conditioning information is unobservable therefore some assumptions have to be made. Hence, the estimation requires additional assumptions about the dynamics of risks. [1] suggests the use of a rolling window ordinary least squares (OLS) estimation of the market model. This data-driven approach has the advantage of no parameterization but requires the prior selection of the window length. The other estimators in the family of recursive least squares have been considered. Time-varying conditional betas have been nonparametrically estimated by [2], [3] and [4] assuming that betas vary smoothly over time and possibly nonlinearly.

In this paper, we try to estimate time-varying risks using nonparametric or pure data driven methodologies. We use the B-spline method which has not been widely used in finance. Since their introduction by [5], varying coefficient models have become an increasingly popular option for dimension reduction in nonparametric regression with multiple predictors.
Some of the previous studies have applied the varying-coefficient model such as functional regression analysis shown in [6] or [7], and the analysis of longitudinal data, e.g., [8]. The kernel estimation is proposed by [9], two-step estimate by [10], and smoothing spline by [11].

The different methods have different advantages. Smoothing spline is basically similar to kernel regression. The advantages of splines are their computational speed and simplicity, as well as the clarity of controlling curvature directly. Kernels however are easier to program, easier to analyze mathematically and extend more straightforwardly to multiple variables however it takes more time to run. It is also well known that the spline regression has a lot of advantages. The optimal convergence rate, in the nonparametric condition quartile function, partially linear model and partially linear model with mixed effect [12], can be obtained by the B-spline estimators of the nonparametric function.

In addition, asymptotic normality about B-spline estimators of the nonparametric regression is established in [13]. Thus, the B-spline approximation is considered in this paper. One important contribution of this paper is the application in Indonesian stock market. The literature on beta estimation is mainly applied to developed countries data. In contrast, Indonesia is an emerging market that provides the opportunity of analyzing the sensitivity of different beta estimators in this market.

2. The nonparametric time-varying risk estimator

2.1. General model

Let \( y \) be the response variable, \( x \) and \( t \) be the covariant vectors. The most general approach to model \( y \) is as a multivariate smooth nonparametric function \((x,t)\). Unfortunately, smoothing estimation of a general multivariate nonparametric regression may require large sample sizes when the dimensionality of the covariant is high, and the smooth results may be difficult to interpret. This phenomenon is called ‘the curse of dimensionality’.

A useful alternative is to impose some specific structure between \( y \) and \((x,t)\). As a dimensionality-reduction approach, we consider the varying-coefficient model of [5], which has the form

\[
y = x_1 \beta_1(t_1) + \cdots + x_p \beta_p(t_p) + \varepsilon
\]

where \( E(\varepsilon) = 0, E(\varepsilon^2) = \sigma^2 \) and \( \beta_i(\cdot), i = 1, \ldots, p \) are the smoothing functional coefficients. When \( t_i \)'s are the same, models (1) can also be written as

\[
y = x_1 \beta_1(t) + \cdots + x_p \beta_p(t) + \varepsilon
\]

A number of methods are proposed to estimate the functional coefficient in the model (2). Among them are the penalized least squares, the local likelihood for generalized varying-coefficient model, kernel estimation, and also smoothing spline. In this paper we will focus on B-spline method due to some advantages such as their computational speed and simplicity, as well as the clarity of controlling curvature directly [12].

Suppose we have observations \( y_1, \ldots, y_n \) from the varying-coefficient model (2), and denote the observed values of \( x_j \) and \( t \) for \( i^{th} \) case as \( x_{ij} \) and \( t_i \). Then, model (2) has the form

\[
y_i = x_{i1} \beta_1(t_i) + \cdots + x_{ip} \beta_p(t_i) + \varepsilon_i
\]

The functional coefficients, defined on \([a,b]\) are assumed \( \beta_i(t) \in C^m[a,b] \) i.e., \( \beta_i(t) \) has \( m \) continuous derivatives. The function \( \beta_i(t) \) will be approximated by function \( s_i(t) \) in the class \( S(m,Z) \). The set \( S(m,Z) \) is the collection of all polynomial spline of order \( m \) (degree \( m-1 \)) with the knots \( Z = \{a = z_0 < z_1 < \cdots < z_k < z_{k+1} = b\} \). By using the least squares criterion, the B-spline estimator of order \( m \) for \( \beta_i(t) \) is defined to the least squares minimizer \( \hat{\beta}_i(t) \) corresponding to

\[
\sum_{i=1}^{n} \left( y_i - \left( x_{i1} \beta_1(t_i) + \cdots + x_{ip} \beta_p(t_i) \right) \right)^2 =
\]
\[
\min_{s \in S(m, Z)} \sum_{i=1}^{n} \left( y_i - \left( X_{i1}s_1(t_i) + \cdots + X_{ip}s_p(t_i) \right) \right)^2
\] (4)

2.2. B-Spline estimation of the time-varying model

Let \( Z = \{a = z_0 < z_1 < \cdots < z_k < z_{k+1} = b\} \) be a subdivision of the interval \([a, b]\) by \( k \) distinct points. For each (fixed) set of knots \( Z \), the class \( S(m, Z) \) is a linear functional space of dimension \( N(= k + m) \). Any function \( s(t) \in S(m, Z) \) has the form

\[
s(t) = \sum_{i=1}^{N} a_i \pi_i(t) = \pi'(t) a
\] (5)

where \( \pi'(t) \) are normalized B-splines and \( \pi(t) = (\pi_1(t), \ldots, \pi_N(t))' \). Suppose that the non-decreasing sequence \((v_i)\) is obtained from \((z_{i1})_{i=1}^{k+1}\) by repeating \( z_0 \) and \( z_{k+1} \) each exactly \( m \) times. For \( m \geq 2 \), the B-spline basis for the family \( S(m, Z) \) is formed by the following \( k + m \) normalized B-spline [15].

\[
\pi_i(x) = (v_{i+m} - v_i)[v_i, \ldots, v_{i+m}](v - x)_{m-1}^\nu
\]

\[ i = 1, \ldots, N \]

where \( [v_i, \ldots, v_{i+m}] \phi \) denote the \( m^\nu \) divided differences on the \( (m+1) \) points \([v_i, \ldots, v_{i+m}]\) of the function \( \phi \) and \( a^\nu_i \) is mean \( a^\nu \) if \( \alpha > 0 \) and 0 otherwise.

The B-spline estimators of functional coefficients of model (2) are defined to be

\[
\sum_{i=1}^{n} \left( y_i - (x_{i1}\hat{\beta}_1(t_i) + \cdots + x_{ip}\hat{\beta}_p(t_i)) \right)^2 = \min_{\alpha \in \mathbb{R}^{N} \cap \sum_{i=1}^{n} \sum_{l=1}^{m} \left( y_i - (x_{i1}\pi'(t_i)\alpha_1 + \cdots + x_{ip}\pi'(t_i)\alpha_p) \right)^2
\] (6)

Let

\[
Y = (y_1, \ldots, y_n)'
\]

\[
D = \begin{bmatrix}
X_{11}\pi'(t_1) & \cdots & X_{1p}\pi'(t_1) \\
\vdots & \ddots & \vdots \\
X_{n1}\pi'(t_n) & \cdots & X_{np}\pi'(t_n)
\end{bmatrix}
\]

\[
\alpha = (\alpha_1', \ldots, \alpha_p')'
\] (7)

The estimator of the \( l^\text{th} \) functional coefficient \( \beta_l(t), l = 1, \ldots, p \) is

\[
\hat{\beta}_l(t) = \pi'(t) \hat{\alpha}_l
\] (8)

Hence,

\[
\hat{\beta}(t) = I_p \otimes \pi'(t) \cdot \hat{\alpha} = I_p \otimes \pi'(t) \cdot (D'D)^{-1}D'Y
\] (9)

where \( \otimes \) is Kronecker product.

3. Data

The data comprise of the market value, book-to-market value and returns on 407 stocks traded on the Indonesian Stock Exchange over the period of July 1997 - December 2010. From this set of stocks, those with the longest possible time series were chosen. Hence the sample reduced to 177 stocks. The returns are computed assuming continuous compounding of returns. This analysis uses monthly data that spans 144 observations over the period from January 1999 to December 2010. The JSX Composite Index is used as a proxy for the market portfolio while SBI (Interest rate of Bank Indonesia) is used as a proxy for the risk-free rate.

The portfolio formation may conceal the security related information present in individual stocks. This is important especially for emerging market research such as in Indonesia where the market is driven by a few blue chip stocks. Therefore this study will perform tests at the portfolio level as well as at the firm level. This study analyses size-sorted portfolio. When forming size-sorted portfolios,
market value is used to sort the stocks into 15 portfolios, from the lowest to the highest based on the trading volume. The first and the last portfolio consists of 17 stocks, while the rest comprises 11 stocks each. The portfolio return is calculated as the equally weighted average return of the stocks in the portfolio.

For each year from July of year $t-1$ to June of year $t$, we rank the stocks based on their size and book-to-market ratio. We then use these two rankings to calculate a 50% breakpoint for size, and 30% and 70% breakpoints for book-to-market. The stocks are subsequently sorted into two size groups and three book-to-market groups based on these breakpoints. In addition, the stocks above the 50% size breakpoint are designated B (for big) and the remaining 50% are designated S (for small). In addition, the stocks above the 70% book-to-market breakpoint are designated H (for high), the middle 40% are designated M (for medium) and the firms below the 30% book-to-market breakpoint are designated L (for low). We form six value-weighted portfolios, S/L, S/M, S/H, B/L, B/M and B/H as the intersection of size and book-to-market groups. Note that the number of firms in each of the six portfolios varies. This procedure is repeated every year. SMB is the equally-weighted average of the returns on the small portfolios minus the returns on the big stocks portfolios.

Similarly, HML is the equally-weighted average of the returns on the value stock portfolios minus the returns on the growth stock portfolios: For each month from July of year $t-1$ to June of year $t$, stocks are ranked by their size and prior performance. The size is based on the market value at the end of June in year $t-1$ whereas the prior performance is based on the previous 11-month nominal stock return lagged one month. One month is skipped between portfolio formation and holding periods in order to avoid any potential microstructure biases. Winners (W) are the top 30% of the total stocks with the highest average prior performance. Losers (L) are the bottom 30% of the total stocks with the lowest average prior performance. Neutral (N) are the remaining 40% of the stocks. Six value-weighted portfolios (S/W, S/N, S/L, B/W, B/N, and B/L) are formed at the intersection of size and prior performance. The equally weighted monthly returns on the six portfolios are calculated each month over the 12 months following portfolio formation. WML (winner minus loser) is the equally-weighted average of the returns on the winner stock portfolios minus the returns on the loser stocks portfolios.

4. Results and discussion

4.1. Model estimated

The asset pricing model relates the expected return on an asset to its systematic risk or betas. These betas are the sensitivity of the asset return to changes in the return of the market portfolio and some other factors such as the size effect, value effect and momentum effect. Three well-known asset pricing models are given in Equation

- **CAPM**

$$R_{it} = \alpha_i + \beta_i^M MR_t + \epsilon_{it}$$

- **Fama-French**

$$R_{it} = \alpha_i + \beta_i^M MR_t + \beta_i^S SMB_t + \beta_i^H HML_t + \epsilon_{it}$$

- **Carhart**

$$R_{it} = \alpha_i + \beta_i^M MR_t + \beta_i^S SMB_t + \beta_i^H HML_t + \beta_i^W WML_t + \epsilon_{it}$$

where:

- $R = \text{firm or portfolio excess return}$
- $\beta = \text{systematic risk}$
- $MR = \text{market excess return}$
- $SMB = \text{Small minus Big}$
- $HML = \text{High minus Low}$
- $WML = \text{Winner minus Loser}$
The CAPM specifies a relation between the expected risk premium on an individual asset (portfolio) and the associated systematic risk the expected excess return on a risky asset is equal to the expected excess return on the market portfolio multiplied by the systematic risk (market beta) of that particular asset. Fama and French in [16] expand the CAPM by adding size (SMB) and value factors (HML). It considers the empirical fact that value and small cap stocks outperform the market on regular basis. By including these two additional factors, the model adjusts for the tendency for poor performance, making this model a better model for evaluating manager performance. Fama and French model does well in explaining many existing market anomalies, except for the short-term momentum strategy documented by [17]. In order to capture the one-year momentum anomaly, Carhart in [18] adds a fourth factor into the Fama French, which is nearly orthogonal to the three factors. Carhart’s fourth factor is based on longing on winner-stocks as well as shorting on loser-stocks on the basis of their returns over the previous year. To mimic such a momentum factor (WML) or premium, WML is defined as the difference between the return on a portfolio of winner-stocks and the return on a portfolio of loser-stocks.

4.2. Findings
There are 177 firms included in the analysis but we will discuss several selected firms in more detail. For comparison purposes, the selected firms are the same firms as discussed in the previous paper. The firms are PT Bakrieland Development Tbk, PT Buana Finance Tbk, PT Davomas Abadi Tbk, PT Indal Aluminium Industry Tbk, PT Maskapai Reasuransi Indonesia Tbk, and PT Prima Alloy Steel Universal Tbk.

Figure 1 shows the estimated values of the nonparametric time-varying betas together with their confidence intervals. These curves were computed using cubic B-splines with selected knot of 8. As illustrated in Figure 1, the estimated betas are not stable over time and most of them are significantly different from zero. However, there is no clear pattern of how these betas vary over time. Most of the betas have fluctuated around their mean. In some stocks such as Bakrieland Dev, Maskapai Reasi Indo and Prima Alloy Steel, the betas fluctuated less compared to Davomas Abadi and Buana Finance which have fluctuated more. Generally, the market betas and the value-effect betas slightly decrease as time increases while the size-effect betas and momentum-effect betas are mostly stable overtime. In summary, 167 out of 177 stocks (95%) market betas are significantly different from zero, 90% size-effect betas are significantly different from zero, 91% value-effect betas are significantly different from zero and 67% momentum-effect betas are significantly different from zero.

The results reveal that generally the nonparametric time-varying beta estimates differ across the stocks. Figure 2 presents the boxplot of nonparametric time-varying beta for all firms. From Figure 2 we can see that different stocks have different beta characteristics. The beta of some stocks have more spread, large range and interquartile range, few outliers, and evidence of asymmetric distribution, while some firms have more dense beta distributions.

Generally, from 177 firms, we observe structural breaks during 2001-2002 and 2009-2010 where the magnitude of the betas change the direction. These are the periods where the presidential elections were held in Indonesia. Figure 3 presents the boxplot of the four betas for 15 size-sorted portfolios. The market beta in this case show evidence of fluctuation. Moreover, the betas tend to slightly increase as the portfolio size is larger. In contrast, the size-effect betas tend to decrease as the portfolio size increases. On the other hand, the value-effect betas and the momentum-effect betas of the portfolios appear to be less sensitive to portfolio size. There is also evidence of asymmetric in the betas.

M, S, H = estimator index for market, size and value respectively;
i = cross section index; t = time index
Figure 1. Nonparametric time-varying beta estimates for selected firms
Figure 2. Boxplot of nonparametric time-varying betas for firm-level analysis

Figure 3. Boxplot of nonparametric time-varying betas for size-sorted portfolio
5. Conclusion

This paper investigates the estimation of time-varying asset pricing model using B-Spline method. The flexibility of the nonparametric setting avoids the problem of misspecification derived from selecting a functional form. The results reveal that the estimated betas are not stable over time and most of them are significantly different from zero. Different stocks (portfolios) have different beta characteristics. The betas of some stocks (portfolios) are more spread while the betas of some others have relatively dense distributions. Further, there is evidence that in many cases distributions are asymmetric. There are instances where the betas fluctuate in opposite directions over short time intervals. Generally the nonparametric betas of large market cap portfolios are larger than in the small market cap portfolios. The accuracy of betas, as the measure of systematic risk is crucial in determining investment strategies and pricing individual equities in security analysis and company valuation. The results show that the beta risk changes through time with changing economic environment. Dynamics of time variation of betas also differ across industries. For investors, the time-varying of the risks are important features that may be considered in their asset allocation and risk management decisions.

References
[1] Fama E F and MacBeth J D 1973 Risk return and equilibrium: Empirical tests Journal of Political Economy 81 607
[2] Esteban M V and Orbe-Mandaluniz S 2010 A nonparametric approach for estimating betas: the smoothed rolling estimator Applied Economics 42 1269
[3] Li Y and Yang L 2011 Testing conditional factor models: A nonparametric approach Journal of Empirical Finance 18 972
[4] Ang A and Kristensen D 2012 Testing conditional factor models Journal of Financial Economics 106 132
[5] Hastie T and Tibshirani R 1993 Varying-coefficient models Journal of the Royal Statistical Society Series B (Methodological) 55 757
[6] Hoover D R, Rice J A, Wu C O and Yang L-P 1998 Nonparametric smoothing estimates of time-varying coefficient models with longitudinal data Biometrika 85 809
[7] Eubank R, Huang C, Maldonado Y M, Wang N, Wang S and Buchanan R 2004 Smoothing spline estimation in varying-coefficient models Journal of the Royal Statistical Society: Series B (Statistical Methodology) 66 653
[8] Wu C O and Chiang C T 2000 Kernel smoothing on varying coefficient models with longitudinal dependent variable Statistica Sinica 10 433
[9] Wu C O, Chiang C T, and Hoover D R 1998 Asymptotic confidence regions for kernel smoothing of a varying-coefficient model with longitudinal data Journal of the American statistical Association 93 1388
[10] Fan J and Zhang W 1999 Statistical estimation in varying coefficient models The Annals of Statistics 27 1491
[11] Chiang C T, Rice J A and Wu C O 2001 Smoothing spline estimation for varying coefficient models with repeatedly measured dependent variables Journal of the American Statistical Association 96 605
[12] He X, Zhu Z Y, and Fung W-K 2002 Estimation in a semiparametric model for longitudinal data with unspecified dependence structure Biometrika 89 579
[13] Zhou S, Shen X and Wolfe D 1998 Local asymptotics for regression splines and confidence regions Ann Statist 26 1760
[14] Chambers J M and Hastie T J 1992 Statistical methods in S Wadsworth and Brooks/Cole 196-246
[15] De Boor C 1978 A practical guide to splines, Mathematics of Computation
[16] Fama E F and French K R 1996 Multifactor explanations of asset pricing anomalies The Journal of finance 51 55
[17] Jegadeesh N and Titman S 2011 Momentum Annu Rev Financ Econ 3 493
[18] Carhart M 1997 On persistence in mutual fund performance The Journal of finance 52 57