

Final Project Report

Mingyue Zhang

Wei-Han Huang

1 Introduction

This project studies a seasonality trading strategy introduced by Mark Kamstra, Lisa Kramer, and Maurice Levi [1], focusing on how the length of daylight may affect investors' portfolio selection and the market return. The paper suggests a strong connection between investors' mental health and their risk-taking tendencies, jointly, this change in decision making may influence the behavior of the whole stock market.

The study concentrates on a depressive disorder named seasonal affective disorder (SAD), caused by the shortening of daylight as we approach the winter solstice and leads to more risk-averse portfolio management. However, the seasons, as well as the length of daylights in the Northern Hemisphere are the complete opposite of those in the Southern Hemisphere. The authors thus propose a trading strategy to exploit this opportunity, by investing in several markets in different latitudes, in specific, long the market in Northern Hemisphere from September and switch to the South in March.

2 Methodology

Although it is difficult to prove a psychological theory, the paper provided both medical evidence and clinical studies to support the link between SAD and heightened risk aversion. The seasonal variation in the length of daytime can therefore translate into recurring variation in equity returns. Furthermore, the authors attempt to model the return based on an AR(2) process and other variables that can also affect the sunlight, such as cloud coverage precipitation, and capture the influence of daylight on human sentiment, risk tolerance, and hence stock markets.

2.1 SAD Measurement

Suppose H_t , as the time from sunset to sunrise at a particular location, define the measurement of SAD effect, SAD_t as follows:

$$SAD_t := \min(H_t - 12, 0)$$

Note that this definition is a little bit different from the one in the original article[1] where SAD_t would be negative only in the fall and winter period¹, but the difference is trivial.

The length of daylight, H_t , is calculated using the code in sun.py file.

Country	Index	City	Latitude
USA	/SP	New York	41°N
USA	^NYA	New York	41°N
USA	NDAQ	NEW YORK	41°N
Sweden	EWD	Stockholm	59°N
UK	EWU	London	51°N
Germany	EWG	Frankfurt	50°N
Canada	EWC	Toronto	43°N
Japan	^N225	Tokyo	36°N
New Zealand	^NZ50	Auckland	37°S
Australia	^AXJO	Sydney	34°S
South Africa	EZA	Johannesburg	26°S

2.2 Model

To estimate the influence of SAD effect, we are using the following model:

$$r_t = \beta_0 + \alpha_1 r_{t-1} + \alpha_2 r_{t-22} + \beta_{SAD} SAD_t + \beta_{Monday} D_t^{Monday} + \beta_{Tax} D_t^{Tax} + \beta_{Fall} D_t^{Fall} + \beta_{Precipitation} Precipitation_t + \beta_{Temperature} Temperature_t + \epsilon_t$$

r_t is the log-return of index at time t , $\beta_{Monday} D_t^{Monday}$ is a dummy variable which equals one when period t is the trading day following a weekend (usually a Monday) and equals zero otherwise, $\beta_{Tax} D_t^{Tax}$ is a dummy variable which equals one for a given country when period t is in the last trading day or first five trading days of the tax year² and equals zero otherwise, and $\beta_{Fall} D_t^{Fall}$ is a dummy variable which equals one for a given country when period t is in the fall³ and equals zero otherwise. The environmental factors, each measured in the city of the exchange, are millimeters of precipitation ($Precipitation_t$), and temperature in degrees Celsius ($Temperature_t$).

¹The fall and winter period is defined as September 21 to March 20 for the Northern Hemisphere and March 21 to September 20 for the Southern Hemisphere

²The tax year commences on January 1 in the United States, Canada, Germany, Japan, and Sweden. The tax year starts on April 6 in Britain, on July 1 in Australia, on March 1 in South Africa, and on April 1 in New Zealand.

³Fall is defined as September 21 to December 20 in the Northern Hemisphere and March 21 to June 20 in the Southern Hemisphere.

3 Data collection

3.1 Equity Data

For this project, we try to replicate the model and trading strategy with out-of-sample and, more importantly, recent data to test its performance in the current market.

We use the daily stock index return data of United States, Sweden, Britain, Germany, Canada, New Zealand, Japan, Australia, and South Africa. All indices are value-weighted and are obtained from Yahoo Finance through Python API. The tickers of the data are: /SP, EWC, EWD, EWG, EWU, EZA, NDAQ, ^ AXJO, ^ N225, ^ NYA, ^ NZ50.

One of the most intricate and tedious processes of the project is perhaps the acquisition of the data for all the cities' weather. We use the database provided by the National Centers for Environmental Information of the National Oceanic and Atmospheric Administration. Yet, each city often has multiple weather stations and each facility measures different variables, for instance, some may only measure the atmospheric pressure and humidity, some have temperatures but lack information on precipitation . Nevertheless, they all have one thing in common: they all have days with no data. The weather stations in Tokyo, as well as the two nearby cities Chiba and Kawasaki, have fairly complete historical data, instead of the whole year of 2005. Almost every city we search, instead of Toronto and Stockholm, is missing data for half a month in April 2015.

Consequently, we focused on searching for weather stations locating at airports, since they tend to have the most complete data. The missing dates are estimated with linear interpolation. We then combine the weather information with the market data acquired earlier using SQL and completed the data collection from Jan 1, 2006 to Oct 31, 2020.

4 Trading Strategy

The paper suggests that the average annual return in market in the northern hemisphere are often low in autumn, and drastically improve until reaching its peak after the winter solstice, the day with the shortest period of daylight and longest night of the year. However, market in the southern hemisphere does not follow this pattern.

They then compare the two trading strategies, both invest in Sweden and Australia from 1980 to 2000: 1. Placing 50 percent of the portfolio in Swedish index and another 50 in the Australia index for twenty will result in an average annual return of 13.2 percent. 2. Repeatedly placing 100 percent of the portfolio in Swedish index from March to September and moving it into Australia index from September to March will results in a 21.1 percent average annual return. We will call this the SAD strategy.

5 Results

5.1 Regression Results

We did the regressions for each indices and the results of SP 500 and ASX 200 are shown below and the others are shown in the appendix.

Unfortunately, unlike the regression results in the original article, we didn't find significant SAD effect in our results: the values of the β_{SAD} are rather small and if we did the t-test, we couldn't reject the null hypothesis that the coefficient is actually 0. The outcome of t-test suggests that after we have controlled the effects of temperature, season, rainfall, tax-loss and the date, SAD couldn't explain the performance of stock markets.

Moreover, if we look at the R-square of the linear model, we would find that only about 1% of the variations of the indices returns have been explained by this model, which indicates that we probably shouldn't use this model to explain the behaviors of markets. Actually, although the authors of [1] indeed found significant positive SAD effects, their models also have low R-square: 2% at most.

Dep. Variable:	/SP	R-squared:	0.005
Model:	OLS	Adj. R-squared:	0.002
Method:	Least Squares	F-statistic:	1.934
Date:	Wed, 23 Dec 2020	Prob (F-statistic):	0.0509
Time:	01:25:35	Log-Likelihood:	7319.1
No. Observations:	3204	AIC:	-1.462e+04
Df Residuals:	3195	BIC:	-1.457e+04
Df Model:	8		

	coef	std err	t	P> t	[0.025	0.975]
const	-0.0042	0.002	-1.734	0.083	-0.009	0.001
fall_north	0.0002	0.001	0.173	0.863	-0.002	0.003
Monday	-0.0020	0.001	-1.627	0.104	-0.004	0.000
PRCP_USA	0.0013	0.001	1.124	0.261	-0.001	0.003
TAVG_USA	6.78e-05	3.6e-05	1.881	0.060	-2.86e-06	0.000
SAD_Japan	-0.0012	0.001	-1.351	0.177	-0.003	0.001
tax_USA	-0.0053	0.003	-1.760	0.079	-0.011	0.001
return_lag1	-0.0367	0.018	-2.076	0.038	-0.071	-0.002
return_lag2	0.0015	0.018	0.083	0.934	-0.033	0.036

Omnibus:	1018.046	Durbin-Watson:	2.001
Prob(Omnibus):	0.000	Jarque-Bera (JB):	25179.838
Skew:	-0.944	Prob(JB):	0.00
Kurtosis:	16.603	Cond. No.	2.47e+03

Dep. Variable:	$\wedge AXJO$	R-squared:	0.009
Model:	OLS	Adj. R-squared:	0.007
Method:	Least Squares	F-statistic:	3.783
Date:	Wed, 23 Dec 2020	Prob (F-statistic):	0.000199
Time:	01:27:55	Log-Likelihood:	9566.8
No. Observations:	3204	AIC:	-1.912e+04
Df Residuals:	3195	BIC:	-1.906e+04
Df Model:	8		

	coef	std err	t	P> t 	[0.025	0.975]
const	-0.0026	0.001	-2.350	0.019	-0.005	-0.000
fall_south	0.0003	0.001	0.600	0.549	-0.001	0.001
Monday	-0.0006	0.001	-1.070	0.285	-0.002	0.001
PRCP_Australia	1.42e-05	2.53e-05	0.562	0.574	-3.53e-05	6.37e-05
TAVG_Australia	0.0002	6.36e-05	2.420	0.016	2.92e-05	0.000
SAD_Japan	0.0004	0.000	1.124	0.261	-0.000	0.001
tax_Australia	0.0013	0.001	0.939	0.348	-0.001	0.004
return_lag1	-0.0787	0.018	-4.450	0.000	-0.113	-0.044
return_lag2	0.0241	0.018	1.366	0.172	-0.011	0.059

Omnibus:	704.562	Durbin-Watson:	1.997
Prob(Omnibus):	0.000	Jarque-Bera (JB):	8661.777
Skew:	-0.687	Prob(JB):	0.00
Kurtosis:	10.937	Cond. No.	1.64e+03

5.2 Strategy Results

For most of the market we tested, the average monthly return through out the years does result in a seasonal change as the article suggested, experiencing its lowest point in fall and peaks in early January.

However, unlike the authors claimed, the markets in the southern hemisphere also experience the same cycle and we could easily reach this conclusion from Figure 2.

The SAD strategy, however, does work for the Sweden verses Australia example. However, only a tiny improvement from 2.2 to 2.3 percent average annual return. Moreover, if we switch Sweden with other market in the North, or switch Australia with other market in the South, the trading strategy does not necessarily work. We even try a new method call Reverse SAD, which we will use the exact opposite strategy, longing markets in the North from March to September and markets in the South from September to March. Yet, some combination resulted with a way higher return using the reverse SAD and the results are shown on Figure 2.

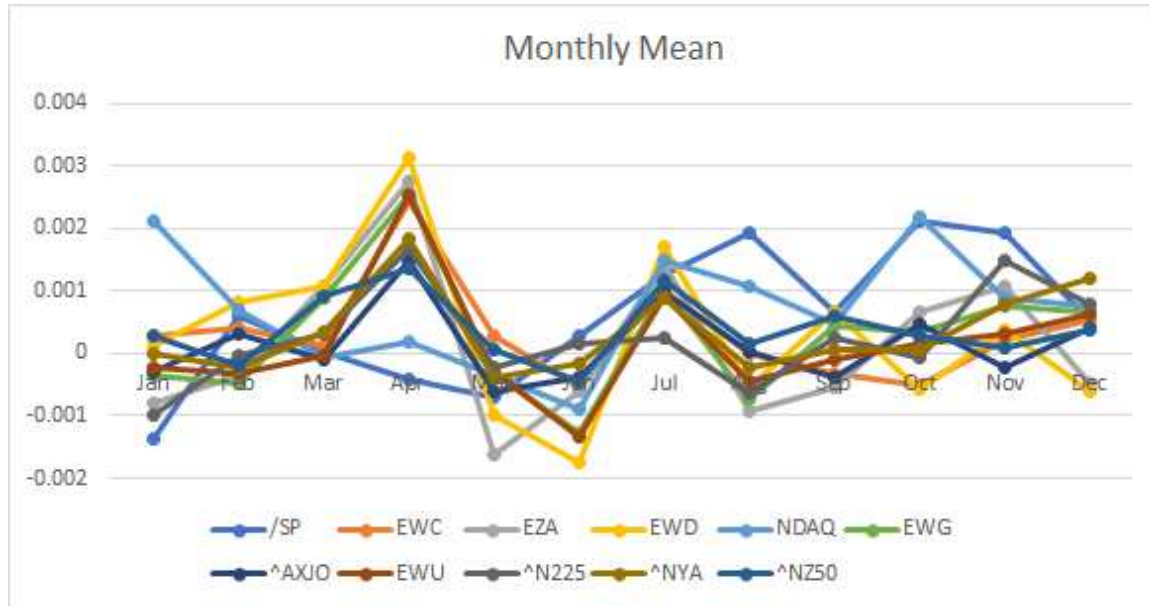


FIGURE 1: Monthly Mean for All the Markets

		Sweden EWD	Germany EWG	UK EWU	USA /SP	USA NDAQ	USA ^NYA	Canada EWC	Japan ^N225	
Australia	^AXJO	2.2%	1.6%	-0.5%	3.8%	8.4%	2.7%	1.3%	2.1%	50/50
		2.3%	1.4%	0.4%	1.5%	-0.7%	2.8%	5.2%	1.0%	SAD Strategy
		2.1%	1.7%	-1.8%	6.1%	23.4%	2.4%	-1.3%	3.4%	reverse SAD
New Zealand	^NZ50	10.1%	9.5%	7.4%	11.7%	16.4%	10.6%	9.3%	10.1%	50/50
		7.5%	6.1%	4.5%	6.2%	2.8%	8.2%	11.9%	5.4%	SAD Strategy
		9.6%	9.0%	2.3%	17.0%	49.2%	10.3%	3.4%	12.2%	reverse SAD
South Africa	EZA	0.5%	-0.2%	-2.2%	2.0%	6.7%	0.9%	-0.4%	0.4%	50/50
		0.4%	-0.3%	-1.1%	-0.3%	-2.0%	0.7%	2.6%	-0.7%	SAD Strategy
		-0.2%	-0.5%	-3.1%	2.7%	15.5%	0.0%	-2.7%	0.8%	reverse SAD

FIGURE 2: Strategy Results

6 Conclusion and Discussion

As we mentioned before, using the recent data (from 2006 to 2020) we didn't find significant SAD effects in our regression results while large positive SAD effects have been detected in [1] when the model is applied on the data ranging from 1928 to 2000.

However, does this mean the SAD effects have disappeared? We don't think so. Indeed, we can observe significant SAD effects within the old period, but we should also notice that the R-squares of all the regression are below 2%. Although it seems that the authors have already added lots of explanatory variables to the model, they can only explain a subtle part of the data: maybe it is no SAD effect within the equity market at all! Then why the trading strategies are rather successful using SAD effects? Well, first of all, this trading strategy is very simple: we would hold one stock market index from Northern Hemisphere during spring and summer and then switch to another index from Southern Hemisphere when fall comes. It actually has a very little things to do with SAD effects. Over a very long period of time, like 1928 to 2000, every stock index has enormous increase, which means we can

always get outstanding outcome even though we just arbitrarily select one index and hold it for several decades. Thus, it is likely that the good results of the strategy applied on the old data aren't the outcomes of employing SAD effects. Instead, they are only the reflection of the growth of global economy.

There are also a few limitations of the model that is worth discussing. For instance, the model only uses the weather data of the location of the exchange, New York City for NASDAQ for example, as the variables. It is impossible for use such a small variable to determine the market of the whole market of the USA. Perhaps a country's average would be more reliable, yet, as we have demonstrated, a complete data like that would be extremely difficult to acquire. Moreover, the SAD trading strategy uses the fact that the market often drops the most in September and peak in January, as the daylight are getting longer, however, it would not explain the phenomenon where the market is slightly improving from September to December, when the daylight is only getting shorter and shorter. Perhaps the amount of sunlight isn't the real factor driving the market.

References

- [1] Mark J Kamstra, Lisa A Kramer, and Maurice D Levi. Winter blues: A sad stock market cycle. *American Economic Review*, 93(1):324–343, 2003.

A Regression Results

Dep. Variable:	EZA	R-squared:	0.011
Model:	OLS	Adj. R-squared:	0.009
Method:	Least Squares	F-statistic:	4.486
Date:	Wed, 23 Dec 2020	Prob (F-statistic):	1.97e-05
Time:	01:30:18	Log-Likelihood:	7356.3
No. Observations:	3204	AIC:	-1.469e+04
Df Residuals:	3195	BIC:	-1.464e+04
Df Model:	8		

	coef	std err	t	P> t 	[0.025	0.975]
const	0.0008	0.002	0.381	0.703	-0.003	0.005
fall_south	0.0004	0.001	0.339	0.734	-0.002	0.003
Monday	-0.0024	0.001	-1.979	0.048	-0.005	-2.22e-05
PRCP_South Africa	2.076e-05	7.17e-05	0.290	0.772	-0.000	0.000
TAVG_South Africa	-3.813e-05	0.000	-0.323	0.747	-0.000	0.000
SAD_Japan	-0.0001	0.001	-0.199	0.842	-0.001	0.001
tax_South Africa	-0.0027	0.003	-0.980	0.327	-0.008	0.003
return_lag1	-0.0980	0.018	-5.543	0.000	-0.133	-0.063
return_lag2	-0.0120	0.018	-0.678	0.498	-0.047	0.023

Omnibus:	635.749	Durbin-Watson:	1.998
Prob(Omnibus):	0.000	Jarque-Bera (JB):	6538.357
Skew:	-0.637	Prob(JB):	0.00
Kurtosis:	9.881	Cond. No.	800.

Dep. Variable:	\wedge NZ50	R-squared:	0.004
Model:	OLS	Adj. R-squared:	0.001
Method:	Least Squares	F-statistic:	1.529
Date:	Wed, 23 Dec 2020	Prob (F-statistic):	0.142
Time:	01:30:52	Log-Likelihood:	10843.
No. Observations:	3204	AIC:	-2.167e+04
Df Residuals:	3195	BIC:	-2.161e+04
Df Model:	8		

	coef	std err	t	P> t 	[0.025	0.975]
const	0.0003	0.001	0.373	0.709	-0.001	0.002
fall_south	9.799e-05	0.000	0.247	0.805	-0.001	0.001
Monday	-0.0004	0.000	-1.052	0.293	-0.001	0.000
PRCP_New Zealand	2.163e-05	2.23e-05	0.968	0.333	-2.22e-05	6.54e-05
TAVG_New Zealand	6.897e-06	5.38e-05	0.128	0.898	-9.85e-05	0.000
SAD_Japan	1.952e-05	0.000	0.081	0.935	-0.000	0.000
tax_New Zealand	0.0013	0.001	1.354	0.176	-0.001	0.003
return_lag1	0.0474	0.018	2.678	0.007	0.013	0.082
return_lag2	-0.0118	0.018	-0.664	0.507	-0.047	0.023

Omnibus:	642.873	Durbin-Watson:	1.999
Prob(Omnibus):	0.000	Jarque-Bera (JB):	21064.194
Skew:	0.025	Prob(JB):	0.00
Kurtosis:	15.561	Cond. No.	1.78e+03

Dep. Variable:	EWU	R-squared:	0.015
Model:	OLS	Adj. R-squared:	0.013
Method:	Least Squares	F-statistic:	6.231
Date:	Wed, 23 Dec 2020	Prob (F-statistic):	5.07e-08
Time:	01:34:45	Log-Likelihood:	8611.8
No. Observations:	3204	AIC:	-1.721e+04
Df Residuals:	3195	BIC:	-1.715e+04
Df Model:	8		

	coef	std err	t	P> t	[0.025	0.975]
const	0.0015	0.001	1.341	0.180	-0.001	0.004
fall_north	-0.0007	0.001	-0.919	0.358	-0.002	0.001
Monday	-0.0013	0.001	-1.667	0.096	-0.003	0.000
PRCP_UK	-4.465e-06	7.86e-05	-0.057	0.955	-0.000	0.000
TAVG_UK	-8.664e-05	7.08e-05	-1.224	0.221	-0.000	5.21e-05
SAD_Japan	0.0004	0.001	0.811	0.418	-0.001	0.001
tax_UK	0.0010	0.002	0.546	0.585	-0.003	0.005
return_lag1	-0.1073	0.018	-6.068	0.000	-0.142	-0.073
return_lag2	0.0342	0.018	1.936	0.053	-0.000	0.069

Omnibus:	927.722	Durbin-Watson:	1.997
Prob(Omnibus):	0.000	Jarque-Bera (JB):	13934.951
Skew:	-0.959	Prob(JB):	0.00
Kurtosis:	13.035	Cond. No.	853.

Dep. Variable:	EWD	R-squared:	0.018
Model:	OLS	Adj. R-squared:	0.016
Method:	Least Squares	F-statistic:	7.392
Date:	Wed, 23 Dec 2020	Prob (F-statistic):	8.56e-10
Time:	01:35:14	Log-Likelihood:	7840.2
No. Observations:	3204	AIC:	-1.566e+04
Df Residuals:	3195	BIC:	-1.561e+04
Df Model:	8		

	coef	std err	t	P> t 	[0.025	0.975]
const	0.0016	0.001	1.689	0.091	-0.000	0.003
fall_north	-0.0019	0.001	-1.913	0.056	-0.004	4.84e-05
Monday	-0.0012	0.001	-1.210	0.227	-0.003	0.001
PRCP_Sweden	3.145e-05	0.000	0.295	0.768	-0.000	0.000
TAVG_Sweden	-8.842e-05	6.32e-05	-1.399	0.162	-0.000	3.55e-05
SAD_Japan	9.576e-05	0.001	0.131	0.895	-0.001	0.002
tax_Sweden	-0.0002	0.003	-0.072	0.943	-0.005	0.005
return_lag1	-0.1263	0.018	-7.142	0.000	-0.161	-0.092
return_lag2	0.0002	0.018	0.009	0.993	-0.035	0.035

Omnibus:	692.487	Durbin-Watson:	1.999
Prob(Omnibus):	0.000	Jarque-Bera (JB):	8777.846
Skew:	-0.658	Prob(JB):	0.00
Kurtosis:	11.001	Cond. No.	611.

Dep. Variable:	EWC	R-squared:	0.005
Model:	OLS	Adj. R-squared:	0.002
Method:	Least Squares	F-statistic:	1.863
Date:	Wed, 23 Dec 2020	Prob (F-statistic):	0.0615
Time:	01:35:48	Log-Likelihood:	8717.5
No. Observations:	3204	AIC:	-1.742e+04
Df Residuals:	3195	BIC:	-1.736e+04
Df Model:	8		

	coef	std err	t	P> t 	[0.025	0.975]
const	0.0008	0.001	1.123	0.261	-0.001	0.002
fall_north	-0.0015	0.001	-1.838	0.066	-0.003	9.71e-05
Monday	-0.0015	0.001	-1.868	0.062	-0.003	7.29e-05
PRCP_Canada	5.398e-05	4.57e-05	1.180	0.238	-3.57e-05	0.000
TAVG_Canada	-2.019e-05	3.78e-05	-0.534	0.593	-9.43e-05	5.39e-05
SAD_Japan	0.0001	0.001	0.224	0.822	-0.001	0.001
tax_Canada	0.0009	0.002	0.441	0.659	-0.003	0.005
return_lag1	-0.0198	0.018	-1.121	0.262	-0.054	0.015
return_lag2	0.0289	0.018	1.631	0.103	-0.006	0.064

Omnibus:	822.319	Durbin-Watson:	1.999
Prob(Omnibus):	0.000	Jarque-Bera (JB):	15345.311
Skew:	-0.735	Prob(JB):	0.00
Kurtosis:	13.620	Cond. No.	919.

Dep. Variable:	EWG	R-squared:	0.009
Model:	OLS	Adj. R-squared:	0.006
Method:	Least Squares	F-statistic:	3.488
Date:	Wed, 23 Dec 2020	Prob (F-statistic):	0.000512
Time:	01:36:11	Log-Likelihood:	8297.1
No. Observations:	3204	AIC:	-1.658e+04
Df Residuals:	3195	BIC:	-1.652e+04
Df Model:	8		

	coef	std err	t	P> t	[0.025	0.975]
const	-0.0002	0.001	-0.237	0.813	-0.002	0.002
fall_north	-0.0007	0.001	-0.779	0.436	-0.002	0.001
Monday	-0.0010	0.001	-1.085	0.278	-0.003	0.001
PRCP_Germany	-2.084e-05	8.29e-05	-0.251	0.801	-0.000	0.000
TAVG_Germany	3.754e-05	5.19e-05	0.724	0.469	-6.42e-05	0.000
SAD_Japan	-0.0004	0.001	-0.681	0.496	-0.002	0.001
tax_Germany	-0.0010	0.002	-0.446	0.656	-0.005	0.003
return_lag1	-0.0772	0.018	-4.366	0.000	-0.112	-0.043
return_lag2	0.0391	0.018	2.212	0.027	0.004	0.074

Omnibus:	660.209	Durbin-Watson:	1.998
Prob(Omnibus):	0.000	Jarque-Bera (JB):	11631.508
Skew:	-0.489	Prob(JB):	0.00
Kurtosis:	12.283	Cond. No.	781.

Dep. Variable:	^N225	R-squared:	0.004
Model:	OLS	Adj. R-squared:	0.001
Method:	Least Squares	F-statistic:	1.416
Date:	Wed, 23 Dec 2020	Prob (F-statistic):	0.184
Time:	01:36:32	Log-Likelihood:	8682.7
No. Observations:	3204	AIC:	-1.735e+04
Df Residuals:	3195	BIC:	-1.729e+04
Df Model:	8		

	coef	std err	t	P> t	[0.025	0.975]
const	0.0008	0.001	0.694	0.488	-0.001	0.003
fall_north	-9.033e-05	0.001	-0.111	0.912	-0.002	0.002
Monday	0.0002	0.001	0.228	0.820	-0.001	0.002
PRCP_Japan	2.817e-05	2.27e-05	1.243	0.214	-1.63e-05	7.26e-05
TAVG_Japan	-4.435e-05	5.15e-05	-0.861	0.389	-0.000	5.67e-05
SAD_Japan	7.449e-05	0.001	0.123	0.902	-0.001	0.001
tax_Japan	-0.0003	0.002	-0.171	0.864	-0.004	0.004
return_lag1	-0.0517	0.018	-2.920	0.004	-0.086	-0.017
return_lag2	0.0080	0.018	0.453	0.651	-0.027	0.043
<hr/>						
Omnibus:	853.099	Durbin-Watson:	1.999			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	15779.626			
Skew:	-0.785	Prob(JB):	0.00			
Kurtosis:	13.758	Cond. No.	1.26e+03			
<hr/>						