DATA 201: Time Series Analysis
Linear Regression Model
Lecture 6: Application: Are Hedge Fund Returns Nonlinear?

Lulu Wang

Data Analytics
Dickinson College

2/7/2025

Introduction

CAPM

Functional Forms

Application: Are Hedge Fund Returns Nonlinear?

Objectives

- Review the Linear Regression Model (LRM)
- Discuss its estimation in Python
- Explore applications in financial data

I will review the regression model in broad terms and more details can be found in an introductory statistics/econometrics textbook, such as:

- Stock and Watson, *Introduction to Econometrics*, Pearson
- Wooldridge, *Introductory Econometrics*, A Modern Approach, South-Western

The Linear Regression Model (LRM)

- The linear regression model is given by:

$$Y_t = \beta_0 + \beta_1 X_t + \varepsilon_t,$$

  where:
  - Y_t: dependent variable at time t
  - X_t: independent variable / factor / predictor at time t
  - β_0, β_1: coefficients to be estimated
  - ε_t: error term (mean zero, variance σ^2)
- The **expected** (or average) value of Y_t given X_t is:

$$E(Y_t \mid X_t) = \beta_0 + \beta_1 X_t.$$

Application: Stock Return Risk Analysis

- **Objective**: Explain excess return of a stock (risk premium) by decomposing it into different risk components (factors).
- **Two type of risk**
  - *Systematic Risk*: Market-wide influences (e.g., macroeconomic conditions).
  - *Idiosyncratic Risk*: Firm-specific factors (e.g., company news, management changes).

Interpretation and CAPM

- **Interpretation**:
  - β_0: the expected value of Y_t when $X_t = 0$
  - β_1: the expected change of Y_t for a unit change of X_t
- The Capital Asset Pricing Model (CAPM) is an example of an LRM:

$$R_t^i = \beta_0 + \beta_1 R_t^{\mathrm{MKT}} + \varepsilon_t,$$

where R_t^i and R_t^{MKT} represent the excess stock and market returns, respectively.

Newey–West Example in Python

```
=== OLS Results (Default Standard Errors) ===
                            OLS Regression Results
==============================================================================
Dep. Variable:                 ep_crsp   R-squared:                       0.034
Model:                             OLS   Adj. R-squared:                  0.023
Method:                  Least Squares   F-statistic:                     3.061
Date:                 Mon, 03 Feb 2025   Prob (F-statistic):             0.0837
Time:                         13:02:36   Log-Likelihood:                -392.57
No. Observations:                   89   AIC:                             789.1
Df Residuals:                       87   BIC:                             794.1
Df Model:                            1
Covariance Type:             nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         -0.1903      5.274     -0.036      0.971     -10.673      10.292
DP             2.1865      1.250      1.749      0.084      -0.298       4.671
==============================================================================
Omnibus:                        3.676   Durbin-Watson:                   1.852
Prob(Omnibus):                  0.159   Jarque-Bera (JB):                3.340
Skew:                          -0.474   Prob(JB):                        0.188
Kurtosis:                       3.021   Cond. No.                         10.9
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

```
=== OLS Results (HAC/ Newey-West Standard Errors) ===
                            OLS Regression Results
==============================================================================
Dep. Variable:                 ep_crsp   R-squared:                       0.034
Model:                             OLS   Adj. R-squared:                  0.023
Method:                  Least Squares   F-statistic:                     2.135
Date:                 Mon, 03 Feb 2025   Prob (F-statistic):              0.148
Time:                         13:02:41   Log-Likelihood:                -392.57
No. Observations:                   89   AIC:                             789.1
Df Residuals:                       87   BIC:                             794.1
Df Model:                            1
Covariance Type:                   HAC
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const         -0.1903      5.626     -0.034      0.973     -11.218      10.837
DP             2.1865      1.496      1.461      0.144      -0.746       5.119
==============================================================================
Omnibus:                        3.676   Durbin-Watson:                   1.852
Prob(Omnibus):                  0.159   Jarque-Bera (JB):                3.340
Skew:                          -0.474   Prob(JB):                        0.188
Kurtosis:                       3.021   Cond. No.                         10.9
==============================================================================

Notes:
[1] Standard Errors are heteroscedasticity and autocorrelation robust (HAC) using 1 lags
```

- The coefficient estimates $(\hat{\beta}_0, \hat{\beta}_1)$ remain the same.
- Standard errors and p-values change to account for heteroskedasticity and/or autocorrelation.

Nonlinear regression models

- The Linear Regression Model (LRM) assumes a linear relationship between X and Y.
- A linear model implies that a **1-unit increase** in X results in a **constant** expected change in Y by β_1.
- However, some relationships are **nonlinear** (e.g., quadratic, logarithmic, exponential).
- Nonlinearity means that the **effect of X on Y varies** depending on the level of X.

Polynomial models

- One way to introduce **nonlinearity** is through the **Quadratic Model**:

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 X_t^2 + \varepsilon_t$$

- The **effect** of changing X by one unit on Y is given by:

$$\beta_1 + 2\beta_2 X$$

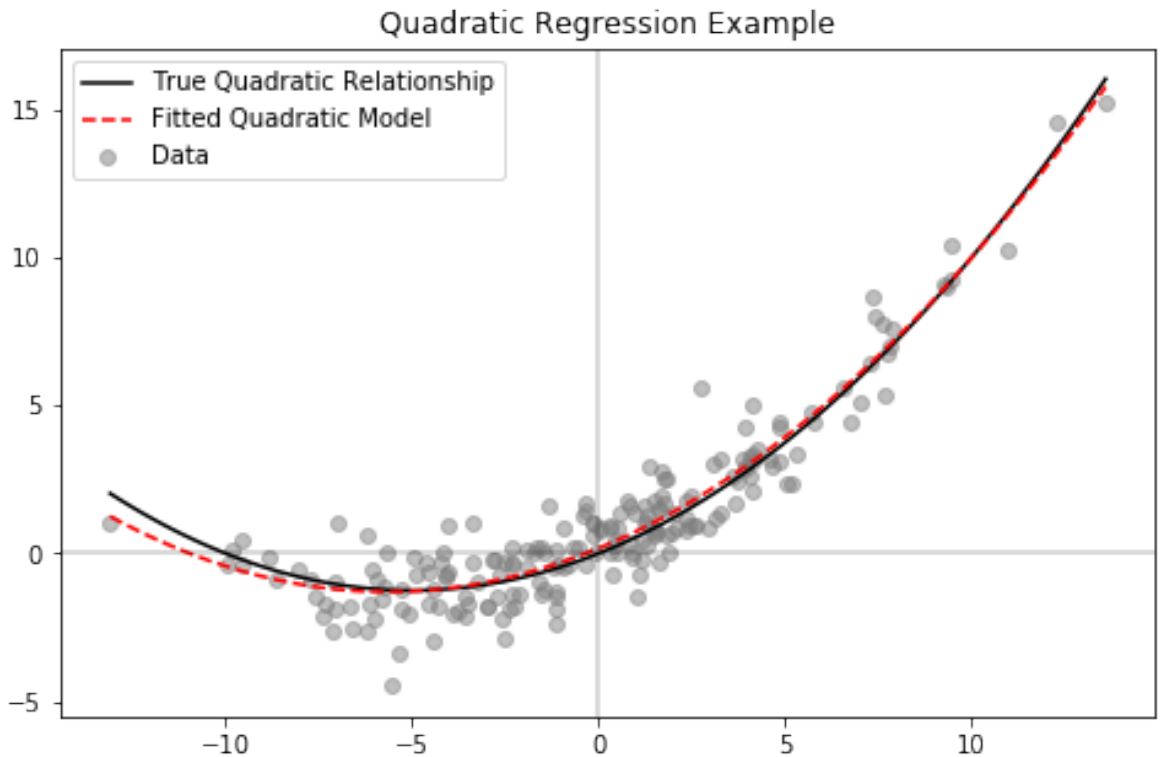
(depends on X)

- Quadratic regression can still be estimated using **OLS** by adding X^2 as a regressor.

Simulating a Quadratic Model in Python

Example: Simulate $Y_t = 0.5X_t + 0.05X_t^2 + \varepsilon_t$, with:

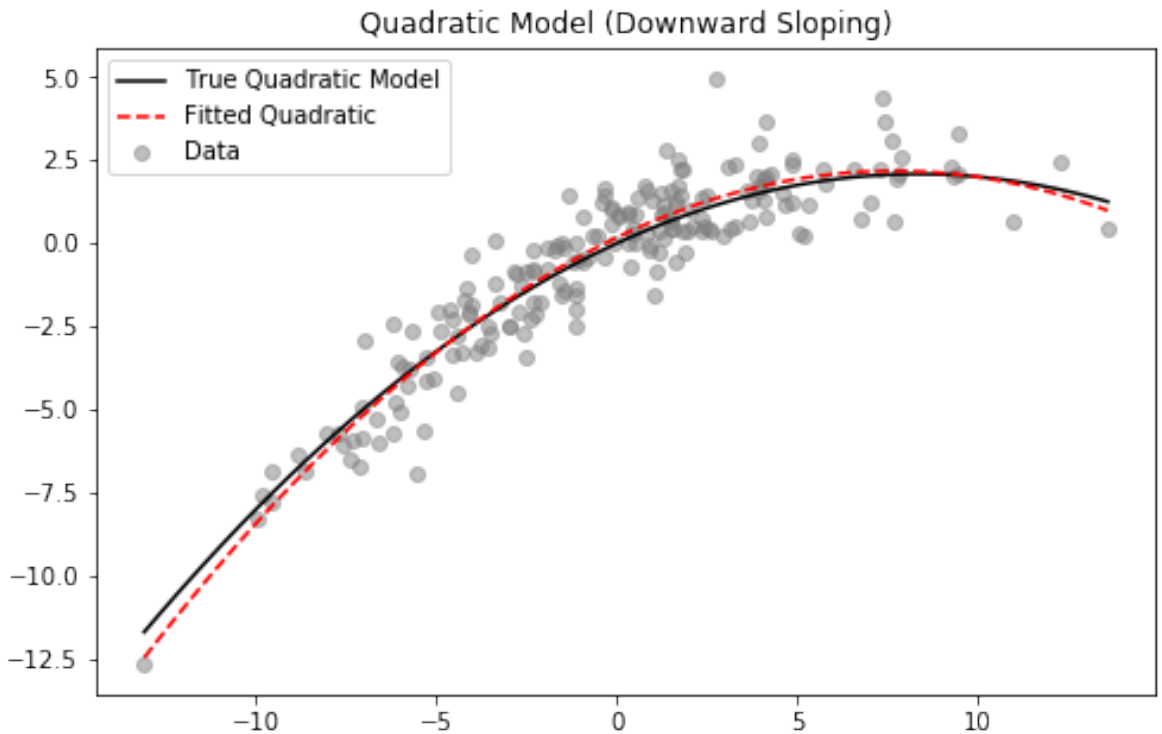
$$X_t \sim N(0, 25), \quad \varepsilon_t \sim N(0, 1)$$



Quadratic Model: Downward Sloping Parabola

- If the coefficient of X^2 is **negative**, the parabola slopes downward at the extremes.
- Below is a simulated quadratic model:

$$Y_t = 0.5X_t - 0.03X_t^2 + \varepsilon_t$$

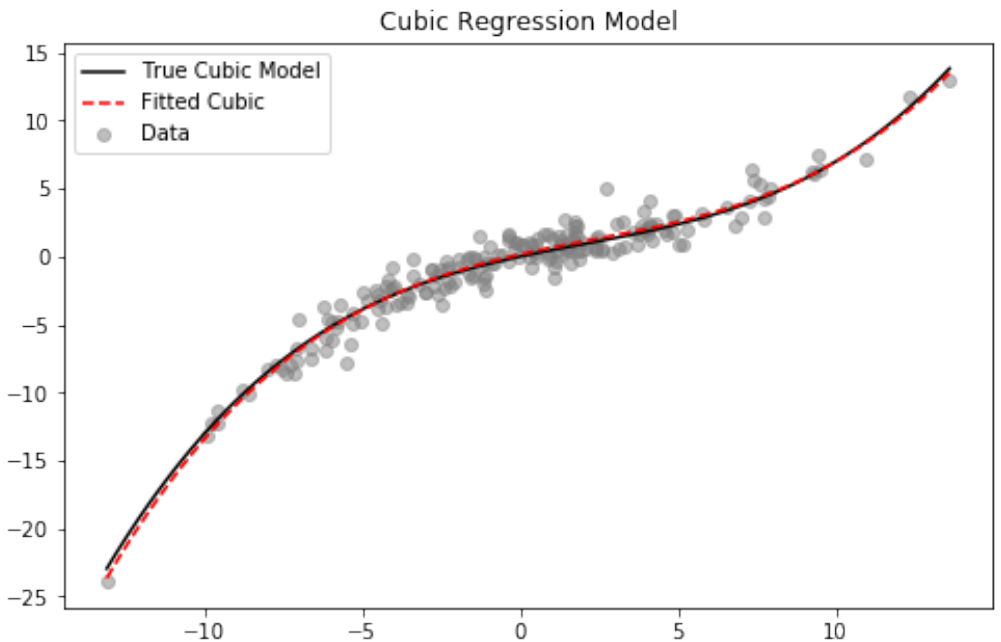


Cubic Regression Model

- A **cubic model** is useful when an **additional curvature** is needed to explain the relationship.
- The model is:

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 X_t^2 + \beta_3 X_t^3 + \varepsilon_t$$

- Including **higher-order terms** may introduce correlation among regressors, requiring careful evaluation.

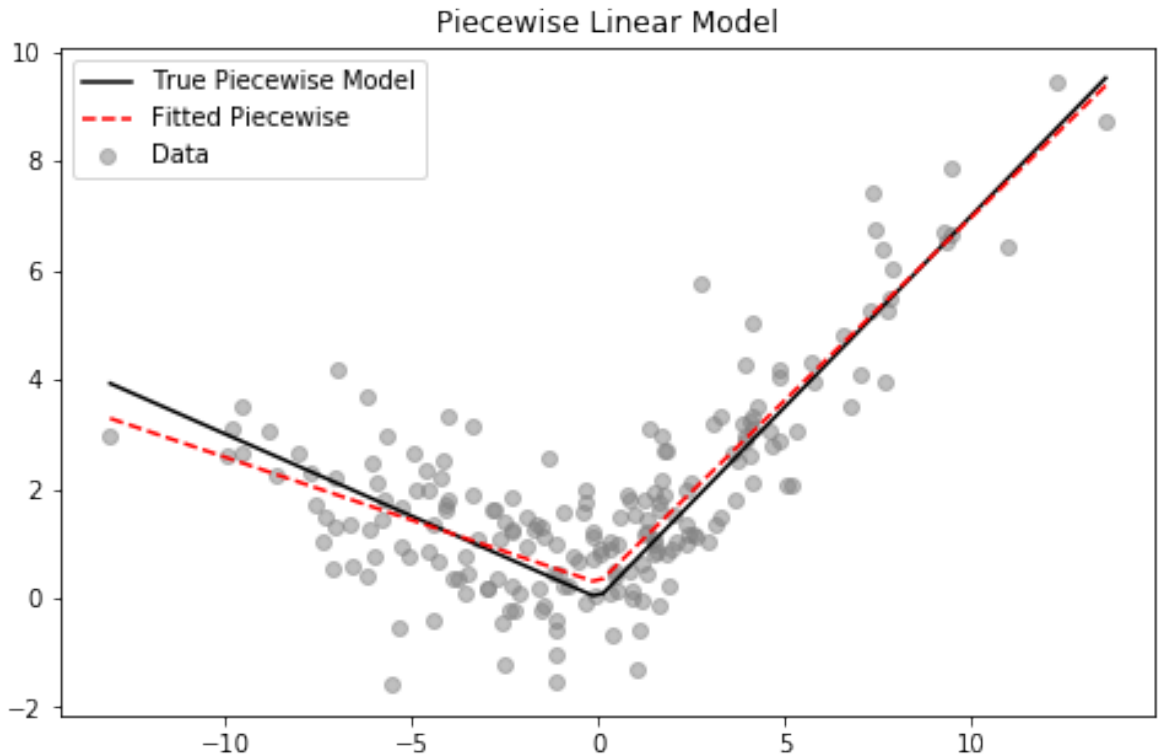


Piecewise Linear Model

- This model assumes different slopes **below/above** a threshold m:

$$Y_t = \beta_0 + \beta_1 X_t I(X_t \geq m) + \beta_2 X_t I(X_t < m) + \varepsilon_t$$

- Interpretation:
  - The **effect** of X_t on Y_t is different for $X_t \geq m$ vs. $X_t < m$.
  - The slopes are determined by β_1 and β_2.



Application: Are Hedge Fund Returns Nonlinear?

- **Hedge Funds (HF)** are investment portfolios that operate across multiple asset classes, including equities, bonds, commodities, and currencies.
- HFs can employ **derivative products** (options, futures, swaps) and **leverage** to control risk and amplify returns.
- Unlike traditional mutual funds, hedge funds are designed to deliver **absolute returns**, often independent of market direction.

Hedge Fund Indexes

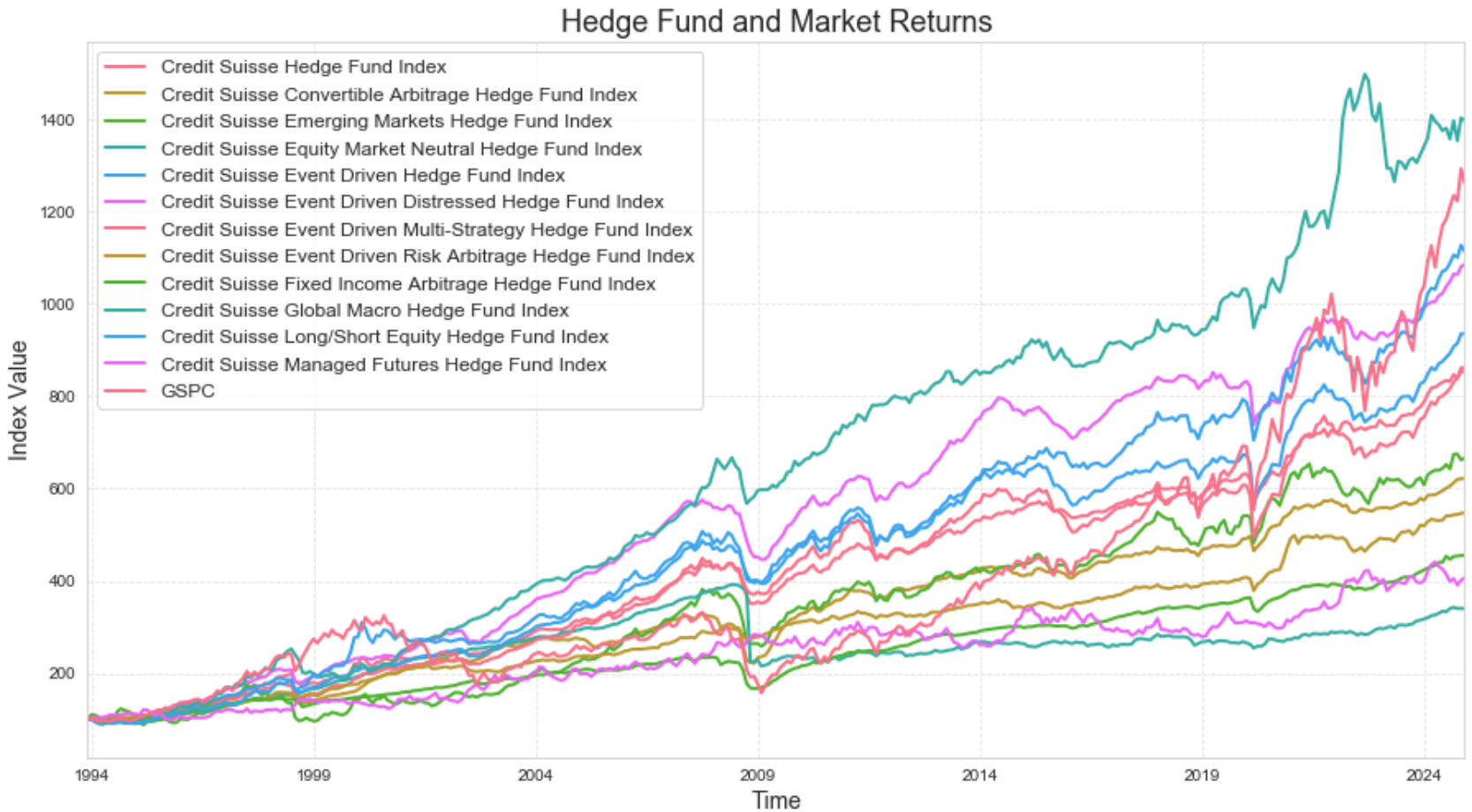
- The data for this analysis is sourced from the Hedge Fund Indexes, which aggregate the performance of a large set of hedge funds.
- These indexes are constructed by **averaging returns** across multiple funds, providing a broad overview of hedge fund performance.
- The indexes include an overall *Hedge Fund Index* as well as indexes tailored to specific strategies.
- **Data Availability**: The data is **publicly available** but may require registration for access.

Overview of Hedge Fund Strategies

- The indexes cover a range of **specific hedge fund strategies**, including:
  - **Dedicated Short Bias**: Focuses on short-selling to profit from declining stock prices.
  - **Emerging Markets**: Invests in developing economies with higher growth potential and volatility.
  - **Equity Market Neutral**: Balances long and short equity positions to minimize market exposure.
  - **Event Driven**: Capitalizes on corporate events like mergers, acquisitions, or bankruptcies.
  - **Global Macro**: Invests based on macroeconomic trends, using a wide range of asset classes.
  - **Long/Short Equity**: Combines long positions in undervalued stocks with short positions in overvalued stocks.
- **Unfortunately,** these indexes aggregate across many funds that makes it difficult to capture nonlinearities that might be present at the individual HF level

Time series of HF Index Value

- Plotting the index value for the strategy

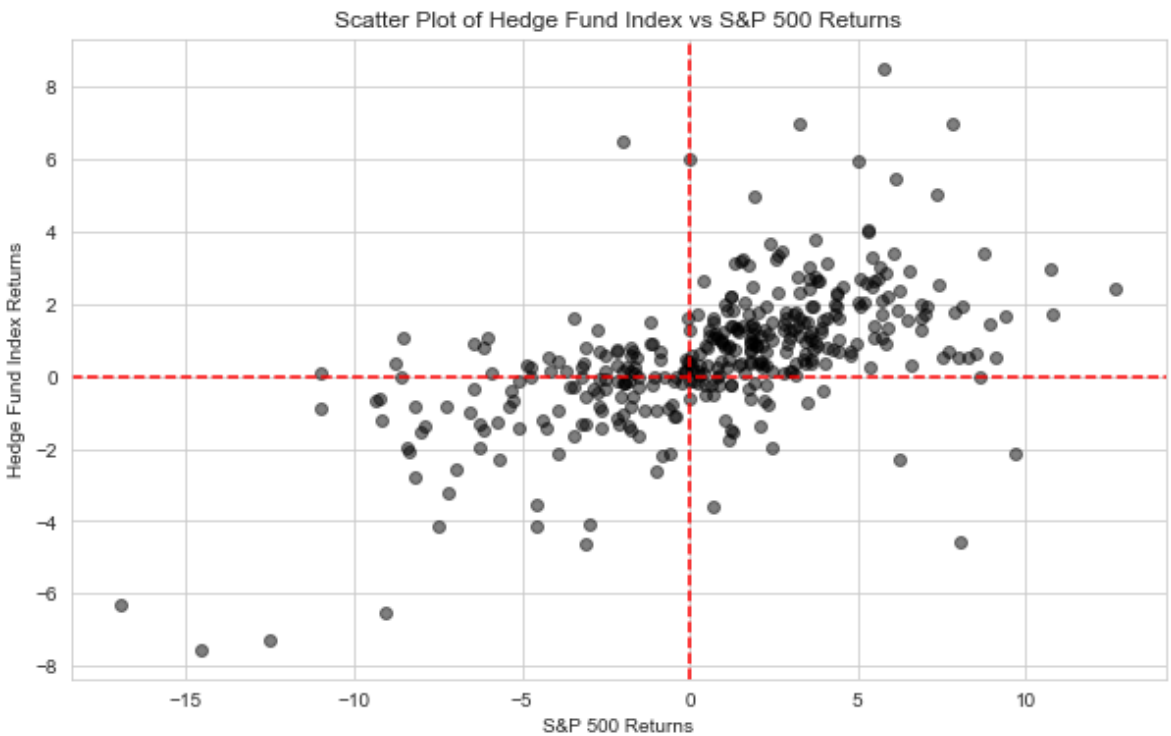


Descriptive statistics

	Mean	Std Dev	Skewness	Kurtosis	Max	Min
0						
Credit Suisse Hedge Fund Index	0.60	1.87	-0.25	3.89	8.53	-7.55
Credit Suisse Convertible Arbitrage Hedge Fund Index	0.51	1.70	-2.59	18.51	5.81	-12.59
Credit Suisse Emerging Markets Hedge Fund Index	0.58	3.59	-0.77	6.93	16.42	-23.03
Credit Suisse Equity Market Neutral Hedge Fund Index	0.37	2.43	-12.80	211.87	4.34	-40.45
Credit Suisse Event Driven Hedge Fund Index	0.62	1.87	-2.32	13.62	6.75	-13.47
Credit Suisse Event Driven Distressed Hedge Fund Index	0.66	1.76	-1.98	11.95	6.90	-12.45
Credit Suisse Event Driven Multi-Strategy Hedge Fund Index	0.60	2.07	-2.17	12.94	6.89	-15.59
Credit Suisse Event Driven Risk Arbitrage Hedge Fund Index	0.47	1.27	-0.61	5.38	6.05	-6.35
Credit Suisse Fixed Income Arbitrage Hedge Fund Index	0.42	1.40	-4.60	36.93	4.33	-14.04
Credit Suisse Global Macro Hedge Fund Index	0.74	2.50	0.13	4.25	10.60	-11.55
Credit Suisse Long/Short Equity Hedge Fund Index	0.68	2.50	-0.06	3.84	13.01	-11.43
Credit Suisse Managed Futures Hedge Fund Index	0.43	3.17	0.04	0.02	9.95	-9.35
GSPC	0.78	4.35	-0.60	0.93	12.68	-16.94

LRM for the HF Index

- Are the HF returns sensitive to movements in the U.S. equity market?
- Positive correlation, although the HF returns range between -8% and +8% while the equity index between 20% and 12%.



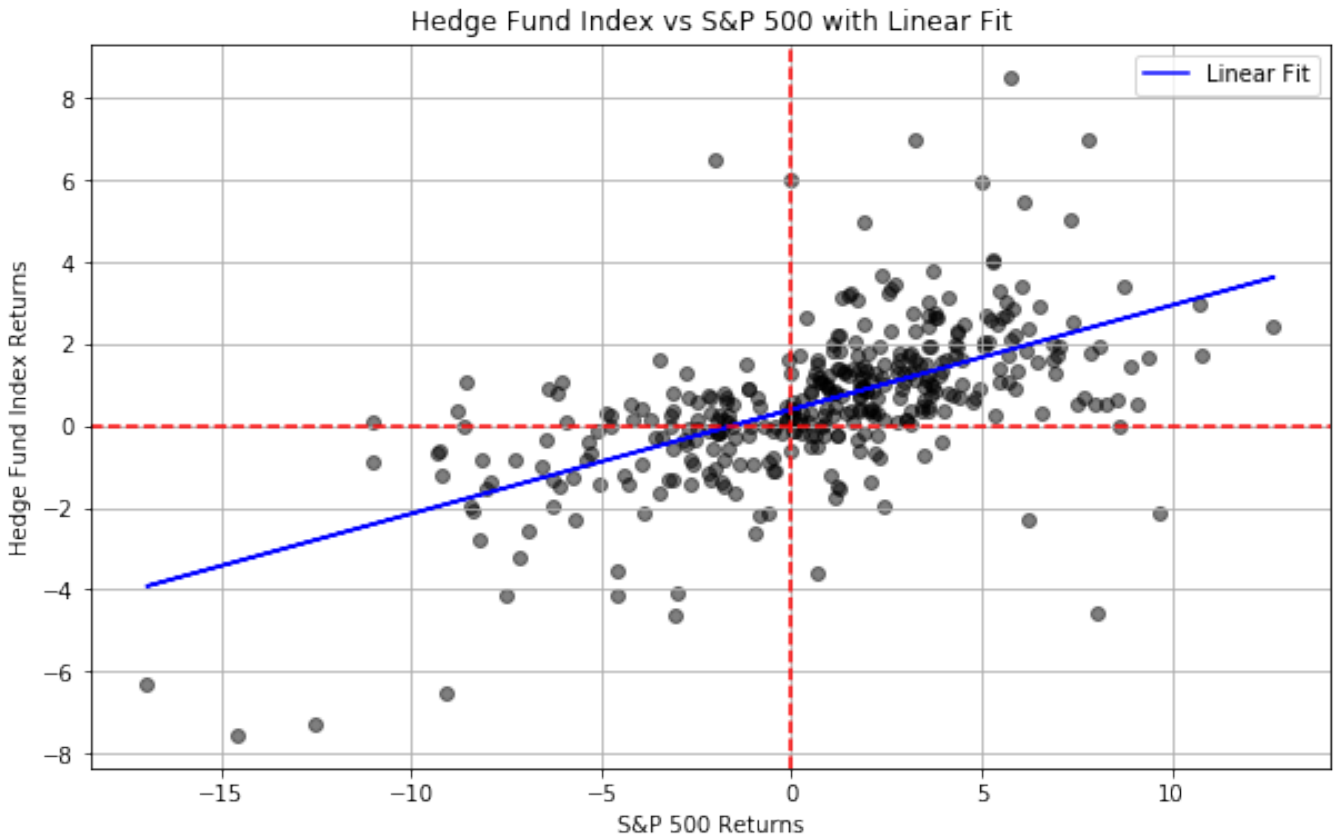
LRM for the HF Index (Cont.)

- The R^2 of the regression is 0.35
- The exposure of the HF Index to market risk is 0.2546
- Large and positive alpha
- Both alpha and beta are statistically significant at 1%

```
=== OLS Results (HAC/ Newey-West Standard Errors) ===
                            OLS Regression Results
==============================================================================
Dep. Variable:     Credit Suisse Hedge Fund Index    R-squared:                       0.349
Model:                                        OLS    Adj. R-squared:                  0.347
Method:                             Least Squares    F-statistic:                     90.34
Date:                            Thu, 06 Feb 2025    Prob (F-statistic):           2.60e-19
Time:                                    22:33:44    Log-Likelihood:                -681.01
No. Observations:                             372    AIC:                             1366.
Df Residuals:                                 370    BIC:                             1374.
Df Model:                                       1
Covariance Type:                              HAC
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.3972      0.087      4.556      0.000       0.226       0.568
GSPC           0.2546      0.027      9.505      0.000       0.202       0.307
==============================================================================
Omnibus:                       45.224   Durbin-Watson:                   1.671
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              296.038
Skew:                           0.158   Prob(JB):                     5.20e-65
Kurtosis:                       7.359   Cond. No.                         4.49
==============================================================================
```

LRM for the HF Index (Cont.)

- We can add the fitted linear relationship given by $0.3972 + 0.2546 R_t^{GSPC}$ to the previous scatter plot.



Quadratic model for the HF Index

- To estimate a quadratic model we need to add the quadratic term to the linear regression.

$$R_t^{HF} = \alpha + \beta_1 R_t^{GSPC} + \beta_2 [R_t^{GSPC}]^2 + \varepsilon_t$$

Quadratic model for the HF Index (Cont.)

- Is the relationship between the S&P 500 and the HF Index nonlinear?
  - The p-value of the squared term is 0.010 so we reject the null hypothesis that the coefficient is equal to zero; so, the relationship is nonlinear.
  - The adjusted R^2 for the linear model is 0.35 and for the quadratic model is 0.38 and we conclude that the quadratic is preferred.

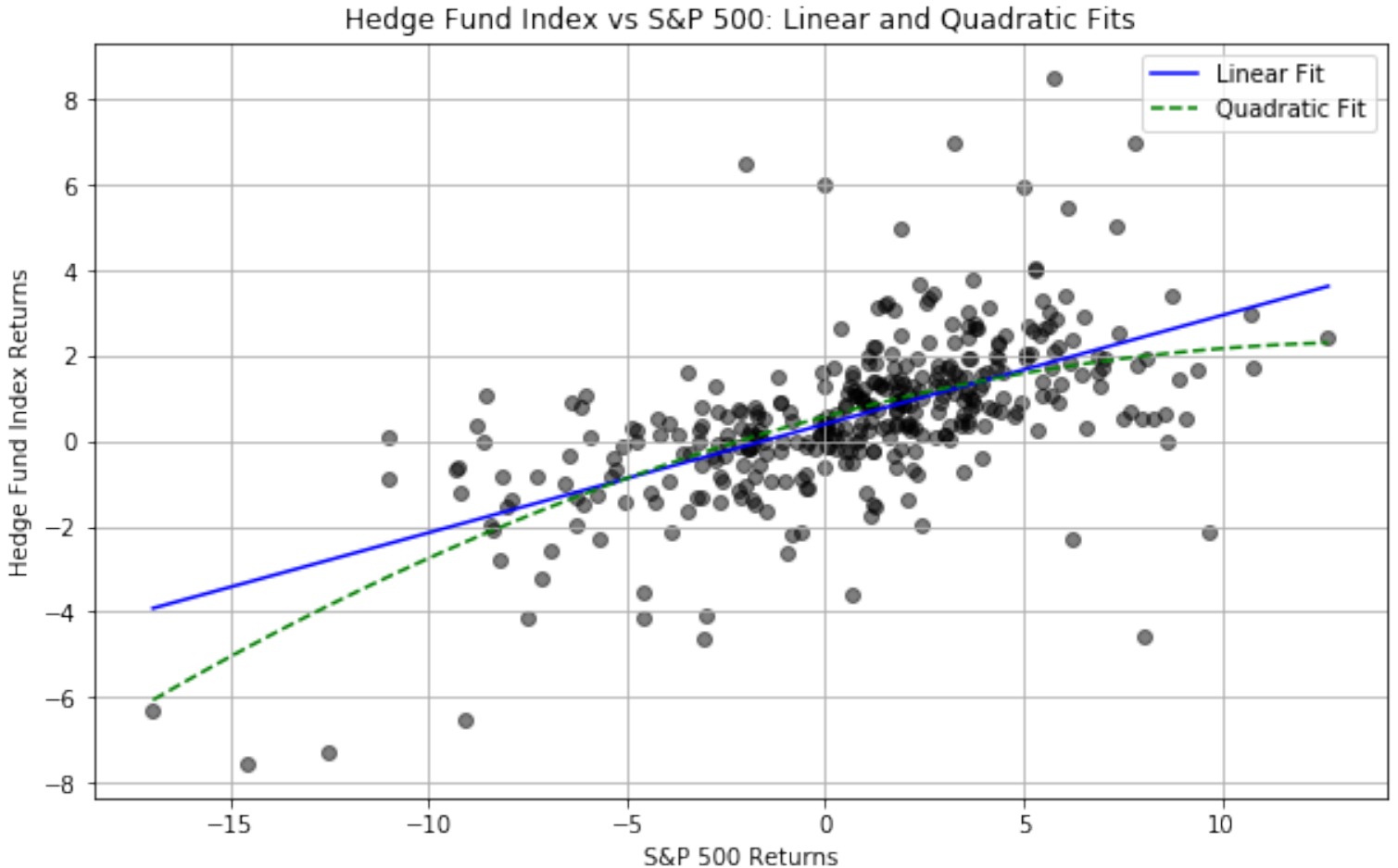
OLS Regression Results

Dep. Variable:	Credit Suisse Hedge Fund Index	R-squared:	0.369
Model:	OLS	Adj. R-squared:	0.365
Method:	Least Squares	F-statistic:	57.35
Date:	Thu, 06 Feb 2025	Prob (F-statistic):	2.05e-22
Time:	21:39:42	Log-Likelihood:	-675.37
No. Observations:	372	AIC:	1357.
Df Residuals:	369	BIC:	1368.
Df Model:	2		
Covariance Type:	HAC		

	coef	std err	z	P>|z|	[0.025	0.975]
const	0.5717	0.087	6.579	0.000	0.401	0.742
GSPC	0.2458	0.024	10.142	0.000	0.198	0.293
GSPC_squared	-0.0086	0.003	-2.576	0.010	-0.015	-0.002

Omnibus:	52.352	Durbin-Watson:	1.726
Prob(Omnibus):	0.000	Jarque-Bera (JB):	276.598
Skew:	0.420	Prob(JB):	8.66e-61
Kurtosis:	7.140	Cond. No.	44.2

Quadratic model for the HF Index (Cont.)



- This means that for large absolute market returns (both positive and negative), hedge fund returns are expected to be lower than what a linear model would predict.

Sources of Nonlinearity in Hedge Fund Returns

- **Nonlinear exposures** arise from:
  - **Derivatives**:
    - Use of options can create asymmetric payoff structures (e.g., *limited downside risk, unlimited upside potential*).
    - Futures and swaps can alter exposure to risk factors dynamically.
  - **Leverage**:
    - Amplifies returns, making the fund more sensitive to favorable movements in underlying assets.
    - However, it also increases risk, potentially leading to large losses during adverse market conditions.
- **Contrast with Mutual Funds**:
  - Mutual fund managers are **strictly regulated**, limiting their ability to use derivatives and leverage.
  - This results in more **linear exposure** to market risk factors.