

## Article

# The Dynamic Interplay of Consumption and Wealth: A Systems Analysis of Horizon-Specific Effects on Chinese Stock Returns

Faezeh Zareian Baghdad Abadi <sup>1</sup>, Ali Hashemizadeh <sup>2</sup> and Weili Liu <sup>3,\*</sup>

<sup>1</sup> China Center for Special Economic Zone Research, Shenzhen University, Shenzhen 518060, China; zareian@email.szu.edu.cn

<sup>2</sup> Key Laboratory for Resilient Infrastructures of Coastal Cities, Ministry of Education, Shenzhen University, Shenzhen 518060, China; ahz@szu.edu.cn

<sup>3</sup> College of Economics, Shenzhen University, Shenzhen 518060, China

\* Correspondence: liuwl@szu.edu.cn

## Abstract

This paper investigates the predictability of stock returns in the Chinese market through the lens of consumption–wealth dynamics within a broader financial system. We focus on two key state variables derived from modern consumption-based asset pricing models: the ratio of log surplus consumption (*scr*), from the habit-formation framework, and the log consumption–wealth ratio (*cay*), from the long-run cointegration framework. Using quarterly data from the CSI 300 index between 2012Q1 and 2018Q4, our system-based analysis reveals a horizon-dependent pattern of predictability. The results show that *scr* is a strong short-term predictor of excess stock returns, reflecting cyclical changes in risk aversion, whereas *cay* demonstrates superior predictive power over mid- to long-term horizons, consistent with its role as a proxy for long-run expectations. Interestingly, combining *scr* and *cay* does not improve predictive performance, suggesting that the economic mechanisms they capture are distinct rather than complementary in the Chinese market. These findings provide evidence on how interconnected macro-financial variables shape stock return dynamics, highlighting the importance of considering temporal horizons when modeling financial systems.

**Keywords:** asset pricing systems; consumption–wealth dynamics; financial systems; China



Academic Editors: Martins Adelaide, Susana Silva and Bruno Veloso

Received: 1 October 2025

Revised: 15 November 2025

Accepted: 20 November 2025

Published: 25 November 2025

**Citation:** Zareian Baghdad Abadi, F.; Hashemizadeh, A.; Liu, W. The Dynamic Interplay of Consumption and Wealth: A Systems Analysis of Horizon-Specific Effects on Chinese Stock Returns. *Systems* **2025**, *13*, 1066. <https://doi.org/10.3390/systems13121066>

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

### Background

The pricing of capital assets, especially stock, in the investment process is one of the most critical issues facing investors and capital market players [1]. Pricing is closely tied to risk, which is why financial economists have developed various models for risk measurement [2]. One of these patterns is the Capital Asset Pricing Model (CAPM). The model was initially welcomed, but after a while, it was sharply criticized by financial experts. Among the models that have been extended to address the defect of the previous model, the consumption-based CAPM (CCAPM) is considerable. In the first steps, Rubinstein, Lucas, Breeden, and Grossman and Shiller expanded this model theoretically to clarify the relationship between return and risk. This model comprehended static CAPM and intertemporal CAPM by connecting equity risk premia directly to systematic consumption risk. The essential advantage of the consumption-based model is that consumption is readily measured.

The central core of various asset pricing models is constructed with a consumption growth ratio. Some researchers presented a theoretical model relating consumption to another variable named the price–dividend ratio. They argued that consumption could make a balancing connection with this ratio. This connection is the enhancement of future dividends and the enhancement of expected future interest rates due to expected future consumption growth. Additionally, since stock prices affect wealth directly, their oscillation affects consumption through a variety of possible alternative avenues. Second, growth in the value of assets in each type of investment will increase its portion of the total savings portfolio of householders. As the third reason, although asset prices do not affect consumption directly, common macroeconomic factors do drive both of them.

CCAPM employs a firm theoretical basis constructed on the financial concepts of utility maximization for pricing assets. Due to the extremely nonlinear nature and non-partibility of consumption over time, accurate estimation is generally difficult. Nonlinear Two-Stage Least Squares (N-TSLS) is a powerful tool for explaining CCAPM problems in practical research. But selecting instrumental variables has a considerable effect on its estimates, and different sets of variables yield different results. If the variables are not chosen correctly, the results will not be reliable. When returning to standard CCAPM models, these have not been appropriately executed in practice. The results obtained from the market data of different countries show unsatisfactory results. Researchers modified consumption-based CAPM to produce more reliable empirical results in order to estimate international stock market returns.

China's financial system has changed quickly, and there have been ongoing changes to make the market more efficient, raising foreign investment, and deepening connections between the macro and micro levels of finance. As the world's second-largest economy, it is important to understand how return predictability works in the Chinese stock market [3]. This is important for both local asset pricing research and worldwide investment allocation. Consequently, these are our motivations for selecting the Chinese stock market as a case study for a study on the consumption-based CAPM. Although many types of research contribute to the literature on CCAPM, attention to the Chinese market has been much lower than its importance. China is the primary driver of world trade, economy, and export and import [4]; therefore, this country's stock markets have received enormous attention in theoretical and experimental research [5]. China's institutional and investor properties are unique, and experts have examined different aspects of the Chinese stock market. But capital pricing models for this market need further research and investigation. In this research, we aim to address the shortage of the related literature and investigate the power of CCAPM in explaining excess returns in the Chinese stock market. For this, two consumption state variables have been used, including the *log surplus consumption ratio* (called *scr*), and the *log consumption–wealth ratio* (called *cay*). The CCAPM approach and long-run hazards frameworks, have undergone extensive testing in developed countries; nevertheless, empirical data in significant developing markets are still scarce. This difference is particularly significant in China, where distinctive institutional characteristics, changing market involvement, and policy-induced liquidity cycles might modify the interplay between consumption, wealth, and anticipated returns. Our work enhances CCAPM research by investigating whether surplus consumption (*scr*) and the consumption–wealth ratio (*cay*), two theoretically based state variables, provide unique and horizon-dependent prediction information for stock returns in China.

Our study enhances the literature in several aspects. The capacity for consumption growth is fundamental to several asset price models. Certain academics proposed a theoretical model that connects consumption with the price–dividend ratio. They contended that consumption may exhibit an offsetting connection with this ratio, in which heightened

anticipated future consumption growth elevates projected future dividends while concurrently raising expected future interest rates. Moreover, stock prices are thought to influence consumption via other potential channels. This may mostly be seen via a direct wealth impact. Secondly, an appreciation in the asset prices of a particular investment or savings vehicle will augment its proportion within a household's overall savings portfolio. Third, asset prices may not directly influence consumption; rather, both may be influenced by a shared macroeconomic component. This research enhances the current literature by using stock returns rather than portfolio returns as the dependent variable. The primary variable for this study is consumption, with size and inflation serving as control factors. This study employs applied research to ascertain the correlation between consumption and stock returns using statistical parameters, utilizing empirical data from the China stock market.

## 2. Literature Review and Hypothesis Analysis

### 2.1. Consumption-Based CAPM

The topic chosen to be covered in this research is the *study of a consumption-based asset price model using the Chinese stock market*. CCAPM is an essential topic in financial issues. As a matter of fact, researchers in the field of economics have also made great efforts to provide such a framework. The ability of CCAPM models to explain asset prices and business cycle data at the same time has advanced significantly in recent decades [6]. A consumption-based CAPM was initially suggested by Lucas and others. They established this model to connect the consumption rate to the excess stock market [7]. Risk is the principal factor of diversity in expected asset returns. Quantifying risk accurately is a crucial first step in risk management, and this entails calculating correlations and volatility throughout the whole investable asset universe [8]. The commerce volume and transaction expenses can be relinquished in these models. To define the relationship between risk and return, these factors are assumed as [9]:

$$E(R_i) = R_f + \beta_i (R_m - R_f) \quad (1)$$

where  $E(R_i)$  is the expected return investors desire for the  $i$ th asset;  $R_f$  is the return of the risk-free rate, calculated by period average background;  $R_m$  is the return of the market portfolio, calculating by risk and the period average background; and  $\beta_i$  is a systematic risk measurement for the  $i$ th asset.

The following describes these assumptions:

1. Many adequate investors are price-takers;
  2. The investment period is the same for all investors;
  3. Taxes and transaction costs are ignorable;
  4. The risk-free rate is a real reference rate for borrowing and lending;
  5. Expected return and variance have the most attention from the perspective of investors;
- (a) A high *mean* and low *variance* are preferable;
  - (b) The market included all generally exchanged assets.

Vendrame et al. (2018) [10] examined various variables to demonstrate their power for predicting stock returns. Among the tested variables, three variables, including the price–dividend ratio, the default spread, and the term spread, have been displayed to present better performance. Such variables are linked with the business cycle. Therefore, it is concluded that the expected return and risk premiums change during the trading cycle. They argued that most changes in the ability to predict returns were due to changes in risk premiums, not betas. Furthermore, it is necessary to focus on time variation in risk premiums [10]. Some researchers like Fama and French, Campbell, and Hodrick provided essential pieces of evidence that the aforementioned variables can account for the time

variation of returns. However, some evidence of the inability of the price–dividend ratio to predict expected returns led to other variables being considered and tested to identify new predictor variables [11]. As an example, Boudoukh et al. (2007) explained that the net payout ratio brings more satisfaction because it more accurately records the amount of money distributed to shareholders and showed that it is more suitable for estimating future returns [12].

Therefore, we considered the dividend yield to determine whether it has any proper power for predicting consumption growth. The use of the mentioned variable has not been proposed recently and was suggested in Campbell and Schiller’s proposed model; they based their research on claims that dividend yields in the late 1990s overestimated the value of the stock market. Similar arguments are seen in Rapach and Wohar who explained the behavior of dividend yields, and also investigated the reasons for their deviation. Additionally, Campbell and Shiller (1988) proposed a model that can expand the value proposition model for stock prices to cover a component of market sentiment, allowing prices to deviate from principles [13,14].

The consumption-based CAPM assesses assets based on their exposure to the risk of macroeconomic factors; unlike the traditional CAPM, which only judges their returns against the returns of other assets. The method in question thus establishes a connection between the financial sector and the economy at a tangible level. Given its principles, assets that exhibit positive growth in consumption are less noteworthy to decision-makers because they do not provide security in difficult situations. Hence, investors must be compensated with the expected return on this kind of asset. Despite the correct methodical basis, empirical support for the standard CCAPM in terms of describing stock returns is brittle and unreliable [15].

There are different interpretations of CAPM pricing rules based on consumption. For example, some researchers have claimed that the price of an asset with a random repayment is less than the expected repayment, while a random repayment is directly related to consumption. Although this interpretation of the CCAPM pricing rule is not theoretically complete, there is considerable evidence to support it, and it must be acknowledged to be right given some additional conditions [16].

Chang et al. (2019) [17] contributed to the body of research on consumption-based CAPM, building on the work of Lettau and Ludvigson (2001) [18] regarding the cay and the prediction of equity market variations. They accomplished this by (a) employing time-varying cointegration techniques to establish a time-varying index of the consumption–aggregate wealth ratio (cayTVP); (b) evaluating the predictive efficacy of the cayTVP measure against that of the cay and Markov-switching cay indices concerning excess stock returns and volatility in the United States from 1952:Q2 to 2015:Q3; and (c) examining the predictive capacity of the three indicators of the consumption–aggregate wealth ratio utilizing a nonlinear causality-in-quantiles test developed by Balcilar et al. (2017) [19]. They discovered that there is cointegration between labor income, asset wealth, and consumption, although in a time-varying manner. Generally speaking, cayTVP is significant in bear markets, cay performs well in normal to excellent market conditions, and cayMS typically performs well in bull markets and in periods of relatively bad performance. They thus came to the conclusion that, depending on the situation of the stock market, investors may learn valuable information from all three metrics [18,19].

Adhering to the well-known consumer-based CAPM, the over-performance of a market portfolio is a popular parameter that illustrates discrimination in cross-sectional returns. Labor income and consumption increase have perfect connections [20]. Marquez et al. (2014) [21] proposed a fundamental consumption-based model. Despite other models in asset pricing research that focus on cumulative liquidity risk, their proposed frame-

work is implemented by answering the representative customer–investor improvement challenge based on the idea of general consumption risk. Their best outcomes were obtained for their proposed framework under the ultimate risk characteristics.

Despite the standard CCAPM being deducted from the power utility, where the single parameter is consumption growth, it can improve the performance of the consumption-based model by comprising two factors, which are called innovation in market returns and innovation in consumption growth. It is possible to overcome a popular indictment of the mentioned model, in which the marginal utility of consumption is based only on current data on consumption. Then, the marginal consumption utility becomes a function of both consumption growth and the returns of the market [15].

In another study, Jacobs and Wang tested the importance of the idiosyncratic risk of consumption for cross-sectional changes in asset returns. These researchers understood that cross-partial variance in consumption increase and total growth rates were both important pricing parameters. The outcomes they obtained proposed that two-factor CCAPM outperforms the CAPM. Additionally, it performed nearly like the Fama–French three-factor framework [22].

## 2.2. Chinese Stock Market

China has a distinct environment in different aspects, such as social, economic, and cultural, in comparison with other countries of the world. Additionally, it has grown significantly in global trade and international commerce as a consequence of its enormous development [23]. China is the most prominent driver of the world economy and international trade by leading imports and exports around the world. Hence, all of the aspects of the economy of China (such as the Chinese stock market) are desirable areas to which experts and researchers can extend theoretical and empirical studies. Due to its unique institutional and shareholder properties, experts have explored its stock market from many various points of views. This includes the Shanghai stock exchange and the Shenzhen stock exchange. Recently, as China's economy has grown, so has its stock market [24]. This market has considerable differences compared to the U.S. market. The main difference is that the estimated high risk is due to the low correlation between the macroeconomy and the stock market in China, not the low volatility of market returns as in the United States [25]. Undoubtedly, determining the factors affecting stock returns in the volatile Chinese stock market is considerably more difficult [26].

The consumption pattern in China is another significant issue to consider when analyzing the stock market. People with less economic and financial knowledge may not reduce their consumption as their wealth and income are impaired. On the contrary, individuals with good financial literacy should strive to mitigate unexpected losses and adjust their consumption to align with the current situation. As a result, they are more susceptible to behavioral biases and market disruptions. Hence, heterogeneous household characteristics influence the decision-making process in consumption [27].

Focusing on this, Huang (2018) [28] compared the outcomes of traditional CAPM, including four different models. The four types are the Sharpe–Lintner CAPM, the Fama and French three-factor model, the Fama and French five-factor model, and the Carhart four-factor framework. The five-factor model is generally more effective than other models in terms of depicting the individual returns of Chinese stock markets. However, the inclusion of investment and profitability components only marginally enhances the model's outcomes, suggesting that both factors possess a maximizing effect but little explanatory power regarding stock returns. By calculating whether equity return dispersion, which is determined by the cross-partial normal deviation of stock returns, bears a significant price risk, Chen et al. (2015) [29] expanded on empirical research on asset pricing analyses for

the markets in question. Even after controlling for market, size, book-to-market, and idiosyncratic volatility impacts, the investigators noted that stock return dispersion is really systematically priced in the cross-section of returns in this improving market.

Le et al. (2025) [30] examined the impact of investor regret on asset prices within the Chinese stock market using data from 2000 to 2021. They concluded that equities related to increased investor remorse often provide better future returns. This study emphasized the influence of behavioral variables, such as regret, on stock values, particularly in developing markets like China. Given the complicated and sometimes contradictory outcomes in the previous research on assigning stock returns in China, the dispersion of stock returns as a systematic risk ratio presents a more significant proxy for risk in this improving market, where many aspects have changed dramatically [29]. Table 1 summarizes some research about the Chinese stock market and the stock markets of some other countries.

**Table 1.** Some CCAPM studies on the Chinese stock market and the stock markets of some other countries.

| Author(s)                            | Method(s)                                          | Feature(s) and Results                                                                                                                                                                                                                                                                                                                                                        | Case Study                                         | Source |
|--------------------------------------|----------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------|--------|
| Kang et al., (2011)                  | Conditional CAPM using macroeconomic indices.      | <ul style="list-style-type: none"> <li>- The conditioning variable had a durable power to predict market excess returns.</li> <li>- The suggested variable strongly forecasts excess stock returns.</li> <li>- Using the Hansen method for generalized empirical likelihood (GEL) estimation.</li> </ul>                                                                      | Korea                                              | [31]   |
| Itoy and Noda (2011)                 | Standard CCAPM                                     | <ul style="list-style-type: none"> <li>- That both factors predicted the CCAPM, the degree of risk aversion, and the rate of time discount change over time.</li> </ul>                                                                                                                                                                                                       | Japan                                              | [32]   |
| Liu and Wan (2012)                   | Linear and nonlinear Granger causality tests       | <ul style="list-style-type: none"> <li>- Examined the co-movement of the Shanghai stock market and exchange rates of Chinese currency (CNY).</li> <li>- The stock price and exchange rate are significantly cross-correlated.</li> <li>- Examined the degree to which changes in asset portfolio stock prices reflect the trading patterns of diverse individuals.</li> </ul> | Shanghai Stock Exchange                            | [33]   |
| Koutmosa and Songb (2014)            | A range-based autoregressive volatility framework. | <ul style="list-style-type: none"> <li>- Information-driven investors believe that periods of declining stock prices are entirely tied to low-volume trading time, whereas rising stock prices are often associated with a significant quantity of trading time.</li> </ul>                                                                                                   | Shanghai Stock Exchange                            | [34]   |
| Vendrame, Guermat, and Tucker (2018) | Conditional CAPM                                   | <ul style="list-style-type: none"> <li>- There are severe time deviations in betas across their approaches and frameworks.</li> <li>- The regime-switching framework deducts the proper prediction of one-day-ahead value-at-risk.</li> </ul>                                                                                                                                 | USA, Germany, England, France, China, and Malaysia | [10]   |

Researchers have examined CCAPM in developed markets extensively; however, research on its application to China's distinctive stock market context remains scarce. China's market is distinctive because it has a lot of individual investors, people who do not know much about finance, and policies that change all the time. These things might modify how consumption and wealth connect to stock returns. Previous research has seldom analyzed the functionality of state indicators, such as the *scr* and the *cay*, over various temporal periods in China. Moreover, several prior studies inadequately addressed critical concerns such as structural fractures in the market and potential endogeneity among variables, and failed to conduct comprehensive diagnostic testing to guarantee dependable outcomes. We address these shortcomings in this research by using a CCAPM framework that incorporates *scr* and *cay*, utilizing sophisticated techniques to manage endogeneity and dynamics, and rigorously evaluating the model's validity via diagnostic assessments. This method helps us learn more about how stock returns are affected by wealth and consumption in the changing Chinese market.

### 2.3. Hypothesis Development

Even though the consumption-based CAPM significantly reduces the pricing errors of the portfolio-based framework, the standard strength utility CCAPM and the HF-CCAPM are able to describe the cross-partial variation in equity risk premia, whereas the CAPM is unable to do so. The power utility CCAPM can account for over 50% of the change in risk premia, and the habit formation model can account for over 90%. Therefore, according to Auer (2013), consumption-based parameters that include the surplus consumption coefficient and the consumption enhancement rate are important indicators [35].

In another study, Chen and Ludvigson state that habit formation models provide a very good explanation of the overall behavior of the stock market when it comes to the content of equilibrium consumption. The most notable habit model was extended by them. According to this hypothesis, individuals progressively broaden their purchasing patterns, making risk aversion anti-cyclical and changeable over time. The significant US equity premium for the US stock market may be explained by the model. One feature of the framework is that, although the risk-free rate is small and steady, the average risk aversion over time is significant. Therefore, without taking the risk-free rate into account, the suggested framework resolves the equity premium problem with considerable risk aversion. The enhanced consumption-based CAPM can forecast returns in global stock markets. Li and Zhong (2010) [22] investigate consumption-based CAPM with habit development, prediction, and the cross-partial of returns from global stock markets using quarterly data collected from 17 different nations. They asserted that the returns from the majority of established equity markets could be accurately described by the local and global stock markets. Their cross-partial analyses of CCAPM in habit development state that the model performs better than a three-factor international model comprising the unconditional global CCAPM, CAPM, and the conditional world CAPM. Thus, the following theory is taken into consideration:

**Hypothesis 1:** *The  $scr_t$  significantly predicts future stock returns and assumes that expected returns are inversely related to  $scr_t$ .*

A cointegration relation between log labor income ( $y_t$ ), log asset holdings ( $a_t$ ), and log consumption ( $c_t$ ) was discovered by certain researchers, and the  $CAY_T$  model they generated may predict excess returns on the U.S. stock market across short and intermediate time horizons. Additionally, cross-sectional stock market returns may be described by  $cay_t$  in CCAPM. Additionally, they contrasted its implementation outcomes with those of the Fama–French three-factor framework. Li and Zhong (2010) [22] discovered a positive

linear association between  $cay_t$  and the extra stock market return. According to McMillan (2013) [13], there is a strong relationship between dividend yield and consumption enhancement; hence, when the dividend yield is low, spending is high. Through predictions about future economic behavior, he contended that there is a steady association between the current dividend yield and future spending. Then, in order to predict future stock returns, it is necessary to combine the  $cay_t$  and the expected future growth in labor income. This will demonstrate that not only do the anticipated changes in labor income have a high degree of predictability in terms of future returns, but that the combination of the  $cay_t$  and expected changes in salaries accurately describes the change in the cross-section of returns. In the long term, however, the dividend yield responds to anticipated future behavior. According to him, there is a clear correlation between the effect of income from modern stock returns and increased spending. He did point out, nevertheless, that the positive coefficient's statistical significance varies throughout the six models. Ultimately, he focused on single-market regressions, even though the main result relies on panel estimation [13]. The following theory is then taken into account:

**Hypothesis 2:**  *$cay_t$  significantly predicts future stock returns and assumes that the expected returns are directly associated with  $cay_t$ .*

If consumption adheres to a random walk process, the wealth-to-consumption ratio in the consumption-based framework with habit formation should only rely on the consumption-to-habit ratio. Moreover, anticipated stock returns ought to vary only with delayed habit-based consumption if dividends and consumption adhere to random walk processes. Consequently, an inherent correlation should exist between the coefficients of wealth to consumption and consumption to habits, with the consumption-habit ratio elucidating the wealth-consumption ratio's ability to predict stock returns. Nevertheless, evidence indicates that consumption increasing is serially correlated, despite actual data often corroborating the random walk theory for dividends [36]. Thus, the following theory is taken into consideration:

**Hypothesis 3:** *The  $cay_t$  and  $scr_t$  are supplementary and have more predicting capacity than single-state variables.*

### 3. Data and Models

#### 3.1. Data

As discussed in the previous section, the critical state variables for the study of the consumption effect in stock returns are the  $scr$  and  $cay$ . As we have two crucial state variables, in this section, we define how the required data and measurements are collected. Our data includes *macroeconomic data* and other *financial data*. In China, the *National Bureau of Statistics of China* (NBSC) maintains a proper database of macroeconomic data. Online websites and the *Wind* database are employed to make a consumption and disposable labor income framework. These variables are time series in quarterly terms. For wealth data, we searched many local and international sources, the best available sources being the information supplied by the *CREDIT SUISSE GROUP*. This information is published each year. Here, we used the data from 2012Q1 to 2018Q4. In this period, there are 28 observations. The quarterly data used, and data covering other time periods, are adjusted to fit real per capita data.

There are three reasons why this period was selected. First, it shows a time of major changes in the market that were meant to make it more open, provide more people access to it, and encourage growth based on consumption. Second, during this time period, solid and regular quarterly data on household wealth and spending became accessible. Third, the window leaves out the huge changes that happened during the global financial crisis of 2008–2009 and the COVID-19 epidemic. This lets us concentrate on fundamental macro-financial interactions instead of crisis-specific aberrations.

We have chosen the CSI 300 as our stock market index since it is one of the most widely utilized indexes. One such stock market index is the CSI 300, which attempts to reflect the performance of the 300 most heavily traded companies on the Shanghai stock market by using capitalization as an indicator. The China Securities Index Company, Ltd. is responsible for compiling the index. Quarterly data was used. For the period beginning in 2012Q1 and ending in 2018Q4, we retrieved CSI300 statistics, the risk-free rate, dividend yields, dividend payout coefficients, and TRM from the RESSET database.

In our research, we use the Campbell and Cochrane external habit according to Li and Zhong (2010) [22] and Bianchi et al. to provide estimates of the surplus consumption ratio and consumption–wealth coefficient [37,38].

### 3.2. Variable Selection and Measurements Ratio

We have three kinds of variables including the following:

1. Dependent variable: excess return ( $r_t$ );
2. Independent variables:  $scr_t$  and  $cay_t$ ;
3. Control variables: dividend yield ( $dp_t$ ), dividend payout coefficient ( $de_t$ ) and government bond term spread ( $TRM_t$ ).

#### 3.2.1. Dependent Variable

Excess return is a dependent variable:  $r_{mt}$  is the log real return of the stock market ratio, and  $r_{ft}$  is the log real return on the risk-free coefficient, therefore  $r_t = r_{mt} - r_{ft}$ .

#### 3.2.2. Independent Variables

In this way,  $scr_t$  and  $cay_t$  are independent variables. As for the measurement of  $scr_t$  Campbell and Cochrane explain the existing CCAPM with external habit formation. Therein, surplus consumption is specified as

$$SCR_t = \frac{C_t - X_t}{C_t} \quad (2)$$

where  $C_t$  is defined as the per capita real usage, and  $X_t$  is defined as an external habit level that is obtained from previous consumption terms.

$cay_t$  contains human capital and household asset holding at time  $t$ . Integrated wealth, exclusively human capital, is not directly visible. So, to predict asset returns, we must identify an agent for human capital. We utilize after-tax labor salaries as a proxy for human capital as follows:

$$cay_t = c_t - \beta_a a_t - \beta_y y_t \quad (3)$$

#### 3.2.3. Control Variables

The considered control parameters are  $dp_t$ ,  $de_t$ , and  $TRM_t$ . Table 2 introduces these variables.

**Table 2.** The details of the considered control variables.

| Name                        | Symbol  | Calculation                                                                                                                                                  | Description                                                                                                                                                                                                                             |
|-----------------------------|---------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Dividend yield              | $dp_t$  | This is the coefficient of a corporation's yearly dividend compared to its share price.                                                                      | A stock investment's dividend-only return is predicted by the dividend yield. The yield will increase when the stock price declines and fall when the stock price rises, assuming that the dividend is neither increased nor decreased. |
| Dividend payout ratio       | $de_t$  | This is the overall volume ratio of dividends distributed to investors as a percentage of the firm's net income.                                             | An innovative, growing company with ambitious intentions to expand into new markets and provide innovative results might be excused for its low payout ratio as it is anticipated to reinvest most or all of its profits.               |
| Government bond term spread | $TRM_t$ | The China Bond Pricing Center provides this information as the difference between the China Bond Long Term Index and the China Bond Short-Middle Term Index. | It is a crucial factor that bond funders calculate when gauging the level of const for single or plural bonds.                                                                                                                          |

With these in mind, we establish a model to examine our hypothesis in Section 2.3. There are two approaches to forecasting expected returns. The first approach is analyzing their relationship with one state variable. As we have two state variables, such as  $scr_t$  and  $cay_t$ , based on the first approach there two formations, including the only  $scr_t$  and only  $cay_t$  [13,39].

The formula for the expected excess return on an asset with  $scr_t$  is

$$E_t [r_{i,t+1}^e] = \alpha + \alpha_1 z_t + \beta_1 scr_t + MA_1 + MA_2 \quad (4)$$

where  $z_t$  represents the vector of control variables, the slope ratio ( $\beta_1$ ) is constant, and  $MA_1$  and  $MA_2$  are moving average indicator models.

The return predicting regression equation with  $cay_t$  is as follows:

$$E_t [r_{i,t+1}^e] = \alpha + \alpha_1 z_t + \beta_2 cay_t + MA_1 + MA_2 \quad (5)$$

where  $\alpha$  is stable,  $\alpha_1$  represents the vector of continuous slope ratios of control parameters,  $\beta_2$  is the continual slope coefficient, and  $MA_1$  and  $MA_2$  are moving average indicator models.

If the expected returns are inversely related to  $scr_t$  and directly associated with  $cay_t$ , then Hypotheses 1 and 2 are established. The second approach is exploring this with two state variables [22]. Combining both state variables is matched with this approach. The equation of regression can be declared as

$$E_t [r_{i,t+1}^e] = \alpha + \alpha_1 z_t + \beta_1 scr_t + \beta_2 cay_t + MA_1 + MA_2 \quad (6)$$

where  $\alpha$  is stable,  $\beta_i$  is a stable slope coefficient, and  $MA_1$  and  $MA_2$  is moving average indicator models.

### 3.3. Descriptive Statistical Analysis

The fundamental descriptive statistics for the independent, dependent, and control parameters are declared in Table 3, which was taken from “Eviews.” Log excess returns on stock market indexes  $r$ ,  $dp_t$ ,  $de_t$ ,  $TRM_t$ ,  $scr_t$ , and  $cay_t$  are among the quarterly variables.

**Table 3.** Descriptive statistics of the considered factors.

| Indicators | Mean     | Median   | Maximum  | Minimum   | Standard Error | Autocorrelation |
|------------|----------|----------|----------|-----------|----------------|-----------------|
| $r$        | 3.12027  | 3.09823  | 3.27979  | 2.93912   | 0.01753        | 0.452           |
| $scr_t$    | −0.58781 | −0.59034 | −0.50139 | −0.69505  | 0.01039        | 0.871           |
| $cay_t$    | 14.86361 | 15.04799 | 16.56272 | 12.86988  | 0.210268       | −0.025          |
| $dp_t$     | 6.63106  | 6.56013  | 7.46083  | 6.23385   | 0.05642        | 0.797           |
| $de_t$     | 6.009458 | 5.97481  | 7.03574  | 5.53335   | 0.06795        | 0.790           |
| $TRM_t$    | 6.09590  | 5.80564  | 22.38690 | −10.66830 | 1.67778        | 0.848           |

The correlation matrix (Table 4) indicates that the relationship between excess returns on the stock market and  $scr_t$  is negative, and the relationship between excess returns on the stock market and  $cay_t$  positive, which is consistent with Hypotheses 1 and 2. The described state factors  $scr_t$  and  $cay_t$  are softly associated with several control variables; furthermore, we orthogonalize the control parameters to the state factors in the next section and examine their predicting power.

**Table 4.** Summary of Correlation Matrix Statistics.

|         | $R$      | $scr_t$ | $cay_t$ | $dp_t$  | $de_t$  | $TRM_t$ |
|---------|----------|---------|---------|---------|---------|---------|
| $r$     | 1        |         |         |         |         |         |
| $scr_t$ | −0.25758 | 1       |         |         |         |         |
| $cay_t$ | 0.50318  | 0.21855 | 1       |         |         |         |
| $dp_t$  | 0.28746  | 0.22822 | 0.12673 | 1       |         |         |
| $de_t$  | 0.29778  | 0.24236 | 0.10829 | 0.99497 | 1       |         |
| $TRM_t$ | 0.73501  | 0.65481 | 0.26817 | 0.38058 | 0.39779 | 1       |

Univariate summary statistics suggest that, on average, the change in  $scr_t$  is less than that in other variables, and the difference in  $cay_t$  is greater than in other variables. Regarding the autocorrelation of the parameters at the *first lag*, all other variables except  $cay_t$  display much more autocorrelation. The high stability of the independent factors may be useful for predicting excess market returns in the long term.

The stationarity of the expected return should be checked. For this purpose, we examined our null hypothesis (*the expected return is stationary*) by observing its trend in Figure 1 and the KPSS unit root test. LM-stat is lower than crucial values, and we cannot reject our null hypothesis. Therefore, it is stationary.

We employed the unit root test, which is augmented by Dickey-Fuller in the series on household consumption (M1), net household wealth (M2), and disposable income (M3). We examined every one of these three series, and there is one unit root with a maximum of 6 lags for the Schwarz Info Criterion (SIC) (See Supplementary Materials Tables S1–S3).

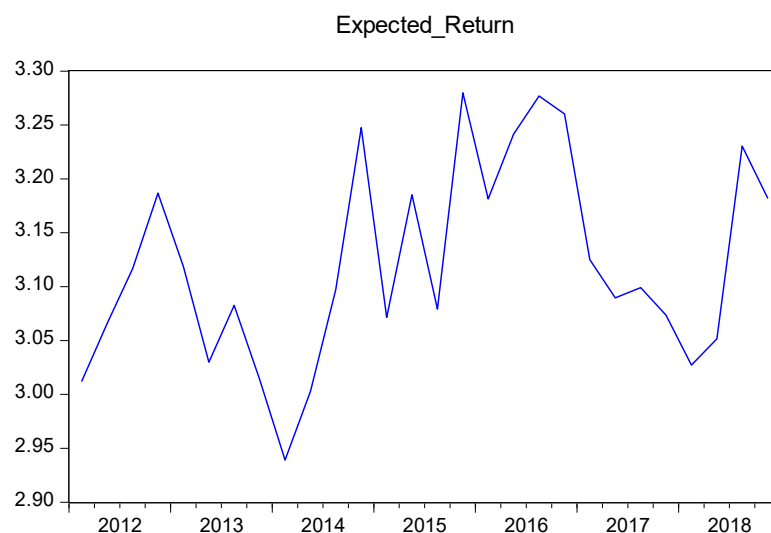


**Figure 1.** Expected return.

As these three variables (M1, M2, and M3) are non-stationary, this research examines the existence of cointegration correlation between the mentioned variables. So, we utilize different types of cointegration experiment: one of the tests is the residual-based cointegration test developed by Phillips and Ouliaris to determine whether the parameters under consideration are cointegrated, and the other is Johansen's maximum eigenvalue test and trace statistics experiment that shows how many cointegration vectors there are in a long-term relationship. Both experiments indicate that, between the three variables, the cointegration vector is unique. The outcome of two tests, the Phillips–Ouliaris cointegration test and the Johansen cointegration test, is presented in Tables S4 and S5. Cointegration experiments show that there is a common trend between M1, M2, and M3. If we want to use  $cay_t$  to test its predictive power regarding stock market returns, first we must calculate the factors of cointegration relevance. To examine endogenously specific nature of the  $c_t$ ,  $a_t$ , and  $y_t$  series, Stock and Watson present a single-equation procedure and use the DLS, which refers to the dynamic least squares presented by Newey and West. The dynamic least squares (DLS) are specified as

$$c_t = \alpha + \beta_a a_t + \beta_y y_t \quad (7)$$

The estimation obtained from Equation (4) generates an extremely consistent estimate of  $\beta$  and the  $\beta$  of the cointegration factors [17]. We used quarterly data from 2012Q1 to 2018Q4 to compute ratios of the trend deviation of the constant, net asset wealth ( $a_t$ ), and labor salaries ( $y_t$ ) from 'Eviews 9', which is shown in Equation (8).

$$c_t = 14.355 + 0.9718a_t - 3.9758y_t \quad (8)$$

(10.155) (12.783) (−8.828)

Relevant t-statistics are written under the ratio in the parenthesis. From the estimation of Equation (7), we can estimate that  $cay_t = c_t - \beta_a a_t - \beta_y y_t$ . Thus

$$cay_t = c_t - 0.9718a_t + 3.9758y_t \quad (9)$$

The expected excess return on asset for both state variables can be written such as:

$$E_t [r_{i,t+1}^e] = \alpha + \beta_1 scr_t + \beta_2 cay_t + MA_i \quad (10)$$

where  $\alpha$  is constant,  $\beta_i$  is constant slope coefficients and  $MA_i$  is moving average indicator modeling. With this in mind, for modeling, we used three different tests, including Akaike, Schwarz, and Hannan–Quinn. These methods have suggested that MA (2) is the best model for forecasting. Using MA (2), the moving average aligns with both theoretical and empirical research that shows how changes in consumption and wealth spread slowly across the financial system [40]. This specification of the moving average accounts for lagged responses in household spending, smoothing, and risk perception, enhancing fit compared to AR alternatives. Additionally, in this research, Nonlinear Two-Stage Least Squares (N-TSLS) was employed to mitigate any endogeneity between state variables and predicted returns [41]. This approach gives consistent estimates by modeling how the dependent variables of this research interact with predictive regressions. This is critical since both are forward-looking indicators that are affected by market expectations.

## 4. Results

### 4.1. Surplus Consumption Ratio

Our goal is measuring the predictability of  $scr_t$  and  $cay_t$  and control factors for unified stock market returns in the short and long term. We have three control variables, including  $dp_t$ ,  $de_t$ , and  $TRM_t$ .

Table 5 shows the OLS regression calculation outcomes with the single state variable  $scr_t$ /with control variables. The regression outcome is shown in columns (1) and (2), respectively.

**Table 5.** Return estimation using  $scr_t$ /with control variables.

| Variable           | (1)                        | (2)                                     |
|--------------------|----------------------------|-----------------------------------------|
| C                  | 2.991591 ***<br>(3.983746) | 2.443527<br>(1.166974)                  |
| $scr_t$            | −0.146233 *<br>(−0.120682) | −0.728282 **<br>(−0.988364)             |
| $de_t$             |                            | −0.162131<br>(−0.173376)                |
| $dp_t$             |                            | 0.175458<br>(0.157883)                  |
| $TRM_t$            |                            | 0.014413 ***<br>(3.432777)              |
| MA(1)              | 0.385215<br>(1.011465)     | −0.596789<br>(−3.58 × 10 <sup>5</sup> ) |
| MA(2)              | 0.230451<br>(0.417315)     | −0.403211<br>(−2.31 × 10 <sup>5</sup> ) |
| Observation        | 14                         | 14                                      |
| R-squared          | 0.120446                   | 0.725737                                |
| Adjusted R-squared | −0.270466                  | 0.405764                                |
| F-statistic        | 0.308115                   | 2.268120                                |

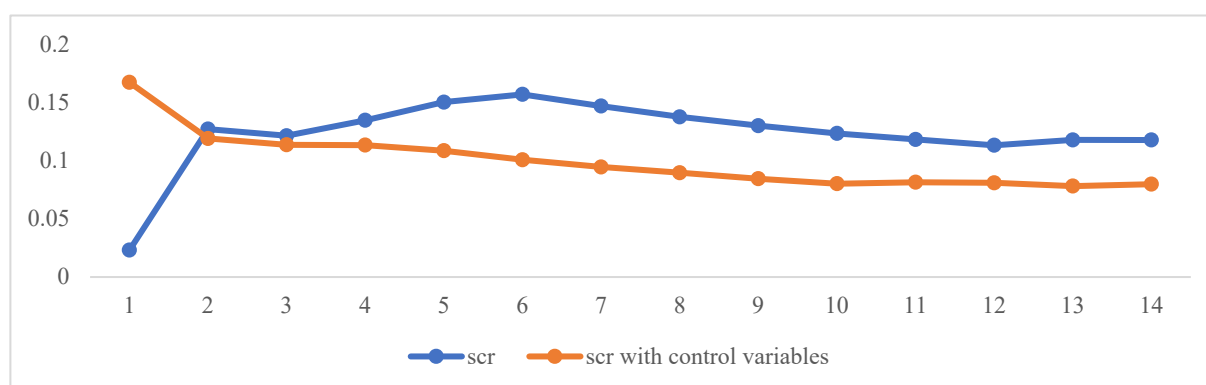
Note: the figures in parentheses are *t*-test values; \*\*\*, \*\*, and \* denote statistical significance at the 0.01, 0.05, and 0.10 level, respectively.

The coefficient of correlation of  $scr_t$  is negative, which shows that its relationship with the expected return is inverse, and Hypothesis 1 is established.

Then, we run *root mean squared error* (RMSE), applying a complete sample to the horizons between 1 and 14. The intended dependent factor is cumulative K-period excess market returns:  $\Sigma r_{t+1}$ . Table 6 indicates the outcomes of the regression using surplus consumption/*with control variables*. Each control factor is orthogonalized to the state variable by using the fitted residuals from the regression of the state factor. Table 6 also reports the Theil Inequality Coefficient (TIC). The lower the amount of RMSE and TIC, the better. This is also shown in Figure 2.

**Table 6.** Single State Variable Forecasting Regressions of  $scr_t$ /with control variables.

| Horizon | 1        |          | 2        |          |
|---------|----------|----------|----------|----------|
|         | RMSE     | TIC      | RMSE     | TIC      |
| Q1      | 0.023192 | -        | 0.167689 | -        |
| Q2      | 0.127344 | 0.020269 | 0.119303 | 0.018486 |
| Q3      | 0.121666 | 0.019395 | 0.113867 | 0.017633 |
| Q4      | 0.134936 | 0.021474 | 0.113568 | 0.017493 |
| Q5      | 0.150575 | 0.02391  | 0.108731 | 0.016671 |
| Q6      | 0.15735  | 0.02496  | 0.10098  | 0.0155   |
| Q7      | 0.14721  | 0.02341  | 0.09474  | 0.01461  |
| Q8      | 0.13788  | 0.021983 | 0.089768 | 0.013916 |
| Q9      | 0.130308 | 0.020812 | 0.084634 | 0.013176 |
| Q10     | 0.123625 | 0.019782 | 0.080349 | 0.01256  |
| Q11     | 0.118442 | 0.018995 | 0.0816   | 0.0128   |
| Q12     | 0.113465 | 0.018225 | 0.081116 | 0.012757 |
| Q13     | 0.11808  | 0.01895  | 0.07827  | 0.0123   |
| Q14     | 0.11796  | 0.01893  | 0.07992  | 0.01254  |

**Figure 2.** RMSE of return forecasting using  $scr$  and  $scr$  with control variables.

Our findings show that the RMSE of  $scr_t$  (Regression 1) at short horizons has the best performance with the lowest value (0.023192). It matches with the theory. After that, it increases from about 0.12 (2nd quarter) to 0.14 (7th quarter), such that the highest value is 0.15 (6th quarter). This shows does not exhibit proper performance in the mid-term horizon. Regarding the long-term horizon, decreased from 0.14 to 0.11, which shows its performance becomes better over the long-term horizon than mid-term. The TIC approach is similar to RMSE in this regression. In the second regression, including  $scr_t$  and control variables, RMSE is a continuously decreasing series (1st to 14th quarter), which shows this regressor is more robust with control variables. Its TIC approach is similar. Regressions (1) and (2) further validate Hypothesis 1.

#### 4.2. Consumption Wealth Ratio

In an identical manner to the previous section, Table 7 shows the OLS regression calculation outcomes with a single state variable:  $cay_t$ /with control variables.

The coefficient of correlation of  $cay_t$  is positive; that means its relationship with an expected return is direct, which verifies Hypothesis 2. The Heteroskedasticity Test of all estimations was conducted according to the Breusch–Pagan–Godfreyof method. The results show that the estimations are homoskedastic (see Tables S6 and S7).

Table 8 indicates the outcomes of the regression using the state variable/*with control variables*. This table also reports the Theil Inequality Coefficient (TIC). The lower the amount of RMSE and TIC, the better. This is also shown in Figure 3.

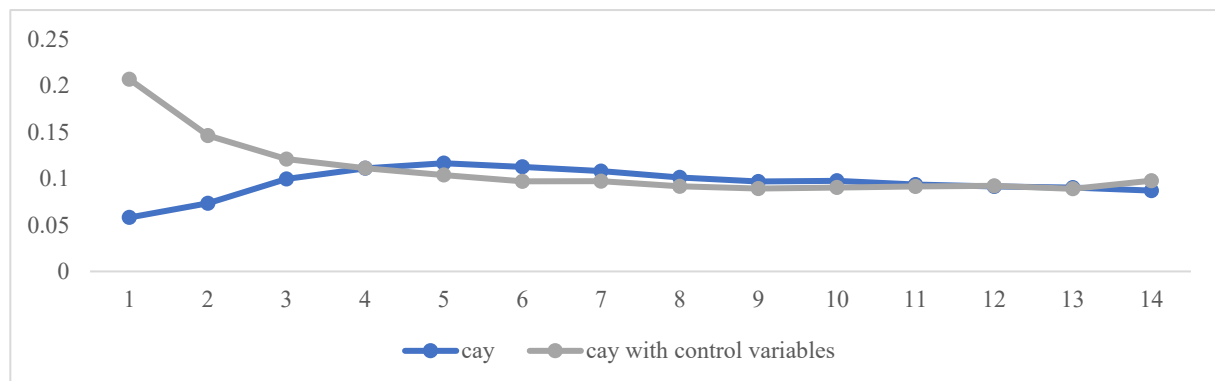
**Table 7.** Return estimation using *cay*/with control variables.

| Variable                | (3)                        | (4)                        |
|-------------------------|----------------------------|----------------------------|
| C                       | 2.435656 ***<br>(11.66133) | −0.079700<br>(−0.071793)   |
| <i>cay</i> <sub>t</sub> | 0.045176 **<br>(3.211853)  | 0.042841 ***<br>(4.550892) |
| <i>dp</i> <sub>t</sub>  |                            | 1.270126 *<br>(2.191582)   |
| <i>de</i> <sub>t</sub>  |                            | −0.977815 *<br>(−2.055565) |
| TRM <sub>t</sub>        |                            | 0.005716<br>(0.654912)     |
| MA (1)                  | 0.591084<br>(1.578985)     | 0.391322<br>(0.000243)     |
| MA (2)                  | 0.694545<br>(1.664657)     | 0.999991<br>(0.000121)     |
| Observations            | 14                         | 14                         |
| R-squared               | 0.600520                   | 0.878058                   |
| Adjusted R-squared      | 0.422973                   | 0.735793                   |
| F-statistic             | 3.382316                   | 6.171977                   |

Note: the figures in parentheses are *t*-test values; \*\*\*, \*\*, and \* denote statistical significance at the 0.01, 0.05, and 0.10 level, respectively.

**Table 8.** Single state variable forecasting regressions *cay*<sub>t</sub>/with control variables.

| Horizon | 3        |          | 4        |          |
|---------|----------|----------|----------|----------|
|         | RMSE     | TIC      | RMSE     | TIC      |
| Q1      | 0.058083 | -        | 0.206886 | -        |
| Q2      | 0.073398 | 0.011565 | 0.146329 | 0.022646 |
| Q3      | 0.099579 | 0.015789 | 0.12099  | 0.018851 |
| Q4      | 0.111052 | 0.017583 | 0.11122  | 0.017335 |
| Q5      | 0.11651  | 0.018387 | 0.103809 | 0.016143 |
| Q6      | 0.11261  | 0.01773  | 0.097    | 0.01507  |
| Q7      | 0.10799  | 0.01707  | 0.09718  | 0.01519  |
| Q8      | 0.101233 | 0.01604  | 0.091631 | 0.014375 |
| Q9      | 0.096831 | 0.01536  | 0.089194 | 0.014019 |
| Q10     | 0.09756  | 0.015488 | 0.090327 | 0.014223 |
| Q11     | 0.09351  | 0.01489  | 0.091458 | 0.01444  |
| Q12     | 0.091498 | 0.014591 | 0.092158 | 0.014576 |
| Q13     | 0.0903   | 0.01438  | 0.08895  | 0.01404  |
| Q14     | 0.08702  | 0.01384  | 0.09766  | 0.01538  |

**Figure 3.** RMSE of return forecasting using *cay* and *cay* with control variables.

In regression 3, the best performance occurred in the first step (0.058083). From the 2nd to 7th quarter, all of the RMSEs increase from 0.073398 to 0.107989. This increase in RMSE shows this regression is not so dominant in mid-term forecasting. For long-term forecasting, the 8th to the 14th quarters show that this regression is more suitable for forecasting than it is for the mid-term horizon. The TIC approach is similar to RMSE.

In the last regression,  $cay_t$  with control variables, its approach from the first to last quarter is decreasing; in comparison with  $cay_t$  (regressor 3), this shows that it is not more suitable forecasting. The approach of TIC in this series is similar to RMSE, which is in line with Hypothesis II.

#### 4.3. Two-State Variable

In this section, we examine a two-state variable forecasting regression in a similar manner to the preceding section. Table 9 shows the return estimation results using  $scr_t - cay_t$ /with control variables. Also, the *Heteroskedasticity Test* using the Breusch–Pagan–Godfreyof method and the relevant correlations was conducted so that they were satisfied (see Table S8).

**Table 9.** Return estimation using  $scr_t - cay_t$ /with control variables.

| Variable           | (5)                        | (6)                        |
|--------------------|----------------------------|----------------------------|
| C                  | 2.237905 ***<br>(3.298284) | −0.755299<br>(−0.606806)   |
| Scrt               | −0.329407 *<br>(−0.359154) | −0.740563 *<br>(−0.708127) |
| Cayt               | 0.044544 **<br>(2.789661)  | 0.036815 **<br>(3.026916)  |
| de <sub>t</sub>    |                            | −0.981177<br>(−1.620132)   |
| dp <sub>t</sub>    |                            | 1.317897<br>(1.844020)     |
| TRM <sub>t</sub>   |                            | 0.004287 *<br>(0.606492)   |
| MA(1)              | 0.826218<br>(0.026679)     | 0.813163<br>(0.000236)     |
| MA(2)              | 0.997404<br>(0.013390)     | 1.000000<br>(0.000118)     |
| Observation        | 14                         | 14                         |
| R-squared          | 0.677445                   | 0.897542                   |
| Adjusted R-squared | 0.475847                   | 0.733610                   |
| F-statistic        | 3.360388                   | 5.475089                   |

Note: the figures in parentheses are *t*-test values; \*\*\*, \*\*, and \* denote statistical significance at the 0.01, 0.05, and 0.10 level, respectively.

In this part, we analyze regressions, including two-state variables, using RMSE and TIC. Table 10 shows the RMSE and TIC using  $scr_t - cay_t$ /with control variables. This is also shown in Figure 4.

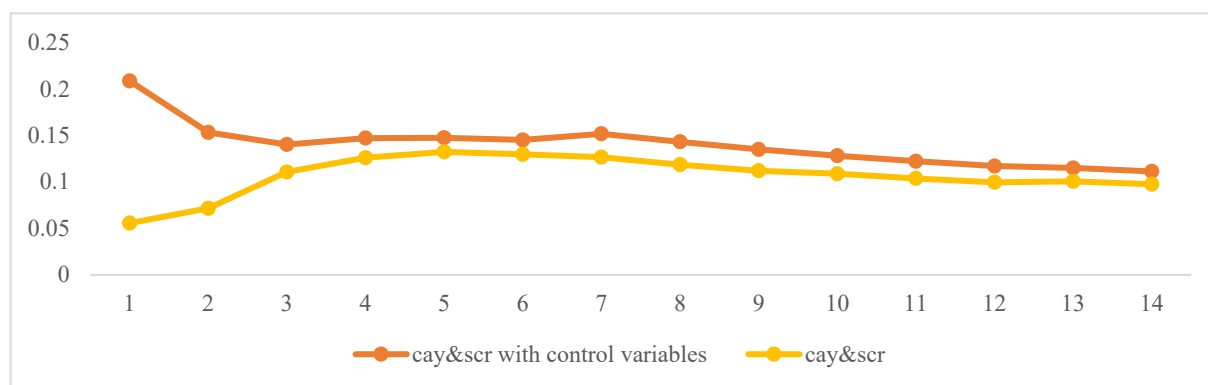
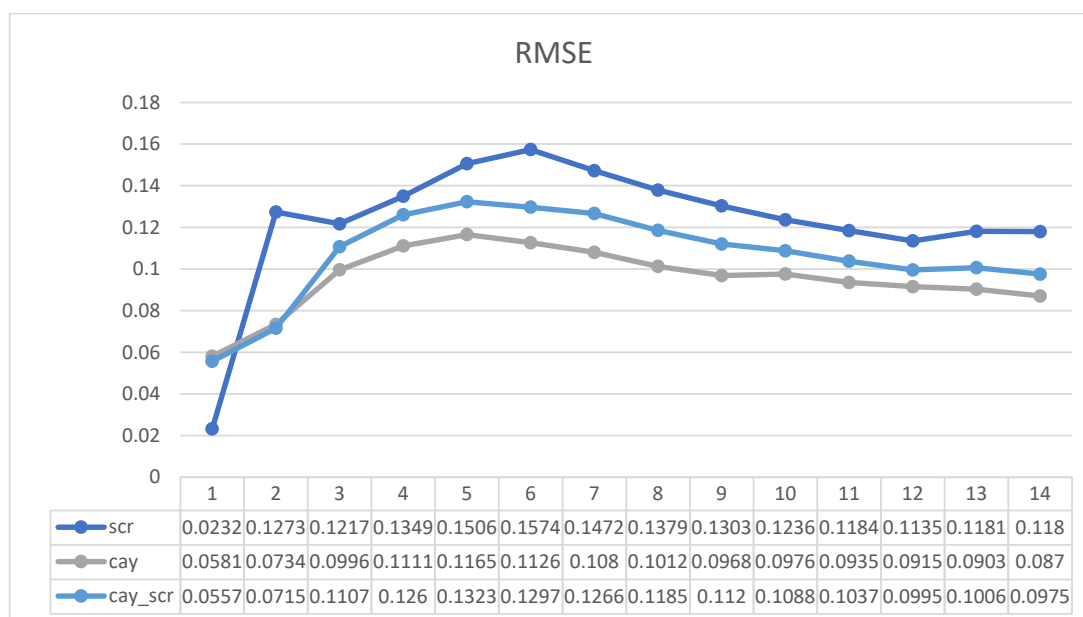


**Figure 4.** RMSE of return forecasting using two-state variables with control variables.

**Table 10.** Two-state variable forecasting regressions with control variables.

| Horizon | 5        |          | 6        |          |
|---------|----------|----------|----------|----------|
|         | RMSE     | TIC      | RMSE     | TIC      |
| Q1      | 0.05567  | -        | 0.208636 | -        |
| Q2      | 0.071538 | 0.011272 | 0.153257 | 0.023814 |
| Q3      | 0.110682 | 0.017573 | 0.140203 | 0.021989 |
| Q4      | 0.126013 | 0.01999  | 0.147173 | 0.023134 |
| Q5      | 0.132302 | 0.020924 | 0.147483 | 0.023151 |
| Q6      | 0.129688 | 0.020466 | 0.14514  | 0.022772 |
| Q7      | 0.12664  | 0.020072 | 0.151785 | 0.023977 |
| Q8      | 0.118537 | 0.01884  | 0.14314  | 0.022701 |
| Q9      | 0.112    | 0.017823 | 0.134958 | 0.021445 |
| Q10     | 0.108754 | 0.017323 | 0.128163 | 0.020406 |
| Q11     | 0.10372  | 0.016576 | 0.12225  | 0.019524 |
| Q12     | 0.099531 | 0.015934 | 0.117102 | 0.018744 |
| Q13     | 0.100626 | 0.016089 | 0.114983 | 0.018376 |
| Q14     | 0.097531 | 0.015582 | 0.111398 | 0.017772 |

Regression 5 in Table 10 indicates the outcomes of RMSE that indicate that  $scr_t - cay_t$  can forecast short-term horizon returns precisely. The RMSE increases from the second quarter to the sixth, indicating that it is not a good means of forecasting in the mid-term horizon. For the long-term horizon, from the 7th to the 14th quarters, it decreases (from 0.12664 to 0.097531), which indicates better performance than that for the mid-term horizon. At regression 6, including two-state variables with control variables, the orientation of the series decreases from the 1st to the 14th step. In the short term and mid-term, its values are high, which is not proper. In the long run, it has better performance. Its TIC approach is similar to its RMSE. Compared to previous regressions, regressions 5 and 6 cannot provide a more definite prediction. So, Hypothesis 3 has been rejected. By comparing the RMSE amounts of different regressions, as shown in Figure 5, over the short-term,  $scr_t$  is the best at forecasting. After the first quarter, to the long-term horizon, the best prediction is achieved by regression 3 ( $cay_t$ ). Figure 6 shows the RMSE trend of the regressions with control variables.  $scr_t$  with control variables is the best at forecasting in the short-term and long-term horizons. Of course, the difference compared to  $cay_t$  with control variables is small. In the mid-term background, from 4th to the 6th quarter,  $cay_t$  has better performance. Following this, a general figure summarizing all regression performances is provided in Figure 7.

**Figure 5.** RMSE of return forecasting using scr and/or cay.

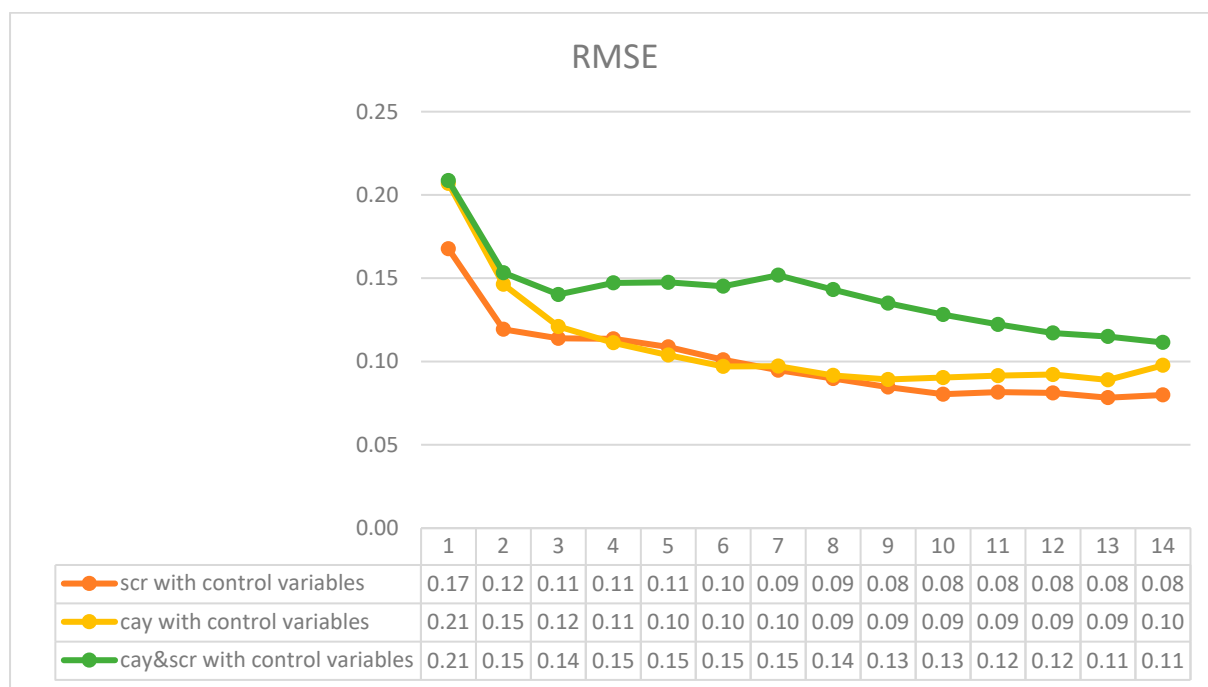


**Figure 6.** RMSE of return forecasting using scr and/or cay with control variables.

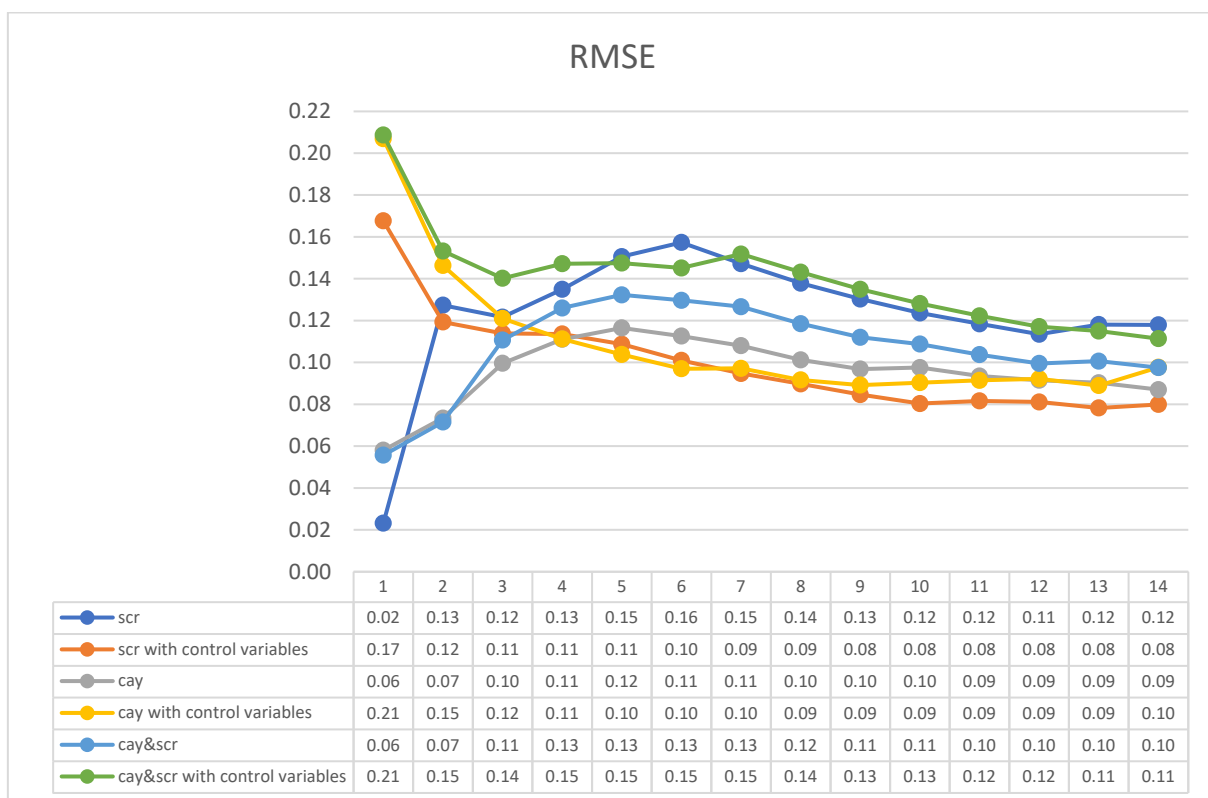


**Figure 7.** RMSE of return forecasting using all regressions.

#### 4.4. Further Diagnostic Test: The Jarque–Bera Statistics

A post-estimation diagnostic test was conducted to assess model adequacy. According to the  $p$ -values (as shown in Table 11), the Jarque–Bera normality test on residuals [42] fails to reject normality. So, this confirms that regression residuals are approximately normally distributed. The detailed components are reported in Tables S9–S11.

**Table 11.** The normality test (Jarque–Bera statistics).

| Variable(s) | Observations | R-Squared | Adjusted R-Squared | F-Statistic | Jarque–Bera Test (p-Value) | Output                 |
|-------------|--------------|-----------|--------------------|-------------|----------------------------|------------------------|
| cay         | 28           | 0.6078    | 0.5396             | 8.91        | 0.3237                     | Normality not rejected |
| scr         | 28           | 0.6079    | 0.5397             | 8.92        | 0.3043                     | Normality not rejected |
| scr-cay     | 28           | 0.6082    | 0.5192             | 6.83        | 0.3231                     | Normality not rejected |

#### 4.5. Regression Test Based on Nonlinear TSLS Estimation

We use the variance ratio to analyze the state variables' illustrative potential. In Equation (11), the regression variables on the right side are orthogonalized. As a result, we can compute the variance decomposition of the required excess returns as follows:

$$\text{Var}\left(E_t\left[\sum_{t=1}^k r_{i,t+k}^e\right]\right) = \text{Var}(a^k z_t) + \text{Var}(\beta_1^k \text{scr}_t) + \text{Var}(\beta_2^k \text{cay}_t) \quad (11)$$

Equation (11) declares that the share of the variation in expected K-period excess returns, that can be described by *scr* and *cay*, can be calculated, respectively, by the following variance ratios:

$$0 \leq VR_{scr}^k = \frac{\text{Var}(\beta_1^k \text{scr}_t)}{\text{Var}(a^k z_t) + \text{Var}(\beta_1^k \text{scr}_t) + \text{Var}(\beta_2^k \text{cay}_t)} \leq 1 \quad (12)$$

$$0 \leq VR_{cay}^k = \frac{\text{Var}(\beta_2^k \text{cay}_t)}{\text{Var}(a^k z_t) + \text{Var}(\beta_1^k \text{scr}_t) + \text{Var}(\beta_2^k \text{cay}_t)} \leq 1$$

We assessed the entire amount of predictable variance in excess returns that can be accounted for by *scr*<sub>*t*</sub> and *cay*<sub>*t*</sub> using Nonlinear Two Stage Least Squares (N-TSLS). We are unable to use the OLS approach as the variance ratio equation is nonlinear. In a nonlinear regression technique, the model with endogenous explanatory variables is examined using N-TSLS. A variable that has a correlation with the regression model's error term is called an endogenous variable. Using an endogenous variable goes against OLS's presumptions. The fundamental idea behind the TSLS technique is to forecast the model's elements using instrumental variables that are uncorrelated with the error term. The instrumental factors have no correlation with the model's error term, in contrast to the endogenous variables. The N-TSLS estimate findings for the explanatory power of *scr* and *cay* are shown in Table 12.

**Table 12.** Single state variable *scr* and *cay* forecasting regressions by N-TSLS.

| Horizon                                                                                                                    | 7        |          | 8        |                                        |
|----------------------------------------------------------------------------------------------------------------------------|----------|----------|----------|----------------------------------------|
|                                                                                                                            | RMSE     | TIC      | RMSE     | TIC                                    |
| Q1                                                                                                                         | 0.006022 | -        | 0.007581 | -                                      |
| Q2                                                                                                                         | 0.01996  | 0.024254 | 0.012914 | 0.022591                               |
| Q3                                                                                                                         | 0.017761 | 0.022696 | 0.013628 | 0.025247                               |
| Q4                                                                                                                         | 0.016289 | 0.021291 | 0.011897 | 0.02218                                |
| Q5                                                                                                                         | 0.014901 | 0.01956  | 0.010943 | 0.020166                               |
| Q6                                                                                                                         | 0.014335 | 0.01904  | 0.012388 | 0.022193                               |
| Q7                                                                                                                         | 0.014696 | 0.01989  | 0.01759  | 0.03198                                |
| Q8                                                                                                                         | 0.014789 | 0.020267 | 0.021422 | 0.038759                               |
| Q9                                                                                                                         | 0.014548 | 0.020049 | 0.022592 | 0.040363                               |
| Q10                                                                                                                        | 0.014553 | 0.020204 | 0.023935 | 0.041983                               |
| Q11                                                                                                                        | 0.015292 | 0.02153  | 0.025849 | 0.045739                               |
| Q12                                                                                                                        | 0.016008 | 0.022835 | 0.028846 | 0.050818                               |
| Q13                                                                                                                        | 0.015185 | 0.021689 | 0.02087  | 0.041536                               |
| Q14                                                                                                                        | 0.016723 | 0.024281 | 0.01439  | 0.032061                               |
| Moment Conditions                                                                                                          |          |          |          | Orthogonal to                          |
| $u_{1t}^k = \sum_{k=1}^K r_{t+k}^e - (a_0^k + a^k z_t + \beta_1^k \text{scr}_t + \beta_2^k \text{cay}_t)$                  |          |          |          | $1, z(-1), \text{scr}_t, \text{cay}_t$ |
| $u_{2t} = [(a^k z_t)^2 + (\beta_1^k \text{scr}_t)^2 + (\beta_2^k \text{cay}_t)^2] VR_{scr}^k - (\beta_1^k \text{scr}_t)^2$ |          |          |          | 1                                      |
| $u_{3t} = [(a^k z_t)^2 + (\beta_1^k \text{scr}_t)^2 + (\beta_2^k \text{cay}_t)^2] VR_{cay}^k - (\beta_2^k \text{cay}_t)^2$ |          |          |          | 1                                      |

The coefficient of  $scr$  ( $\beta_1$ ) is significantly negative. This means that its relationship with the expected return is indirect. The coefficient of  $cay$  ( $\beta_2$ ) is significantly positive. This implies that the relationship of  $cay_t$  with the expected return is direct. Regarding the RMSE, our findings show that the RMSE of  $scr_t$  (Regression 12) at short horizons has the best performance with the lowest value (0.006). This matches with the theory. After that, in 2nd quarter, it increases up to 0.019. And then it decreases from 0.017 (3th quarter) to 0.014 (5th quarter). After that, it is almost constant (about 0.014) from the 6th quarter to the 10th quarter. Next, it increases from 0.015 (11th quarter) up to 0.016 (12th quarter). And forwards it decreases (13th quarter). And finally, it increases again (14th quarter). The TIC approach is similar. Regression 12 further validates Hypothesis 1.

In  $cay_t$ , the best performance occurred in the first step (0.007). From the second to the third quarter, the RMSE increases from 0.012 to 0.013. And then from the forth quarter to fifth quarter the RMSE decreases from 0.011 to 0.010. Next it increases again from 0.012 (6th quarter) up to 0.028 (12th quarter). This increase in RMSE shows this regression is not so dominant in mid-term forecasting. Finally, it decreases from 0.020 (13th quarter) to 0.014 (14th quarter). This shows that forecasting using this regression is more suitable in the long-term than it is in the mid-term horizon. The TIC approach is similar to RMSE, which is in line with Hypothesis II. The next table (Table 13) shows the results of N-TSLS and OLS simultaneously.

**Table 13.** Comparison of RMSE in the proposed method and robustness method.

| Horizon | <i>cay</i> |        | <i>scr</i> |        |
|---------|------------|--------|------------|--------|
|         | N-TSLS     | OLS    | N-TSLS     | OLS    |
| Q1      | 0.0076     | 0.0581 | 0.0060     | 0.0232 |
| Q2      | 0.0129     | 0.0734 | 0.0200     | 0.1273 |
| Q3      | 0.0136     | 0.0996 | 0.0178     | 0.1217 |
| Q4      | 0.0119     | 0.1111 | 0.0163     | 0.1349 |
| Q5      | 0.0109     | 0.1165 | 0.0149     | 0.1506 |
| Q6      | 0.0124     | 0.1126 | 0.0143     | 0.1574 |
| Q7      | 0.0176     | 0.1080 | 0.0147     | 0.1472 |
| Q8      | 0.0214     | 0.1012 | 0.0148     | 0.1379 |
| Q9      | 0.0226     | 0.0968 | 0.0145     | 0.1303 |
| Q10     | 0.0239     | 0.0976 | 0.0146     | 0.1236 |
| Q11     | 0.0258     | 0.0935 | 0.0153     | 0.1184 |
| Q12     | 0.0288     | 0.0915 | 0.0160     | 0.1135 |
| Q13     | 0.0209     | 0.0903 | 0.0152     | 0.1181 |
| Q14     | 0.0144     | 0.0799 | 0.0167     | 0.0870 |

To analyze Table 12, we categorized all quarters into three main categories. The first is the short term. For N-TSLS regression, the RMSE of  $scr_t$  in the first quarter (short-term) is less than  $cay_t$  (0.006 vs. 0.0076). It is similar to OLS regression, such that the result of  $scr_t$  is more reliable than that of  $cay_t$  (0.023 vs. 0.0581). Secondly, in the mid-term, the trend of fluctuations of N-TSLS in comparison to OLS is slightly different, and N-TSLS is less reliable. Last but not least, in the long term,  $cay_t$  results have less error than  $scr_t$  in the long-term horizon (0.0144 vs. 0.0167). This is also similar to OLS regression (0.0799 vs. 0.087). In general, the results obtained from the robustness test verify the results of the research and show that they are significant.

## 5. Conclusions

The primary obstacle that relevant academics and investors face is estimating stock returns using asset price parameters. Thus, this research aims to tackle this difficulty by using actual stock data. Given the high degree of integration between the Chinese stock market and the world market, this research used China as a case study to investigate the predictive capacity of the consumption-based Capital Asset Pricing Model (CAPM) for stock returns. Campbell and Cochrane's habit formation model predicts time varying returns. According to this model,  $scr_t$ -based return predictability is implied. In addition,  $cay_t$  is used in the main stock market return expectation assessments. Logarithmic consumption (ct), logarithmic asset holding (at), and logarithmic labor income (yt) are all part of

the same technique that yields  $cay_t$ . Our research reveals a link between log consumption, log asset ownership, and log labor income via cointegration.

In this research, we apply the consumption-based CAPM via two-state variables such as  $scr_t$  and  $cay_t$ . These state variables are considered separately and together as different regressors. Using control variables (including dividend yield, dividend payout ratio, and government bond term spread) can improve the estimation results. Then, these state variables and control variables are formed into six regressions. We used the OLS procedure to test the forecast of stock market returns at the Shanghai Stock Exchange (SSE) with the above-mentioned regressors from 2012Q1 to 2018Q4. The coefficient correlation of  $scr_t$  is negative, which means its relationship with the expected returns is inverse, and Hypothesis 1 is established. The confidence levels of  $scr_t$  and  $cay_t$  with the control variables' prediction power are 90% and 95%, respectively. When the business cycle troughs, surplus consumption and investor risk aversion are high, while the expected or desired return is low. The coefficient correlation of  $cay_t$  is positive, which means its relationship with the expected return is direct, and Hypothesis 2 is established. The confidence levels in the prediction power of  $cay$  and  $cay$  with control variables are both 99%, respectively. When investors expect more moderate returns on an asset in the coming days, they will decrease consumption in the present, accordingly, to below the long-term horizon connection of consumption, assets, and wealth, to ensure further consumption. For state variables considering the control variable, we find that  $scr_t$  at short horizons has the best performance with the lowest RMSE value and predicts excess stock returns for extremely short timeframes. In spite of this,  $cay_t$  has better forecasting performance than  $scr_t$  for the mid-term and long-term horizons. The trend of estimation results using  $scr_t - cay_t$  is the same as  $cay_t$ , and, of course, its difference compared with  $cay_t$  is small. In the regressions, including state variables with control variables,  $scr_t$  with control variables is better at forecasting in the short-term horizon as well as the long-term horizon than  $cay_t$  with control variables and  $scr_t - cay_t$  with control variables. Generally, among all regressors,  $scr_t$  in the short term,  $cay_t$  in the medium term, and  $scr_t$  with control variables in the long term are the optimal regressors. These results show that combining  $scr_t$  and  $cay_t$  did not improve the prediction power. Therefore, we reject Hypothesis 3. Finally, we tested the robustness of the proposed model using N-2SLS. The results of the robustness check show that our model has significant power. Our research contributes new knowledge to the forecasting of changes to the aggregate stock market returns in the Chinese stock market, and can be considered by investigators and researchers. Future research may extend the time period of the used data to evaluate other policy regimes. Also, the Shenzhen Stock Exchange could be considered to conduct a further analysis of the Chinese stock market.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/systems13121066/s1>, Table S1: Household wealth, Table S2: Household disposable income, Table S3: Nondurable consumption, Table S4: Johansen, Table S5: Philips-Ouliaris Cointegration Test, Table S6:  $scr_t$ , Table S7:  $cay_t$ , Table S8:  $cay_t$  and  $scr_t$ , Table S9:  $scr$  normality test, Table S10:  $cay$  normality test, and Table S11: Double variables normality test.

**Author Contributions:** Conceptualization, F.Z.B.A.; Software, A.H.; Validation, A.H. and W.L.; Investigation, W.L.; Writing—original draft, F.Z.B.A.; Writing—review & editing, A.H. and W.L.; Supervision, W.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** Data is contained within the article or Supplementary Material.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Sun, G.; Li, Y. Intraday and Post-Market Investor Sentiment for Stock Price Prediction: A Deep Learning Framework with Explainability and Quantitative Trading Strategy. *Systems* **2025**, *13*, 390. [CrossRef]
2. Zhang, X.; Yang, X.; He, Q. Multi-Scale Systemic Risk and Spillover Networks of Commodity Markets in the Bullish and Bearish Regimes. *N. Am. J. Econ. Financ.* **2022**, *62*, 101766. [CrossRef]

3. Wang, Y.; Gu, W.; Ma, B.; Sun, G.; Kang, X. Does the opening of the China's stock market improve market liquidity? *Financ. Res. Lett.* **2025**, *85*, 108273. [\[CrossRef\]](#)
4. Liu, Y.; Zhang, L.; Failler, P.; Wang, Z. The Dynamic Evolution of Agricultural Trade Network Structures and Its Influencing Factors: Evidence from Global Soybean Trade. *Systems* **2025**, *13*, 279. [\[CrossRef\]](#)
5. Lyu, J.; Li, Z. Time-Varying Spillover Effects of Carbon Prices on China's Financial Risks. *Systems* **2024**, *12*, 534. [\[CrossRef\]](#)
6. Fehrle, D.; Heiberger, C. The return on everything and the business cycle in production economies. *Econ. Model.* **2024**, *136*, 106742. [\[CrossRef\]](#)
7. Kwan, Y.K.; Leung, C.K.Y.; Dong, J. Comparing Consumption-Based Asset Pricing Models: The Case of an Asian City. *J. Hous. Econ.* **2015**, *28*, 18–41. [\[CrossRef\]](#)
8. Polimenis, V. The historical lepto-variance of the US stock returns. *Data Sci. Financ. Econ.* **2024**, *4*, 270–284. [\[CrossRef\]](#)
9. French, J. A Practitioner's Guide to the CAPM: An Empirical Study. In *Global Tensions in Financial Markets*; Research in Finance; Emerald Publishing Limited: Leeds, UK, 2018; Volume 34, p. 1.
10. Vendrame, V.; Guermat, C.; Tucker, J. A Conditional Regime Switching CAPM. *Int. Rev. Financ. Anal.* **2018**, *56*, 1–11. [\[CrossRef\]](#)
11. Engsted, T.; Hyde, S.; Møller, S. V Habit Formation, Surplus Consumption and Return Predictability: International Evidence. *J. Int. Money Financ.* **2010**, *29*, 1237–1255. [\[CrossRef\]](#)
12. Boudoukh, J.; Michaely, R.; Richardson, M.; Roberts, M.R. On the Importance of Measuring Payout Yield: Implications for Empirical Asset Pricing. *J. Financ.* **2007**, *62*, 877–915. [\[CrossRef\]](#)
13. McMillan, D.G. Consumption and Stock Prices: Evidence from a Small International Panel. *J. Macroecon.* **2013**, *36*, 76–88. [\[Cross-Ref\]](#)
14. John, Y.C.; Shiller, R.J. Valuation Ratios and the Long-Run Stock Market Outlook. *J. Portf. Manag.* **1998**, *24*, 11–26. [\[CrossRef\]](#)
15. Xiao, Y.; Faff, R.; Gharghori, P.; Min, B.K. Pricing Innovations in Consumption Growth: A Re-Evaluation of the Recursive Utility Model. *J. Bank. Financ.* **2013**, *37*, 4465–4475. [\[CrossRef\]](#)
16. Liu, L.; Wang, Z. A Note on an Interpretation to Consumption-Based CAPM. *Econ. Lett.* **2008**, *99*, 443–445. [\[CrossRef\]](#)
17. Chang, T.; Gupta, R.; Majumdar, A.; Pierdzioch, C. Predicting Stock Market Movements with a Time-Varying Consumption-Aggregate Wealth Ratio. *Int. Rev. Econ. Financ.* **2019**, *59*, 458–467. [\[CrossRef\]](#)
18. Lettau, M.; Ludvigson, S. Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia Are Time-Varying. *J. Polit. Econ.* **2001**, *109*, 1238–1287. [\[CrossRef\]](#)
19. Balcilar, M.; Bekiros, S.; Gupta, R. The Role of News-Based Uncertainty Indices in Predicting Oil Markets: A Hybrid Nonparametric Quantile Causality Method. *Empir. Econ.* **2017**, *53*, 879–889. [\[CrossRef\]](#)
20. Fung, K.W.T.; Lau, C.K.M.; Chan, K.H. The Conditional Equity Premium, Cross-Sectional Returns and Stochastic Volatility. *Econ. Model.* **2014**, *38*, 316–327. [\[CrossRef\]](#)
21. Márquez, E.; Nieto, B.; Rubio, G. Stock Returns with Consumption and Illiquidity Risks. *Int. Rev. Econ. Financ.* **2014**, *29*, 57–74. [\[CrossRef\]](#)
22. Li, B.; Zhong, M. Predictability of Stock Returns and Consumption-Based CAPM: Evidence from a Small Open Market. *Eur. J. Econ. Financ. Adm. Sci.* **2010**, *22*, 148–173.
23. Farag, H.; Meng, Q.; Mallin, C. The Social, Environmental and Ethical Performance of Chinese Companies: Evidence from the Shanghai Stock Exchange. *Int. Rev. Financ. Anal.* **2015**, *42*, 53–63. [\[CrossRef\]](#)
24. Chen, C.; Hu, C.; Yao, H. Behavioral Framework of Asset Price Bubbles: Theoretical and Empirical Analyses. *Systems* **2022**, *10*, 251. [\[CrossRef\]](#)
25. Guo, B.; Zhang, W.; Zhang, Y.; Zhang, H. The Five-Factor Asset Pricing Model Tests for the Chinese Stock Market. *Pac.-Basin Financ. J.* **2017**, *43*, 84–106. [\[CrossRef\]](#)
26. Meng, Q.; Qi, J.; Wang, S.; Zhao, X. Disciplining the Factor Zoo: Identifying Pricing Factors in the Chinese Stock Market. *Econ. Model.* **2025**, *155*, 107370. [\[CrossRef\]](#)
27. Zhang, Y.; Jia, Q.; Chen, C. Risk Attitude, Financial Literacy and Household Consumption: Evidence from Stock Market Crash in China. *Econ. Model.* **2021**, *94*, 995–1006. [\[CrossRef\]](#)
28. Huang, T.-L. Is the Fama and French Five-Factor Model Robust in the Chinese Stock Market? *Asia Pac. Manag. Rev.* **2019**, *24*, 278–289. [\[CrossRef\]](#)
29. Chen, C.-D.; Demirer, R.; Jategaonkar, S.P. Risk and Return in the Chinese Stock Market: Does Equity Return Dispersion Proxy Risk? *Pac.-Basin Financ. J.* **2015**, *33*, 23–37. [\[CrossRef\]](#)
30. Le, A.T.; Nguyen, H.; Nguyen, C. Regret to reward: Investor regret and the cross-sectional stock returns in the Chinese market. *Glob. Financ. J.* **2025**, *68*, 101205. [\[CrossRef\]](#)
31. Kang, J.; Kim, T.S.; Lee, C.; Min, B.-K. Macroeconomic Risk and the Cross-Section of Stock Returns. *J. Bank. Financ.* **2011**, *35*, 3158–3173. [\[CrossRef\]](#)
32. Itoy, M.; Noda, A. *CCAPM with Time-Varying Parameters: Some Evidence from Japan*; Keio Economic Society, Keio University: Tokyo, Japan, 2011.

33. Liu, L.; Wan, J. The Relationships between Shanghai Stock Market and CNY/USD Exchange Rate: New Evidence Based on Cross-Correlation Analysis, Structural Cointegration and Nonlinear Causality Test. *Phys. A Stat. Mech. Its Appl.* **2012**, *391*, 6051–6059. [[CrossRef](#)]
34. Koutmos, D.; Song, W. Speculative Dynamics and Price Behavior in the Shanghai Stock Exchange. *Res. Int. Bus. Financ.* **2014**, *31*, 74–86. [[CrossRef](#)]
35. Auer, B.R. Can Habit Formation under Complete Market Integration Explain the Cross-Section of International Equity Risk Premia? *Rev. Financ. Econ.* **2013**, *22*, 61–67. [[CrossRef](#)]
36. Li, Y. The Wealth-Consumption Ratio and the Consumption-Habit Ratio. *J. Bus. Econ. Stat.* **2014**, *23*, 226–241. [[CrossRef](#)]
37. Møller, S.V. Habit Persistence: Explaining Cross-Sectional Variation in Returns and Time-Varying Expected Returns. *J. Empir. Financ.* **2009**, *16*, 525–536. [[CrossRef](#)]
38. Bianchi, F.; Lettau, M.; Ludvigson, S.C. *Monetary Policy and Asset Valuation*; No. 22572; National Bureau of Economic Research: Cambridge, MA, USA, 2016.
39. Caporale, G.M.; Sousa, R.M. Consumption, Wealth, Stock and Housing Returns: Evidence from Emerging Markets. *Res. Int. Bus. Financ.* **2016**, *36*, 562–578. [[CrossRef](#)]
40. Gardberg, M.; Pozzi, L. Aggregate consumption and wealth in the long run: The impact of financial liberalization. *J. Appl. Econom.* **2022**, *37*, 161–186. [[CrossRef](#)]
41. Xie, H. Nonlinear and nonseparable structural functions in regression discontinuity designs with a continuous treatment. *J. Econom.* **2024**, *242*, 105784. [[CrossRef](#)]
42. Jarque, C.M.; Bera, A.K. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Econ. Lett.* **1980**, *6*, 255–259. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.