A seasonal behavioral stock buying pattern in the United States stock exchange.

Tariq Khan
Omaha, Nebraska, USA

Humans, like all life, are sensitive to their environment. Human and animal individuals can be triggered into impulsive and even violent activities from spikes in serum serotonin. Serum histamine level is directly proportional to environmental allergen levels producing dreaded seasonal allergic symptoms. But serum serotonin level, or the control of the level, is inversely proportional to serum histamine levels. Examples are presented where analysis over time of stock buying, as measured in the daily closing level of the United States stock exchange, shows a seasonal “J” shaped pattern that correlates with serotonin falls and spikes (histamine spikes and crashes) induced by the rise and fall of environmental allergens. A logical predictive strategy is thus presented where, excluding very large shocks like wars, market players could “buy low” an exchange-traded fund (ETF) i.e., indexed “stock,” every fall season (between August 25 to October 5) and then subsequently “sell high” (between October 17 and November 10) as increased serotonin in the aggregate population leads to the increase in impulsive and speculative (over-confident) stock buying before “normalizing” around late November.

“"We are all animals, my lady. Most are too afraid to see it!"  
-- Darkness -- Legend

"In all the known history of Mankind, advances have been made primarily in physical technology; in the capacity of handling the inanimate world about Man. Control of self and society has been left to chance or to the vague gropings of intuitive ethical systems based on inspiration and emotion… Psychohistory was the quintessence of sociology; it was the science of human behavior reduced to mathematical equations. The individual human being is unpredictable, but the reactions of human mobs, Seldon found, could be treated statistically.”  
-- Isaac Asimov -- Second Foundation

"We human beings are part of nature and therefore we are more likely to find out about our ‘inner’ nature, to understand ourselves, by looking outside ourselves, at our role and place as animals. In John Gray’s words, ‘A zoo is a better window from which to look out of the human world than a monastery.’ This is not paradoxical, and without some such realignment of approach, the modern incoherence will continue.”  
-- Peter Watson -- Ideas a history of thought and invention, from fire to Freud

“Wouldn’t economics make a lot more sense if it were based on how people actually behave, instead of how they should behave?”  
-- Dan Ariely -- Predictably Irrational: The Hidden Forces That Shape Our Decisions

"Pages and pages of data... efficiency functioning on multiple levels and in multiple dimensions... there it was all the time, staring you in the face. Buried within the message itself.”  
-- S.R. Hadden -- Contact

Serotonin: …serotonergic neurons play an important part in a variety of psychiatric conditions from anxiety disorders to schizophrenia as well as behavioral impulse-related disorders (violence, substance abuse, obsessive control, gambling, etc…)

Humans, like all life forms, are sensitive to their environment. So much so, in fact, that mentally ill or mentally unstable individuals can be triggered into impulsive and violent activities from spikes in serum serotonin (Cetin et. al, 2017) including spikes after seasonal drops in pollen (Fig. 1). But serotonin spikes also affect “normal” humans as well. Further note that serum histamine levels are directly proportional to
environmental allergen levels producing the obvious, and dreaded, seasonal allergic reaction symptoms. But
serum serotonin level, or the control of the level in the brain, is inversely proportional to serum histamine
levels (Hough, 1999, and Munari et. al, 2015, and Ryo et. al., 2006). Also note that male humans have 52%
more serotonin than females (Nishizawa et. al., 1997).

Analysis over time of stock market stock buying, as measured in the daily closing level of the United
States stock exchange (also the S&P 500 and Dow Jones Industrial Average indexes), shows a seasonal "J"
shaped pattern that correlates with serotonin falls and spikes (histamine spikes and crashes) induced by the
rise and fall of environmental allergens.

A logical predictive strategy is thus presented where, excluding very large shocks like wars or a
Presidential election scandal before November voting, or in our specific analysis presented here, the removal
of the 2008 financial crisis and the 2020 global pandemic impact from covid-19, market players could “buy
low” an exchange-traded fund (ETF) i.e., indexed “stock,” every fall season (between an approximate date
range of August 25 to September 5) and then subsequently “sell high” (between an approximate date range of
October 17 and November 10) as increased serotonin in the aggregate population leads to the increase in
impulsive and speculative (over-confident) stock buying before “normalizing” back to the normal regression
curve around the middle to end of November each year (Fig. 2-3).

A scientific motto, often demonstrated to be true, is that “biology drives psychology.” The advent of
near real-time tracking of allergen levels in given cities or zip codes has led to a possible predictive model
from the known human serum biochemistry of histamine and serotonin interactions versus observed year-
over-year acts of violence (mass shootings) from mentally unstable individuals or even stock market overly
optimistic buying akin to impulsive gambling. “Normal” individuals can be observed and analyzed tracking
simple impulsive behaviors from similar, albeit manageable, increases in blood serum serotonin levels
leading to excess confidence and risk taking from the biochemistry of the associated crash in airborne
allergens of pollen that thus lowers serum histamine levels.

Examples of the phenomena can be seen in both fall and spring allergy seasons but the fall season
with ragweed weed pollen is very consistent with the dates of start, peak, and end each year occurring within
a specific range while the spring season involves various trees that pollenate and the start, peak, and end can
vary by a month or two depending on the length of a year’s winter (bomb cyclones etc.…). The fall “J” shaped
event is also more obvious and significant in size, compared to spring example events (Fig. 4). The
observation that the fall pollen start date and stop dates ”moves” down the latitudes (while spring moves up
latitudes) still only leads to a date range that is at most two to three weeks in length and this can be seen in
examples. Fall and spring of 2020 and 2021 are shown (Fig. 5) with the original graph followed by the
diagramed graph for comparison. The same is done using the years 2004 and 2005 (Fig. 6) and fall of 2021
(Fig. 7) to show fall season buy and sell ranges and the “J” shaped drop with subsequent larger increase. Note
it is interesting to observe that the market almost never fully drops below the higher level reached at the top
of the “J.” Part of this is growth in human population, investors, and transactions but this could also help
explain the aggregate growth in the stock market period (S&P 500 year over year 9% growth rate); perhaps
human bias is too optimistic.

The histamine dropping, and thus serotonin spiking, and even post-histamine (post-pollen) serotonin
rebound or two to three week “serotonin hump” (Fig. 8) has its own impact including possible increases in
cardiac arrests (Fig. 9) and consumer spending. Impulsivity from this higher-than-normal serotonin includes
spending, not just on stocks, but also spending “in general” a la gifts - like those for the annual Christmas
holiday in the United States. Thus, those infamous Black Friday (major shopping event after the annual last
Thursday in November Thanksgiving holiday) 5 am store front mobs of shoppers suddenly make a lot more
sense. Shoppers imbued with high serotonin (driving aggression and anger) and with their unbeknownst
impulsive drive to want and to buy act accordingly regardless of context of free will. In reflection what a
convenient if not perfect time of year for a gift-buying holiday season to maximize consumer spending and
economic output.

Also, if our assumed correlation with pollen/histamine/serotonin/stock buying is accurate, then note
how exogenous factors like climate change (increases number of pollination weeks before frost), increased
use of antidepressant (selective serotonin uptake inhibitor SSRI) medications, and even gym related exercise,
stimulants, and steroids could all, in theory, have an effect on a population of investors or stock traders that,
in the end, are still making a decision or bet with serotonin influencing their judgment, to one degree or
another, whether any individual actor understands or admits this. Note too that this investment strategy
must involve an indexed stock as any individual stock is too correlated with events specific to a given
company or industry versus the aggregate effect noticed at scale with the entire market (population) and with indexed stocks.

The goal of this essay is only to propose this possible correlation and to promote additional considerations for behavioral economics if not environmental, sociological, bio-chemical, psychological economics or finance research. The data exists so that experienced analysts can do full regression and statistical analysis to carefully exclude minor shocks and to compare every annual season, as far back as the data allows, to confirm or refute the given hypothesis with formal analysis. Thus, the limitations of this proposal are known and understood in depth.

An example analysis is done using closing daily stock market data from 1915 to 2019. Years 2008 and 2020 are removed as outliers – they lead to excessively low closing values as exogenous events. As a model to proxy the hypothesis, the BUY LOW day is chosen to be the last day of August for each year and then SELL HIGH day is chosen to be the last day of October for each year. The difference between the last market day in October price minus the last day in August price, per the hypothesis, should be a positive value with statistical significance with, at the very least, a weak direct relationship. Removing 2008 and 2020 shocks, which is the observed result from a basic linear regression analysis using Microtrends.net data (Fig. 10-15).

**Figures**

**Fig. 1. United States mass shootings by date versus city, latitude, and fall pollen range (orange).**

![Image of mass shootings by date versus city, latitude, and fall pollen range](source: [link.springer.com/article/10.1007/s10453-019-09601-2])
Fig 2. BUY/SELL strategy timeline dates vs fall seasonal allergen/histamine/serotonin levels.

Source: http://pollen.utls.edu/ragweed.htm

Fig. 3. Example indexed fund Spyder ETF on Robinhood app.
Fig 4. Example Robinhood SPDR S&P 500 ETF indexed fund in 2021 showing “J” shaped movement.

8/20/2021  $100.00 purchased to monitor.
9/16/2021  $100.87 baseline
10/4/2021  $96.69 (low buy)
11/5/2021  $106.10 (high sell) 32 days later (25 trading days) = 9.7% growth rate.

Fig 5. Original and overlay of Fall 2020 to Spring 2021 of Dow Jones (DJIA) index vs pollen/serotonin.

Source: https://www.macrotrends.net/1358/dow-jones-industrial-average-last-10-years
Fig 6. Original and overlay of fall 2004 and 2005 S&P 500 vs serotonin vs BUY/SELL date ranges.
Fig 7. Robinhood SPDR indexed ETF in fall 2021 showing “J-shaped” annual pattern.

Fig 8. 3-month (3M) prior window from 12/2/2021 of SPDR ETF Robinhood app. Original and overlay of annual drop, spike, and 2-3 week serotonin hump of (violence, impulsivity, and spending in blue).
Fig. 9. Example of US cardiac arrests by months vs “serotonin spikes” after tri-annual pollen events (spring tree (juniper), summer grass, fall weed (ragweed)).

Source: https://www.ahajournals.org/doi/full/10.1161/01.CIR.100.15.1630
Fig. 10: Final regression using all years from 2015 to 2019 removing the years 2008 (global financial crisis exogenous hit) and 2020 (covid19 global pandemic) as both are considered large overwhelming outliers. Doing so then leads to a p-value that rejects null hypothesis.

### Regression ANOVA

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Square</th>
<th>Mean Square</th>
<th>F Statistic (df1, df2)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression (between y, and y)</td>
<td>1</td>
<td>732502.624</td>
<td>732502.624</td>
<td>7.9882 (1,102)</td>
<td>0.005666</td>
</tr>
<tr>
<td>Residual (between y, and ŷ)</td>
<td>102</td>
<td>9351189.911</td>
<td>915697.9403</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total (between y, and ŷ)</td>
<td>103</td>
<td>10065602.53</td>
<td>97919.345</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. **Y and X relationship**

   - R Square ($R^2$) equals 0.07263. It means that 7.3% of the variability of Y is explained by X.
   - Correlation (R) equals 0.2695. It means that there is a weak direct relationship between X and Y.

2. **Goodness of fit**

   - Overall regression: right-tailed, $F(1,102) = 7.9882$, p-value = 0.005666. Since p-value < α (0.05), we reject the $H_0$.
   - The linear regression model, $Y = b_0 + b_1X + e$, provides a better fit than the model without the independent variable resulting in $Y = b_0 + e$.

   - The slope (β): two-tailed, $t(102) = 2.8263$, p-value = 0.005666. For one predictor it is the same as the p-value for the overall model.
   - The Y-intercept (b): two-tailed, t(102) = -2.0003, p-value = 1.9959. Hence b is not significantly different from zero. It is still mostly recommended not to force b to be zero.

3. **Residual normality**

   - The linear regression model assumes normality for residual errors. Shapiro will p-value equals 2.598e-11. It is assumed that the data is not normally distributed but since the sample size is large, it should not adversely affect the regression model.

4. **Outliers**

   - Outliers may affect the regression line.

   - If the distribution of the residuals is normal, then the probability of detecting 6 valid outliers or more would be 1.

   - You should only remove outliers if you identify them as errors!

Fig. 11: Data using all years including 2008 and 2020 (1915 to 2020) (below).

Regression line equation
\[ \hat{Y} = -1.332.9407 + 0.6824X \]

Interpretation of the results

Regression ANOVA

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of Square</th>
<th>Mean Square</th>
<th>F Statistic (df, df)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>1</td>
<td>46216.0313</td>
<td>46216.0313</td>
<td>0.0136 (1, 198)</td>
<td>0.9194</td>
</tr>
<tr>
<td>Residual</td>
<td>197</td>
<td>1929393.012</td>
<td>1004.705</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>198</td>
<td>1975609.045</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. Y and X relationship
   R Square (R²) equals 0.832406. It means that 83.2% of the variability of Y is explained by X. The correlation (r) equals 0.91259. It means that there is a very strong direct relationship between X and Y.

2. Goodness of fit
   Overall regression. r² = 0.832406, p-value = 0.000154. Since p-value < 0.05, we accept the H₀.
   The linear regression model: \( Y = b_0 + b_1 X \). It doesn't provide a better fit than the model without the independent variable resulting in \( H_0 \) and \( b_1 \).
   The slope (b₁) is unknown. T(df) = 0.0089, p-value = 0.8884. For our predictor it is the same as the p-value for the overall model.
   The t-value (b₁) is known. T(df) = 0.0089, p-value = 0.9230. Hence it is not significantly different from zero; it is still must likely recommended not to force it to be zero.

3. Residual normality
   The linear regression model assumes normality for residual errors. Shapiro-Wilk p-value equals 0.14824, F value = 0.255. It is assumed that the data is not normally distributed. Since the sample size is large, it should not adversely affect the regression model.

4. Outliers
   Outliers may affect the regression line. If the distribution of the residuals is normal, then the probability of detecting a valid outlier or more would be 0.9995. You should only remove outliers if you identify them as such.

Residuals normality

Calculation
\[
\begin{align*}
    \hat{Y} &= b_0 + b_1 X \\
    b_0 &= \frac{\sum Y - b_1 \sum X}{n} \\
    b_1 &= \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sum (X - \bar{X})^2} \\
    r &= \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum (X - \bar{X})^2 \sum (Y - \bar{Y})^2}} \\
    p &= \Phi^2(\frac{r-0.7745}{0.8440})^{0.973} - 0.332.9407 \\
    f &= \frac{\sum (Y - \hat{Y})^2}{\sum (Y - \bar{Y})^2} \\
    R^2 &= \frac{f}{f + n-2} \\
    \end{align*}
\]

Source: https://www.statskingdom.com/linear-regression-calculator.html
Fig. 12: All years (1915-2019) removing 2008 and 2020 outlier exogenous shocks (below): Now we can reject the H0 null hypothesis.

Regression line equation
\[ \hat{Y} = -5417.8971 + 2.7803X \]

Source: https://www.statskingdom.com/linear-regression-calculator.html
Fig. 13: Regression Test.

Source: https://www.socscistatistics.com/tests/regression/default.aspx
Fig. 14: Excel results using same data where original data has 2008 and 2020 removed. Last day market day of August is used as BUY LOW indexed price and the last market day of October is used as the SELL HIGH index price. The October price minus the August price (indexed closing level) is the gain. This gain is analyzed year over year where the weak direct positive relationship is seen as expected.

Source: How to Add a Regression Line to a Scatterplot in Excel (statology.org)
Fig. 15: Original Data Set.

<table>
<thead>
<tr>
<th>Date</th>
<th>Open</th>
<th>High</th>
<th>Low</th>
<th>Close</th>
</tr>
</thead>
<tbody>
<tr>
<td>8/29/1958</td>
<td>508.83</td>
<td>528.54</td>
<td>505.22</td>
<td>513.39</td>
</tr>
<tr>
<td>8/31/1958</td>
<td>505.18</td>
<td>525.66</td>
<td>502.22</td>
<td>507.22</td>
</tr>
<tr>
<td>9/1/1958</td>
<td>507.32</td>
<td>526.22</td>
<td>503.22</td>
<td>506.22</td>
</tr>
<tr>
<td>9/3/1958</td>
<td>506.66</td>
<td>526.22</td>
<td>502.22</td>
<td>505.66</td>
</tr>
<tr>
<td>9/4/1958</td>
<td>506.66</td>
<td>526.22</td>
<td>502.22</td>
<td>505.66</td>
</tr>
<tr>
<td>9/5/1958</td>
<td>506.66</td>
<td>526.22</td>
<td>502.22</td>
<td>505.66</td>
</tr>
</tbody>
</table>

Source: [https://www.macrotrends.net/charts/stock-indexes](https://www.macrotrends.net/charts/stock-indexes) and [https://www.macrotrends.net/charts/stock-indexes](https://www.macrotrends.net/charts/stock-indexes)

MacroTrends Data Download
Dow Jones Industrial Average: Daily Closing Values
Disclaimer and Terms of Use: Historical data is provided ‘as is’ and solely for informational purposes, not for trading purposes or advice. MacroTrends LLC expressly disclaims the accuracy, adequacy, or completeness of any data and shall not be liable for any errors, omissions, or other defects in, delays or interruptions in such data, or for any actions taken in reliance thereon. Neither MacroTrends LLC nor any of our information providers will be liable for any damages relating to your use of the data provided.
References


