Journal of the Asia Pacific Economy ›
Volume 18, 2013 - Issue 4: Macro and micro dimensions of manufacturing performance in India

ARTICLES

Lunar calendar effect: evidence of the Chinese Farmer's Calendar on the equity markets in East Asia

Wei-han Liu

Pages 560-593 | Published online: 07 Jun 2013

Cite this article https://doi.org/10.1080/13547860.2013.803841

Full Article | Figures & data | References | Citations | Metrics
Reprints & Permissions | Read this article | Share

Abstract

This paper examines the statistical relation between the Chinese Farmer's Calendar (CFC) and public market information in 10 East Asian equity markets during 1995–2004. CATREG and CART, two data mining techniques, are employed and the implications of the outcomes are discussed. The outcomes confirm that the CFC plays a supplementary role to market information in predicting the market rate of return. In addition to confirmation of lunar calendar effect by the CFC, CATREG outperforms in three markets: Taiwan, South Korea and Singapore. According to CART analyses, all the three markets value the funeral category of the CFC advice and this pattern coincides with the traditional wisdom of astrological knowledge. The latter two markets share the same set of CFC items while Taiwan has its own. This contrast indicates the market-dependent observance and different interpretation of the CFC items for determining an

auspicious date for equity trading. The lunar calendar effect in the three equity markets is confirmed.

Keywords:

lunar calendar effect | Chinese Farmer's Calendar | CATREG | Classification Tree Analysis | CART

Notes

1. Confirming empirical analyses include: (1) Floros (2008), (2) Dowling and Lucey (2008), and (3) Kang et al. (2010), etc.

2. The other major festivals in Chinese lunar calendar exercising significant effect include: Dragon Boat Festival (Tomb Festival, in summer season) and Moon Festival (mid-autumn festival, in autumn season).

3. The 24 solar terms are: spring begins (□□), the rains (□□), insects awaken (□□), vernal equinox (□□), clear and bright (□□), grain rain (□□), summer begins (□□), grain buds (□□), grain in ear (□□), summer solstice (□□), slight heat (□□), great heat (□□), autumn begins (□□), stopping the heat (□□), white dews (□□), autumn equinox (□□), cold dews (□□), hoar-frost falls (□□), winter begins (□□), light snow (□□), heavy snow (□□), winter solstice (□□), slight cold (□□), great cold (□□).

4. In imperial days, the Chinese type was certified, decreed and published exclusively by the central government, initially for farming purpose. This lunar calendar is also called the Yellow Calendar (□□) or Imperial Calendar (□□) because yellow is esteemed as the royal color, representing the emperor. CFC is also called the Old Calendar (□□), while the solar system or the Gregorian calendar adopted since the early twentieth century in China is called the New Calendar (□□).

5. Chinese calendar, http://en.wikipedia.org/wiki/Chinese_calendar

6. Among the literatures on calendar effects, some studies have acknowledged and analyzed its different impacts across different cultures. For example, American market participants do not like the number 13, even Friday the 13th(Kolb and Rodriguez 1988; Chamberlain, Cheung, and Kwan 1991). Chinese avoid the number 4 (Brown, Chua, and

Mitchell 2002). and this dislike is also revealed in many aspects, e.g. plate number selection Woo et al. (2008). However, no study so far has thoroughly analyzed the effects of days of lunar calendar on the large markets with participants who observe lunar calendar.

7. Interestingly enough, the logic to decide the daily advice is based on the calculation of the power and number of certain luck stars to appear on a specific date. Yet this mystical or esoteric know-how is only shared only among prestigious astrologists, not in open media. The printed material covering the CFC is only available in traditional Chinese, to the best of my knowledge and search. The book source covers only the outcomes at the introductory level of interpretation for practical purpose, nothing of the theory or calculation details.

8. In their worldwide survey of 76 countries, the people in the East Asian markets show higher degree of power distance index (PDI) than those in the western societies. The higher PDI indicates a pattern of polarization between dependence and counter-dependence of power in the region. PDI is defined as the extent to which the less powerful members of institutions and organizations within a country expect and accept that power is distributed unequally.

9. Nominal and ordinal variables have some shortfalls that make them often difficult to implement and interpret definitively. Many statistical procedures cannot be undertaken accordingly. The drawbacks of nominal and ordinal variables could be circumvented by converting nominal and ordinal variables to interval (numeric) variables. To some degree, optimal scaling can fulfill this purpose and give flexible transformation of the nonlinear relations. That is, optimal scaling derives interval measures from nominal and ordinal measures.

10. Variable with nominal measurement level only have the grouping property, that is, the category values only serve to code the observations into classes. Ordinal variables have the properties of grouping and ordering. Interval (numeric) variables have all the three properties: grouping, ordering and equal relative spacing.

11. The iteration refering to the transformation of the response is estimated from the linear combination of the transformed predictors from the previous iteration. The whole estimation process uses the backfitting procedure (Buja, Hasti, and Tibshirani 1989) in which the transformations are estimated from the partial residual when the response is predicted from all predictors, except the predictor for which the transformation is being

estimated. Specifically, the CATREG algorithm consists of two steps. In the first step, keeping the quantifications of the predictor variables and the regression coefficients fixed, the quantification of the response variable is estimated. In the second step, holding fixed the quantification of the response variable, the quantifications of the predictor variables and the regression coefficients are estimated for one variable at a time.

12. As one of methods for linear regularization of ill-posed problems, Ridge regression reduces coefficient estimates variability by shrinking the coefficients, resulting in more prediction accuracy at the cost of only a small increase of bias. Lasso improves prediction accuracy and model interpretability by combining the nice features of Ridge regression and subset selection. However, Lasso has certain limitation. For example, it tends to select one variable from a group and ignores the others. The number of selected variables is bounded by the sample size. Elastic net regularization further improves Lasso by encouraging grouping of highly correlated variables, while Lasso fails in grouped selection. Elastic net regularization removes the limitation on the number of selected variables.

13. The performance of a fitted regression model can be determined by its future observations. The optimality of the quantifications can be obtained from the observed data to predict future response and the usual criterion is the expected prediction error. For a linear regression, the apparent prediction error is the average loss of the observed data, which is minimized over the regression weights. The apparent prediction error is not an appropriate estimate for expected prediction error because the expected prediction error is estimated from the same data that were used for fitting the model and consequently leads to an optimistic estimate.

14. These resampling methods can serve as out of sample tests to avoid choosing an over-fitted model due to optimal scaling. However, those two resampling methods are handicapped with issues of bias and variability.

15. The traditional approaches include linear discriminant analysis, cluster analysis, logistic regression, nonparametric additive logistic regression, partial least squares classification or neural networks.

16. CART is noted for its recursive consideration of one feature at a time, instead of working on multiple features at a time as most other parametric and nonparametric methods do.

17. The CART algorithm is a recursive procedure; starting at the root node, and then at every internal node, it selects a single feature, and a threshold value to split the group of individuals at the node into two groups to be placed at two new daughter nodes. CART grows the largest tree possible, called a maximal tree that is the tree whose terminal nodes cannot be split any further. A node may not be split any further either because it contains only cases that belong to a single class, or because no reduction in total diversity can be obtained by further splitting. The recursive computational process can be categorized into four steps: (1) specifying the criteria for predictive accuracy; (2) selecting splits; (3) determining when to stop splitting; and (4) selecting the 'right-sized' tree.

18. It still remains an open question as to which criterion is most appropriate for a given data-set.

19. Both methods are for estimating generalization error based on resampling. Cross-validation is markedly superior for small data-sets.

20. The detailed information for these 10 capitalization-weighted indices are available in the following respective websites: Taiwan Stock Exchange Corporation, http://www.twse.com.tw/en/; Hang Seng Indexes Company, http://www.hsi.com.hk/HSI-Net/; Shanghai Stock Exchange, http://www.sse.com.cn/sseportal/en/home/home.shtml; Shenzhen Stock Exchange, http://www.szse.cn/main/en/; Singapore Exchange, http://www.sgx.com/wps/portal/marketplace/mp-en/home; Philippine Stock Exchange, http://www.pse.com.ph/; Korea Exchange, http://eng.krx.co.kr/; Tokyo Stock Exchange, http://www.tse.or.jp/english/; Indonesia Stock Exchange, http://www.idx.co.id/; Bursa Malaysia, http://www.klse.com.my/website/bm/.

21. Observing CFC is not a religious practice. There is no question to record the observance level in any public census or private survey in the East Asian markets, to the best of my knowledge.

22. http://product.datastream.com/extranet

23. Some of the 10 equity markets are in their preliminary stage during the data period. Not all the markets carry the same market information set. I choose the representative five items that are available for all the 10 markets.

24. There are some other possible combinations as the regression choices but they are not reported here. Their estimation outcomes are hardly significant in terms of adjusted

R squares and coefficient estimates.

25. In terms of data transformation, numeric data (market information and market rate of return) are transformed by grouping which I follow the default and assume normal distribution of seven categories. Normal distribution is more appropriate than the only other selection: uniform distribution, for grouping variables of market information and market rate of return. Market status variable is transformed by the grouping method with one equal interval because there are only two outcomes. Step functions are used for nominal and ordinal scaling because there are limited number of categories in the CFC items and market status.

26. Possibly due to the high number of bootstrapping, there is marginal difference in the CATREG estimation outcomes in terms of various scaling levels (numeric, ordinal and nominal) and the discretization method (grouping, ranking and multiplying) of the variables in this study.

27. The high adjusted R squares in Regression 9 can be due to the high collinearity between the variables of the market rate of return and market status.

28. The high adjusted R squares may contribute to the significant correlation between the variables of market rate of return and market status.

29. The Red Guards in the Great Proletarian Cultural Revolution campaigned to ‘Destroy the Four Olds’, intriguing to break the old thought, culture, custom and habit of the traditional Chinese society.

30. Classification table to show the misclassification cost is not available for numeric dependent variable.

31. Figures 3 through 5 show the concise tree structure diagrams without the detailed statistics so as to avoid the reader being overwhelmed, e.g. Chi-square statistics and the number of cases in each node. The full set of output can be made available upon request.