Analysis of 2020 S&P Equity Returns

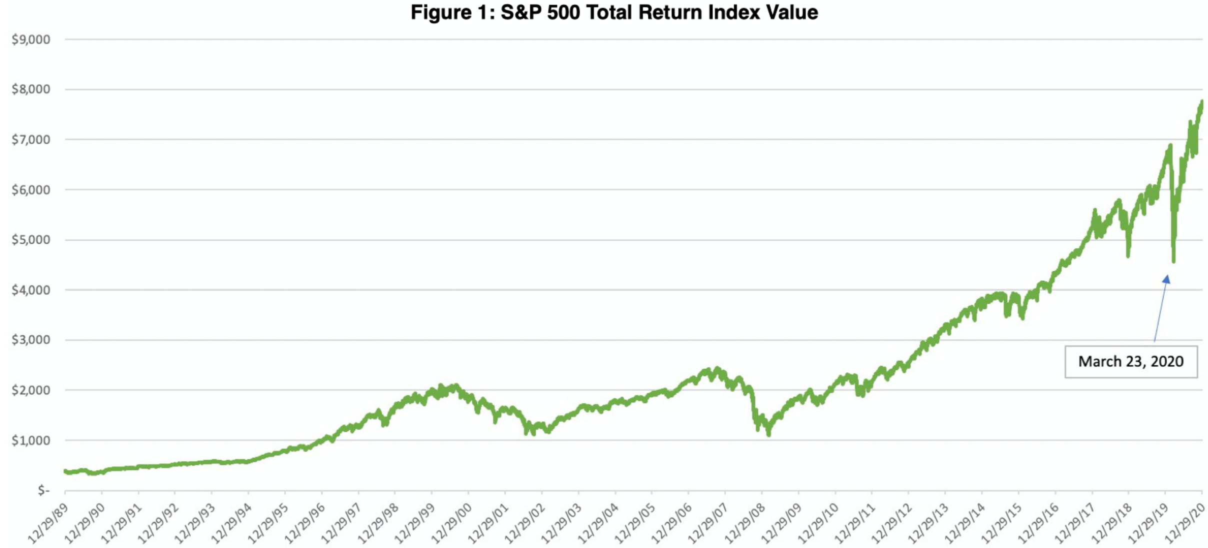
Author: Zak Fischer

Section 1: Overview

Most would agree 2020 was a year like no other. Between the COVID-19 pandemic, a heated U.S. political election, BREXIT, Tiger King, and more – 2020 was a unique year that many of us will always remember, for better or worse. While the world itself was quite strange in 2020, was stock market performance in 2020 also unusual? In this paper, we will explore U.S. equity market performance in 2020.

U.S. equity returns are analyzed in this paper by looking at daily returns in the S&P 500 Total Return Index (SP500TR). A graph of the stock price path over time is plotted below:

Figure 1: S&P 500 Total Return Index Value



During March 2020, COVID-19 concerns escalated in the US and lockdown measures were announced. Market returns in March 2020 were quite extreme and choppy, alternating between days of huge gains and losses.

March 16th was a particularly extreme market selloff. It was the largest point drop ever recorded for the S&P 500. Stocks dropped significantly after heightened concerns that the pandemic could have longer-lasting impacts than initially thought. On the other hand, March 13th and 24th are the two highest point increases ever recorded for the S&P 500. March was a time of heightened volatility with both sharp increases and decreases in equity markets, as seen in the table below:

Table 1: Analysis of Extreme Market Movements in March 2020

Day	Return
3/16/20	-11.98%
3/12/20	-9.49%
3/24/20	9.39%
3/13/20	9.32%

As investors prepare for the future, it is important to understand whether these volatility surges and sharp movements are understood and captured by commonly used models. For example, the popular Black-Scholes model assumes that stock price returns are normally distributed and stock price levels are lognormally distributed. Section 3.1 explores in more detail that directly using this distribution assumption does not adequately explain the stock market performance observed in 2020:

Table 2: Z-Scores of Extreme Market Movements in March 2020

Day	Return	Z-Score
3/16/20	-11.98%	-10.97
3/12/20	-9.49%	-8.70
3/24/20	9.39%	8.53
3/13/20	9.32%	8.46

The Z-scores shown in the table above are extremely large in magnitude. For a normal distribution, 68.27% of observations are within one standard deviation, 95.45% within two standard deviations, and 99.73% within three standard deviations.

The odds of a Z-score above 8-10 is extremely small and even hard to comprehend. For example, the odds of a Z-score above 10 is $13 \cdot 10^{-22}$.

In other words, the returns observed in March cannot be adequately explained by a normal distribution. And, there were *four* days in the month that cannot be explained by a normal distribution with a magnitude of Z-score above 8.

This paper will look to see if 2020 equity returns were different from prior years, and also look at which distributions besides a normal distribution may describe equity market performance more clearly. Besides the normal distribution, other distribution choices are explored, including: Student's t, Regime Switching Lognormal (RSLN-2), Extreme Value Theory (EVT) and Poisson Jump.

The models are judged on their ability to explain extreme returns by looking at the probability of experiencing equity returns at or more extreme than those observed in 2020. Additionally, the fit of the body of the distribution is tested in Section 2 using the Kolmogorov–Smirnov test.

The table below summarizes the results of this paper[1]:

Table 3: Results Summary

Model	Explains the Body of Returns	Explains Extreme Events
Normal	✗	✗
Student's t	✓	✓
RSLN-2	✓	∼
EVT	✗	✓
Poisson Jump	✗	✓

[1] ✓ indicates a strong fit, ∼ indicates a mediocre fit, and ✗ indicates a poor fit.

Below are a list of the key insights explored throughout this paper:

- The normal distribution does not have heavy enough of tails to explain 2020 equity return behavior. The four most extreme daily equity movements in 2020 have probability of less than 1-in-1 trillion, which is unreasonable and assigns insufficient weight to extreme events.
- 2020 equity returns had both negative skew and highly positive excess kurtosis (8.04), which again indicates the normal distribution does not have sufficiently heavy tails.
- The Student's t distribution is able to explain both the body of the distribution and extreme events. This paper uses a degrees of freedom parameter of 3, which gives a substantially fatter tail than the normal distribution. This allows for much more robust modeling of extreme events compared to the normal distribution. It is important to keep in mind the Student's t distribution is still not perfect and does have disadvantages, which are further explored in Section 3.7
- The RSLN-2 model explains the body of returns well, using two states: a low-volatility bull market and a high-volatility bear market. One of the unfortunate shortcomings of this model is that it is unable to explain large equity surges because the bull market is low-volatility. Therefore, RSLN-2 does not fully describe extreme market movements. This shortcoming is apparent from the extreme return analysis, but is not as important for the Kolmogorov–Smirnov test which assesses the fit of the entire distribution.
- The Extreme Value Theory (EVT) model explains extreme returns fairly well, but does not adequately fit the entire distribution of returns.
- The Poisson Jump is able to explain extreme jumps, but does not adequately describe the body of the distribution.
- The Kolmogorov–Smirnov test explores the fit for the entire distribution, while the extreme return probability analysis explores whether models can adequately explain sharp market movements. Importantly, a model may describe most returns well but fail in the extremes and still pass the Kolmogorov–Smirnov test. Therefore, both types of analysis are important to consider. For example, for insurance companies preparing contingency plans for large market movements, extreme returns are crucial to understand. In contrast, those focusing on pricing may have a more holistic focus on the entire distribution.
- Qualitative analysis should also be considered to get a broad perspective of risks. Using different distribution assumptions and parameterizations, as well as performing sensitivity tests on key metrics are also recommended.
- Parameters are calculated using 1990-2019 equity return data. Maximum Likelihood Estimation (MLE) is used for most parameters, with further details discussed in Appendix 2.

The remainder of this paper is structured in the following sections:

- **Section 2 (Summary Statistics and Nonparametric Tests)**. 2020 equity return performance is compared to prior years by looking at summary statistics and also by conducting the Kolmogorov–Smirnov test.
- **Section 3 (Model Comparisons)**. The normal distribution assumption is used as a starting point to model equity returns, and additional distributions are also analyzed.
- **Section 4 (Types of Randomness)**. Mild vs wild randomness (Mandelbrot, 86) is discussed as well as Black Swan events.
- **Section 5 (Conclusion)**. Key takeaways and recommendations going forward are summarized.
- **Appendices**. Parameterization methodology and detailed Kolmogorov–Smirnov test results are documented in the appendices.

Section 2: Summary Statistics and Nonparametric Tests

2.1: Summary Statistics

Summary statistics are computed from daily SP500TR performance and then annualized. The first four moments of S&P returns are summarized below:

Table 4: Historical Equity Return Analysis

Year	Average Return	Standard Deviation	Skew	Excess Kurtosis
2020	22.88%	34.50%	-55%	8.04
2019	28.16%	12.47%	-58%	3.27
2018	-3.03%	17.03%	-43%	3.13
2017	19.98%	6.67%	-42%	2.84
2016	12.16%	13.10%	-38%	2.35
2015	2.57%	15.49%	-17%	1.92
2014	13.48%	11.36%	-39%	1.34
2013	28.68%	11.07%	-33%	1.43
2012	15.65%	12.70%	8%	0.84
2011	4.80%	23.29%	-41%	2.75
2010	15.66%	18.06%	-14%	2.03
2009	27.19%	27.27%	4%	2.03
2008	-37.76%	41.05%	19%	4.00
2007	6.62%	15.96%	-44%	1.46
2006	15.17%	10.02%	14%	1.20
2005	5.32%	10.29%	0%	(0.13)
2004	10.95%	11.10%	-9%	(0.12)
2003	26.68%	17.07%	9%	0.81
2002	-21.60%	26.04%	49%	0.80
2001	-10.38%	21.38%	9%	1.52
2000	-7.09%	22.22%	7%	1.38

Observations

- Despite the pandemic, average annual equity returns were still quite high (22.88%).
- Markets were very volatile in 2020 (34.50%). The only year in the table above that was more volatile than 2020 was the financial crisis of 2008.
- Negative skew for the past 9 out of 10 years, indicating a tendency for a heavy left tail containing negative returns.
- Equity return excess kurtosis was quite large at 8.04. Recall that a normal distribution has a kurtosis of 3 and excess kurtosis of 0. Therefore, 2020 equity returns had substantially heavier tails than a normal distribution. In fact, 2020 had excess kurtosis more than twice that of any other year. Distributions with heavier tails than the normal distribution are called *leptokurtic*.

2.2: Non-Parametric Tests

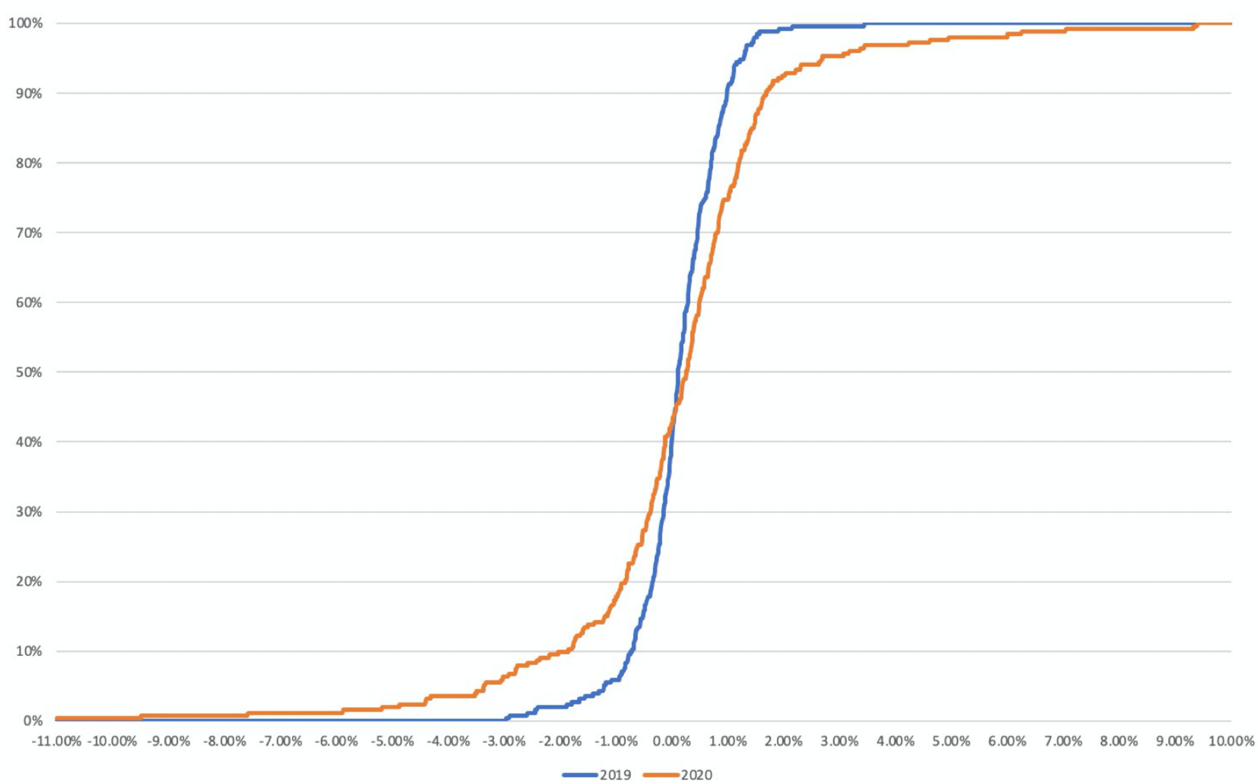
Next, this section compares 2019 and 2020 equity return performance to see if they are statistically significantly different. This can be done using the Kolmogorov–Smirnov test, a non-parametric test, which has its statistic computed as the maximum difference between the cumulative distribution functions for 2019 and 2020 equity returns:

$$D = \sup_x |F_{2019}(x) - F_{2020}(x)|$$

- When comparing 2019 vs 2020 equity returns, the Kolmogorov–Smirnov test results are:
  - H_0: 2019 and 2020 equity returns are drawn from the same distribution (e.g. normal)
  - H_1: 2019 and 2020 equity returns are drawn from different-distributions
  - $D =$ Maximum Cumulative Distribution Function Difference $= 16.97\%$
  - $p = .0014 < 5\%$
- Therefore, the null is rejected because of the small p-value and 2019 vs 2020 equity returns are statistically significantly different.

The cumulative distribution functions of 2019 and 2020 equity returns are visualized below. Notice from the graph that the 2020 equity returns have substantially fatter tails on both sides:

Figure 2: Equity Return CDF (2019 vs 2020)



Note that while the previous analysis compared 2019 vs 2020 equity returns, this exercise can be repeated to compare any two years. Repeating the Kolmogorov–Smirnov test across all pairs of years between 2005 to 2020 yielded the following insights:

- 2008, 2017, and 2020 were the most "unique" years for equity returns, meaning that these years were statistically the most different from other years.
- Each of these three years had different kinds of returns:
  - 2008: $\mu = -37.8\%$, $\sigma = 41.0\%$ (Markets Drop and High Volatility)
  - 2017: $\mu = 29.0\%$, $\sigma = 6.7\%$ (Markets Up and Low Volatility)
  - 2020: $\mu = 22.9\%$, $\sigma = 34.5\%$ (Markets Up and High Volatility)
- Both 2008 and 2017 were statistically significantly different than all other years analyzed (2005 and beyond). 2020 was statistically different from all years 2012 and beyond.
- Another interesting observation was that after standardizing market returns (i.e. subtracting by the mean and dividing by the standard deviation), p-values increased and there were no pairs of years that were statistically significantly different at the 1% level. In other words, much of the difference in returns can be explained by inspecting the first two moments: average returns and volatility.
- *Note: Detailed output from the Kolmogorov–Smirnov test and additional discussions on this topic are included in Appendix 1.*

Section 3: Model Comparisons

3.1: Normal Distribution

The Black-Scholes model assumes that stock returns are normally distributed and stock prices are lognormally distributed. Calibration (discussed further in the appendix) yields $\mu = 11\%$ and $\sigma = 17\%$. As seen earlier and in the table below, directly using this assumption yields a Z-score of magnitude 8 or above. This corresponds to an extremely low probability, lower than 1-in-1 trillion.

Table 5: Extreme Return Analysis with the Normal Distribution

Day	Return	Z-Score
3/16/20	-11.98%	-10.97
3/12/20	-9.49%	-8.70
3/24/20	9.39%	8.53
3/13/20	9.32%	8.46

Therefore, the returns above are not well-explained by a normal distribution. The remainder of this section explores four additional models:

- Student's t-Distribution
- Regime Switching Lognormal (RSLN-2)
- Extreme Value Theory (EVT)
- Poisson Jump Diffusion

3.2: Student's t-Distribution

The Student's t-distribution is similar to the normal distribution. If the degrees of freedom parameter is set to a sufficiently small number, the Student's t-Distribution does differ from the normal distribution in that it does have fatter tails than a normal distribution. Given the large positive excess kurtosis in the historical data in Section 2.1, this seems like potentially a more reasonable distribution choice that may fit the historical data better than the normal distribution.

We can reanalyze the extreme return probabilities but now using the Student's t-Distribution with 3 degrees of freedom:

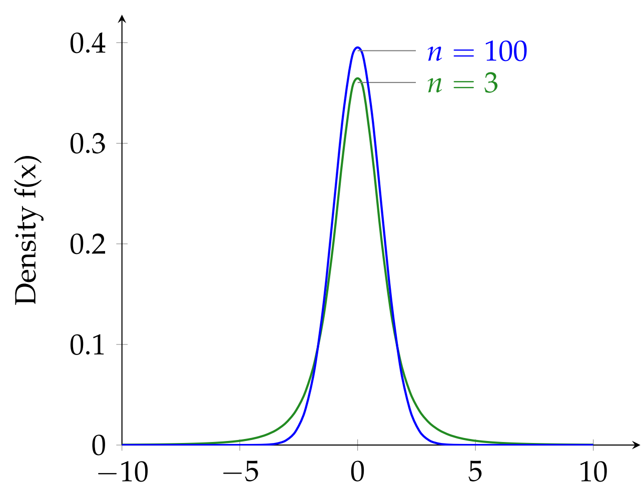
Table 6: Extreme Return Analysis with the Student's t-Distribution

Day	Return	Tail Probability
3/16/20	-11.98%	0.16%
3/12/20	-9.49%	0.32%
3/24/20	9.39%	0.32%
3/13/20	9.32%	0.35%

Recall the tail probabilities associated with these returns for a normal distribution were less than 1 in 1 trillion. The table above using the Student's t-Distribution assigns much more reasonable probabilities between .16% and .35%. The Student's t-Distribution has a much heavier tail, and thus assigned much more reasonable probabilities to extreme returns.

In the figure below, the density of the Student's t-Distribution is plotted for both a degrees of freedom of 3 and 100. Once n is at least approximately 30, then the Student's t-Distribution is similar to a normal distribution. Therefore, the blue curve is quite similar to a normal distribution. When $n = 3$, the tails are substantially fatter, as seen in the figure below:

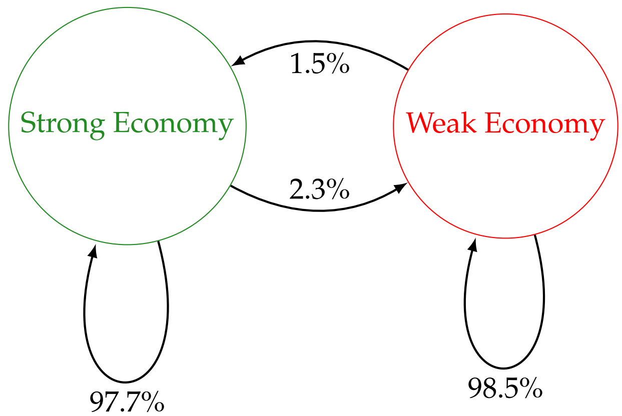
Figure 3: Student's T Density Plot



3.3: Regime Switching Lognormal (RSLN-2)

The RSLN-2 model is similar to a normal distribution, except there are two states that lead to pairs of normal distribution parameters. When the economy is strong, $\mu = 124\%$, $\sigma = 7.2\%$. When the economy is weak, $\mu = -.1\%$, $\sigma = 34.9\%$. The model transitions between a strong and weak economy based on the probabilities in the diagram below. For example, if today is a strong economy, then there is a 97.7% probability of the strong economy state tomorrow and 2.3% probability of the weak economy state tomorrow.

Figure 4: Visualization of RSLN-2 Model



The Z-scores with the RSLN-2 model are computed below:

Table 7: Extreme Return Analysis with RSLN-2

Day	Return	Z-Score
3/16/20	-11.98%	-5.4
3/12/20	-9.49%	-4.3
3/24/20	9.39%	19.6
3/13/20	9.32%	19.4

The above table shows that market drops are explained better, but market gains are explained worse compared to a normal distribution. RSLN-2 often assigns high return, low volatility parameters to the strong economy and low return, high volatility parameters to the weak economy. This typically means the RSLN-2 is capable of modeling highly adverse returns. On the other hand, because the strong economy is typically lower volatility, the RSLN-2 model often does not adequately explain extreme market rallies.

Therefore, adverse returns have more reasonable (but still fairly large) Z-scores of around 4-5. However, the up-market returns are quite extreme at around 19. To summarize, so far, out of the models analyzed, the Student's t-distribution has given the most reasonable tail probabilities.

3.4: Extreme Value Theory

Extreme Value Theory (EVT) is a branch of statistics dealing with the extreme deviations from the median of probability distributions (Levine, 14). A threshold return u is chosen, and then all excess observations past threshold u have a Generalized Pareto Distribution. The Generalized Pareto Distribution (GPD) has two parameters: scale parameter s and shape parameter k. The cumulative distribution function (CDF) is given by:

$$G_{s,k}(x) = \begin{cases} 1 - \left(1 - \frac{kx}{s}\right)^{\frac{1}{k}} & k \neq 0 \\ \\ 1 - e^{\frac{-x}{s}} & k = 0 \end{cases}$$

The CDF for tail returns $|x| \geq u = 5\%$ is given by $F(x) = .05 \cdot G_{s,k}(x-u) + .95$. Note that the .05, .95, s, and k values are determined using the methodology in Levine (14-17). The excess observation past the threshold u has GPD distribution. Using EVT to explain extreme returns in 2020 gives the following probabilities:

Table 8: Extreme Return Analysis with EVT

Day	Return	Tail Probability
3/16/20	-11.98%	.00031%
3/12/20	-9.49%	.00055%
3/24/20	9.39%	.00055%
3/13/20	9.32%	.00056%

Thus, EVT does assign more realistic probabilities to extreme 2020 equity returns than the normal and RSLN-2 distributions, but less reasonable than Student's t. The probabilities above correspond to roughly 1 in 200,000 events. Since four of these extreme events occurred in a month, these probabilities are still likely too low to adequately explain the tails. Thus, while this is better at explaining tail behavior than normal and EVT, it is still far from perfect and still does not assign sufficient tail weight.

3.5: Poisson Jump Diffusion

The stock price process under Black-Scholes follows:

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t$$

The Poisson Jump Diffusion model incorporates the possibility there is a large stock price jump:

- A jump happens when $dN_t = 1$ and there is no jump when $dN_t = 0$
- The jump probability is λdt
- The Poisson Jump Diffusion model incorporates an additional jump term $(e^J - 1)dN_t$, which is the jump size $(e^J - 1)$ times the jump indicator dN_t
- $J \sim N(0, 1)$ which toggles the jump size
- The *expected proportional jump size* is $\kappa = E(e^J - 1)$

The jump-diffusion model can then be based on:

$$\underbrace{\frac{dS_t}{S_t}}_{\text{Percentage Change in Stock Price}} = \underbrace{(\mu - \lambda\kappa)dt}_{\text{Drift Term}} + \underbrace{\sigma dW_t}_{\text{Normal Randomness}} + \underbrace{(e^J - 1)dN_t}_{\text{Jump Term}}$$

Note that the $\lambda\kappa$ subtraction term is simply an adjustment to the drift so that introducing jumps does not change the expected path. Using the Poisson Jump Diffusion model to compute the probability of March 2020 equity returns gives:

Table 9: Extreme Return Analysis with Poisson Jump Diffusion

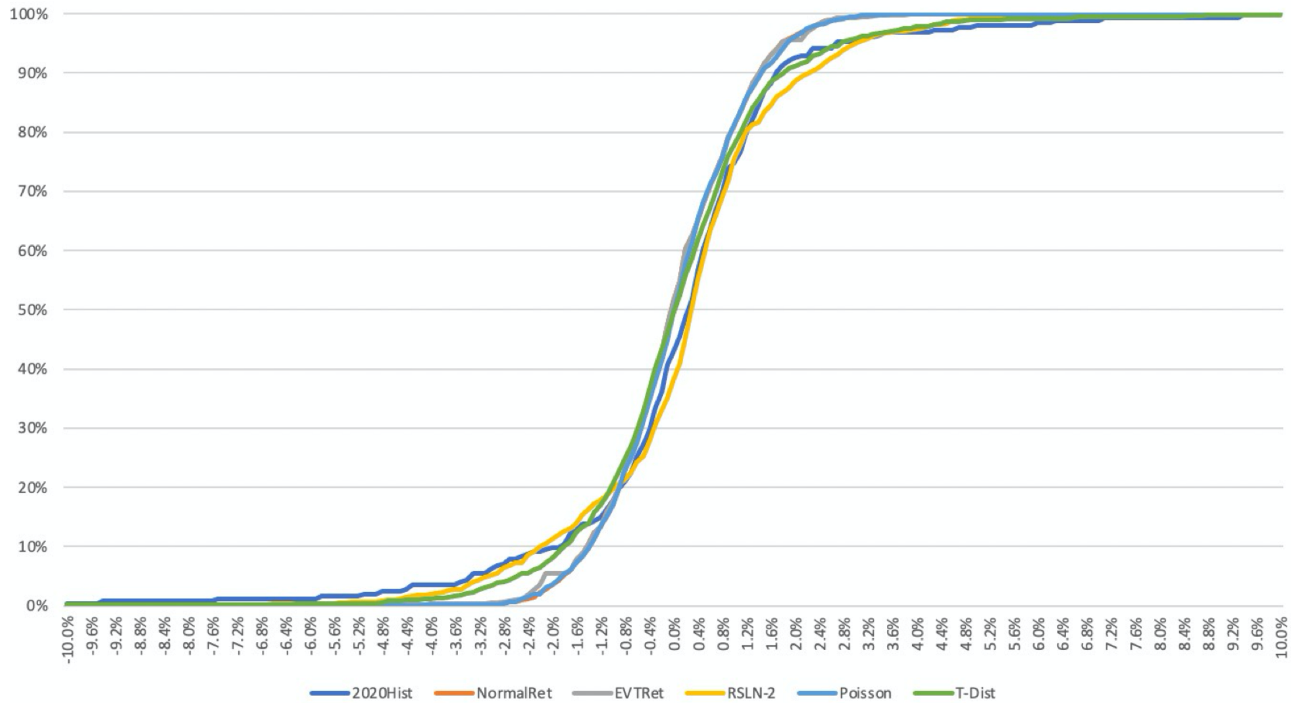
Day	Return	Tail Probability
3/16/20	-11.98%	.0133%
3/12/20	-9.49%	.0137%
3/24/20	9.39%	.0159%
3/13/20	9.32%	.0159%

In the table above, the probabilities of the extreme returns are around .013% - .016%, which seems very reasonable. To summarize, the Poisson Jump Diffusion and Student's t-distribution are the two models that are best able to explain the sharp market moves when comparing the results explored in Section 3.

3.6: Model Comparison

So far, Section 3 has compared different-distribution assumptions (Normal, Student's t, RSLN-2, Extreme Value Theory, Poisson Jump Diffusion) along with their ability to explain the four most extreme daily market moves in March 2020. This analysis gives insight into how well each of the distributions models extreme returns. Next, the Kolmogorov–Smirnov test is run and CDFs are graphed to give broader insight into the fit of the entire distribution (and not just extreme events):

Figure 5: CDF Comparison



The maximum difference between each modeled CDF (normal, Student's t, RSLN-2, Extreme Value Theory, Poisson Jump Diffusion) and the empirical CDF (2020 historical equity returns) is used to compute the Kolmogorov–Smirnov statistic which gives the associated p-values shown in the table below:

Table 10: Kolmogorov–Smirnov Results

Model	p-value
Normal	0.4%
Student's t	23.8%
RSLN-2	43.8%
EVT	0.4%
Poisson Jump	1.7%

The RSLN-2 are Student's t models are the only models that fails to reject the null at a 5% confidence level, indicating that they are potentially overall the strongest of the models in explaining 2020 equity returns.

Model Comparison Conclusions

- The null and alternative hypothesis, using the Kolmogorov–Smirnov statistic (i.e. the maximal difference in CDF) to compare the models is stated below:
  - H_0: Modeled returns (normal, RSLN-2, etc.) explain 2020 historical returns
  - H_1: Modeled returns (normal, RSLN-2, etc.) do not explain 2020 historical returns
- The normal, EVT, and Poisson Jump models all have p-values less than 5%, indicating that they do a poor job of fully describing 2020 returns.
- Note that while Poisson Jump explained sharp movements well, as seen in Section 3.5, it did not explain the body of the distribution well. The reason for this is that the jump probability is only about .03%, and the other 99.97% of draws are similar to a normal distribution (which also does not explain the body of the distribution well)
- Note that Sections 3.1-3.5 explored the ability to explain extreme returns, while Section 3.6 is a more holistic Kolmogorov–Smirnov test which analyzes the fit of the entire distribution and not just extreme returns. The Poisson Jump and Student's t models appear best at explaining extreme returns, while RSLN-2 and Student's t fit the entire 2020 equity return distribution the best.
- The RSLN-2 and Student's t models are the only models that fail to reject the null, indicating they are potentially the strongest of the models in explaining overall 2020 equity returns. Note that this by no means implies these are perfect models, and advantages and disadvantages of each model are discussed next.

3.6: Model Advantages and Disadvantages

To finish Section 3, *the table below highlights advantages and disadvantages of the different models in the context of their ability to model 2020 equity returns:*

Model	Advantages	Disadvantages
Normal	◦ Well-known ◦ Easy to communicate ◦ Intuitive ◦ Few, easy-to-understand parameters	◦ Tails not heavy enough to explain market performance ◦ Likely understates the true level of risk ◦ Unable to model skewness
Student's t	◦ Can explain extreme market jumps and heavy tails ◦ Able to explain 2020 equity returns ◦ Shares much of the intuition from the normal distribution	◦ Judgment needed to determine degrees of freedom parameter ◦ Unable to model skewness
RSLN-2	◦ Able to explain 2020 equity returns ◦ Parameters relatively easy to interpret ◦ Able to explain boom and bust cycles	◦ May not be able to adequately explain sharp market increases ◦ Difficult parameterization ◦ Complex
EVT	◦ Can explain tails heavier than a normal distribution ◦ EVT doesn't solve all problems but you can start to quantify extremes with EVT	◦ Arbitrary threshold u ◦ Hard to interpret ◦ Complex
Poisson Jump	◦ Can explain extreme market jumps and heavy tails	◦ Can fit extreme cases but often not the body of the distribution ◦ Hard to interpret ◦ Complex

Section 4: Types of Randomness (Mild vs Wild)

Despite its disadvantages, the normal distribution is still an enormously popular tool that has gained recognition in a number of fields:

"Conventional studies of uncertainty, whether in statistics, economics, finance, or social science, have largely stayed close to the so-called bell curve, a symmetrical graph that represents a probability distribution. Used to great effect to describe errors in astronomical measurement by the 19th-century mathematician Carl Friedrich Gauss, the bell curve, or Gaussian model, has since pervaded our business and scientific culture, and terms like sigma, variance, standard deviation, correlation, R-square, and Sharpe ratio are all directly linked to it. Neoclassical finance and portfolio theory are completely grounded in it." (Diebold, 86)

Diebold describes that the world in fact has two kinds of randomness: "mild randomness" and "wild randomness". Quantities that follow the normal distribution or other distributions without large deviances have mild randomness. Wild randomness, on the other hand, occurs when "a single observation or particular number can impact the total in a disproportionate way." Variables such as height, weight and IQ all obey mild randomness. In contrast, financial quantities such as wealth and stock returns exhibit wild randomness. As an example, you would not expect any adult to be 100 times taller than another adult. Yet, many people such as Bill Gates are easily more than 100 times wealthier than the vast majority of adults. While the normal distribution is often adequate to describe mild randomness, the normal distribution does not have sufficiently heavy tails to describe wild randomness. Therefore, those that rely on the normal distribution, such as those using the Black-Scholes model for stock returns, need to understand and appreciate the model shortcomings and associated risks.

The fact that the normal distribution does not adequately predict rare events is a well-known phenomenon in many areas, including financial markets. The prior analysis of extreme 2020 equity returns provides a remainder that this is still the case.

Nicholas Taleb, a former options trader, acknowledges this idea in his book *The Black Swan: The Impact of the Highly Improbable*. His book emphasizes our blindness with respect to randomness, particularly large deviations. Taleb emphasizes that attempting to predict rare, "Black Swan" events is often futile; instead, focus on building robustness to extreme events. This means that firms should both be prepared to take full benefit of positive events, but have adequate resources to prevent failure under conditions of adverse change. This closely coincides with the ideas of stress testing and contingency planning. In other words, none of the distributions from Section 3 are perfect. Qualitative methods should be employed to supplement any quantitative modeling to ensure firms have a broad understanding of their risk exposures.

Section 5: Conclusion

Key Takeaways

- March 2020 had four days with returns over an absolute magnitude of 9%. This corresponds to a Z-score when using a normal distribution of over 8. The probability $Z \geq 8$ is extremely small, and so the normal distribution does not adequately explain stock return movements.

- Equity returns in 2020 had extremely heavy tails with an excess kurtosis of over 8, more than twice of any other year in the dataset. Additionally, the large excess kurtosis of 8 indicates much heavier tails than a normal distribution.
- Equity return data frequently exhibits negative skew (e.g. years 2013 and beyond).
- The Poisson Jump model is able to explain extreme returns well from a probabilistic perspective. However, while the Poisson Jump can model thicker tails, it may sacrifice substantial fit in the body of the distribution and not perform well to explain typical cases.
- The RSLN-2 and Student's t-distributions were overall the strongest of the models at explaining 2020 equity returns, and were the only distributions statistically able to explain 2020 market performance when using the Kolmogorov-Smirnov test.
- The Student's t-distribution was the only model to pass the Kolmogorov-Smirnov test and have reasonable extreme return probabilities. Therefore, using these two tests as the criteria, the Student's t-distribution was the best of the models at explaining 2020 equity returns.
- No one particular model is always better than any other. There are advantages and disadvantages of each equity model considered in this paper. Analyzing risk under multiple different-distribution assumptions and parameterizations, as well as supplementing with qualitative analysis, may give a more holistic picture of risk.

Looking Towards The Future

- Practitioners already know that the normal distribution is an imperfect distribution assumption, and does not truly reflect the behavior of the stock market.
- The Black-Scholes options pricing model assumes that stock price returns follow a normal distribution. However, stock price returns have a much heavier tail than the normal distribution (i.e. leptokurtic), and so directly applying the Black-Scholes framework will often underprice options that are far out of the money. This is due to the fact that the normal distribution gives nearly no weight to Z-scores above 8, but these extreme market returns are in fact observed and have material financial implications.
- Those relying on the normal distribution or Black-Scholes model need to understand the shortcomings of these models and that they may underestimate tail risk
- Alternatives to the Black-Scholes model could be considered, such as stochastic volatility models or models that consider smile/smirks in volatility surfaces.
- Risk management professionals can supplement with stress testing and scenario analysis to more rigorously quantify the impact and prepare for extreme events.
- There is no perfect-distribution assumption: common sense, expert judgment and supplemental methods such as sensitivity tests should all be used when modeling.
- Perhaps the only certainty about markets is that they will continue to remain uncertain. Financial planning should consider the reality that market turbulence may persist.

Appendices

Appendix 1: Kolmogorov–Smirnov Test - Detailed Results

Section 2.2 explored the Kolmogorov–Smirnov test, comparing 2019 vs 2020 historical equity returns. This appendix takes a deeper dive and compares all pairs of years from 2005-2020. Note that no distribution assumption is made (e.g. normal) and these tests are based on empirical data. The table below shows the results of the Kolmogorov–Smirnov tests, where each entry is the p-value rounded to the nearest percentage (with values significant at the 5% level highlighted in yellow):

Table 11: Kolmogorov–Smirnov Test - Unadjusted Equity Returns

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
2005	100%															
2006	11%	100%														
2007	16%	11%	100%													
2008	0%	0%	0%	100%												
2009	0%	0%	3%	4%	100%											
2010	11%	12%	90%	0%	9%	100%										
2011	1%	0%	31%	2%	62%	25%	100%									
2012	53%	57%	56%	0%	1%	44%	6%	100%								
2013	14%	1%	20%	0%	0%	20%	0%	8%	100%							
2014	41%	31%	53%	0%	0%	20%	1%	49%	62%	100%						
2015	6%	9%	10%	0%	3%	17%	14%	61%	6%	3%	100%					
2016	54%	65%	60%	0%	0%	41%	3%	79%	17%	76%	9%	100%				
2017	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	2%	100%			
2018	20%	6%	94%	0%	5%	87%	21%	62%	13%	24%	49%	39%	0%	100%		
2019	29%	10%	34%	0%	0%	25%	1%	29%	94%	35%	2%	29%	0%	20%	100%	
2020	0%	0%	1%	1%	96%	4%	31%	0%	0%	0%	0%	0%	0%	2%	0%	100%

For example, 2008 and 2020 are statistically significantly different-distributions, as noted by the circled 1% p-value in the first table. Next, the same exercise is repeated except with standardized returns, where standardization is done by subtracting by the mean and dividing by the standard deviation of equity returns within each year. Interestingly, after standardization, the p-value increases substantially to 41% and the distributions are no longer statistically significantly different after standardization.

Table 12: Kolmogorov–Smirnov Test - Standardized Equity Returns

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
2005	100%															
2006	21%	100%														
2007	21%	11%	100%													
2008	8%	37%	31%	100%												
2009	29%	38%	71%	81%	100%											
2010	25%	39%	85%	88%	97%	100%										
2011	25%	28%	95%	52%	94%	76%	100%									
2012	33%	50%	19%	50%	47%	69%	34%	100%								
2013	25%	18%	10%	3%	35%	20%	29%	16%	100%							
2014	41%	33%	79%	16%	83%	62%	83%	45%	47%	100%						
2015	35%	65%	12%	40%	69%	47%	47%	99%	35%	47%	100%					
2016	20%	33%	31%	67%	69%	89%	54%	49%	25%	76%	54%	100%				
2017	4%	40%	6%	40%	33%	27%	26%	88%	3%	10%	78%	23%	100%			
2018	12%	20%	97%	65%	78%	96%	94%	27%	26%	70%	18%	57%	11%	100%		
2019	20%	52%	49%	28%	89%	62%	76%	87%	47%	97%	89%	69%	31%	57%	100%	
2020	3%	4%	28%	41%	28%	40%	39%	10%	5%	11%	3%	44%	4%	13%	7%	100%

Observations

- In the first table, the following years for equity returns stand out as particularly unique: 2008, 2017 and 2020.
- Remember that each of these three years had different kinds of returns:
  - 2008: $\mu = -37.8\%$, $\sigma = 41.0\%$ (Markets Drop and High Volatility)
  - 2017: $\mu = 29.0\%$, $\sigma = 6.7\%$ (Markets Up and Low Volatility)
  - 2020: $\mu = 22.9\%$, $\sigma = 34.5\%$ (Markets Up and High Volatility)
- Upon standardization[2] , much of the differences can be explained. Notice in the second table, far fewer cells are yellow. And, in fact, none of the differences are significant at the 1% level. Of the cells that remain significant at the 5% level, most are related to 2020, likely related to the large excess kurtosis above 8 in 2020.

Appendix 2: Parameterization Documentation

The analysis presented in this paper is based on daily SP500TR price levels. Calibration was done over the historical period from 1990-2019. This allowed data to be fit based on past historical data to see if it could explain 2020 market movements.

The specific parameterization methodology used for each of the distributions is described below:

1. **Normal Distribution**. Maximum Likelihood Estimation (MLE) was used. This means that:

$$\widehat{\mu} = \bar{x} = \sum_{i=1}^{n} \frac{x_i}{n} \quad \text{and} \quad \widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$$

2. **Student's t-Distribution**. The Student's t-distribution was analyzed using the normal MLE parameters above. The degrees of freedom (DOF) parameter was set to 3. The DOF parameter was set by minimizing the Kolmogorov–Smirnov statistic as seen in the table below:

Table 13: Degrees of Freedom (DOF) Parameter Selection

DOF	K/S Stat	March 16th Probability
1	0.096	5.79%
2	0.085	0.82%
3	0.073	0.16%
4	0.074	0.04%
5	0.074	0.01%
10	0.075	0.00%

[2]Standardization is the process of subtracting the mean and then dividing by the standard deviation. After performing additional tests, we noticed that the increase in p-values was driven more so by the division by the standard deviation than the netting against the mean.

3. **RSLN-2**. MLE was also used to calibrate the RSLN-2 model [3]. Mean and standard deviations for each of the two states are estimated as well as the transition probabilities. Because the maximization algorithm is computationally burdensome, only for RSLN-2, weekly return data was used from 1990 to 2019 instead of daily returns. Note that it is possible an RSLN model with more than 2 regimes may have had stronger results (e.g. adding a third regime for an up-state with high volatility); analysis of this is beyond the scope of this paper.

4. **Extreme Value Theory (EVT)**. Implementing the EVT model requires estimating the s and k parameters. Parameter estimation techniques followed the methodology set forth by Levine (14-17) using the method of moments technique. The threshold parameters u are set based on the 5th and 95th percentile returns as discussed by Levine (14-17).

   The formulas below summarize the process for applying the method of moments:

   Notation

   - $\bar{x}$ is the mean of the excesses
   - w is the mean of the squared excesses
   - s^* is the method of moments parameter estimate for s
   - k^* is the method of moments parameter estimate for k

   Formulas

   - $\bar{x} = \frac{1}{n}\sum_i (x_i - u)$
   - $w = \frac{1}{n}\sum_i (x_i - u)^2$
   - $A = \frac{\bar{x}^2}{w - \bar{x}^2}$
   - $s^* = \frac{1}{2}\bar{x}(A+1)$
   - $k^* = \frac{1}{2}(A-1)$

5. **Poisson Jump Diffusion**. The coefficient of the Wiener increment was computed using MLE. Calibration of the λ paper is beyond the scope of this paper and a fixed value of $\lambda = .075$ was chosen. Because lambda only impacts tail risk, it does not have a material impact on the Kolmogorov–Smirnov test results.

Bibliography

International Actuarial Association. *Stochastic Modeling Theory and Reality from an Actuarial Perspective*. I-21 to I-23. 2010.

Levine, Damon. *Modeling Tail Behavior with Extreme Value Theory*. 14-17. September 2009.

Taleb, Nicholas. The Black Swan: The Impact of the Highly Improbable. Random House, 2007.

Diebold, Francis X. The Known, the Unknown, and the Unknowable in Financial Risk Management: Measurement and Theory Advancing Practice. Princeton University Press, 2010.

[3]The process of calibrating RSLN-2 parameters using MLE is discussed in (International Actuarial Association, I-21 to I-23).