A Quantitative Analysis of the Impact of COVID-19 on the Video Gaming Industry

Dr. Mingwei Sun [1] and James Harrison [1]

Abstract

The impact of the Covid-19 pandemic has been growing at a tremendous speed all over the world since its outbreak in 2020, which resulted in massive uncertainty and disorder in many fields including the economy, society, politics and culture. It is reported that gaming usage has been increasing during the pandemic. In this research, we analyze the impact of Covid-19 on the video gaming industry in the United States. A quantitative analysis which focuses on the behavior of major gaming companies' stock prices and the video gaming user data before and during the Covid-19 pandemic is performed. It shows that daily player counts and average concurrent players increased dramatically with the advent of Covid-19. However, there is only a weak correlation between the concurrent Covid-19 cases and the change in daily players, which suggests that the pandemic introduced many people to video games as a more permanent new hobby instead of as a temporary alternative to entertainment and social interaction. Other important results include that the stock prices of major gaming companies did not alter significantly due to Covid-19. With the exception of a few drops, their stock prices during Covid-19 have continued to grow and recover from drops in a similar manner to before the pandemic.

Keywords: Covid-19, Pandemic, Video Game, Stock Price

1 Introduction

WHO announced Covid-19 as a pandemic in March 2020 (WHO, 2020). Since then, it has had a widespread impact on society and the global economy. In the months following, schools across

[1]Department of Mathematics and Computer Science, Samford University, Birmingham, AL USA.

the United States shut down in-person operations, with many shifting towards virtual learning. Small businesses were also hit hard, with 31% reporting as inoperable in May of 2020. Even as of July 2021, there were 37.5% fewer small businesses open than in January 2020 (Nichols, 2021). Unemployment rose from 3.6% to 13% nationwide (Smith, 2021).

Measures such as quarantining, mask mandates, and social distancing were implemented. This reduced the availability of entertainment and social interaction. One way to remain entertained and socially connected is video games. Self reported responses found that video games had a positive effect on well being, and served as an "enjoyable means of maintaining social contact, and a stress relieving and mentally stimulating escape from the effects of lockdown"(Barr, 2021).

The Cowen Industrial Report for March 16, 2020 predicted that the video game industry would benefit from containment measures, at least in the short-term, and that the video game industry should hold up well during the pandemic. In both the 2001 and 2008-09 recession, the video game industry saw healthy growth, and the report predicts that will be the case during Covid-19 as well (Creutz, 2020). As for the short-term benefit, two weeks after Covid-19 was announced as a pandemic video game sells increased 82% from the previous week, with two games accounting for 71.2% of this increase (Statista, 2021).

Our literature review suggests that the video game industry benefited during the Covid-19 pandemic. In this paper, we examine this assumption by analyzing the video game industry before and during Covid-19.

2 Methodology and Data

2.1 Data Sources

In order to analyze and model the impact of Covid-19 on the video game industry, we gathered daily data on the number of Steam players in the US, the stock price of several major gaming companies, and the number of Covid-19 cases in the US. We gathered the player data and stock data for all days available from September 1, 2018 through August 31, 2021, or 18 months before and 18 months after March 1, 2020. To gather data on players we utilized Steam DB (https://steamdb.info), a 3rd party software that tracks user data for the popular game engine

Steam. Steam is currently one of the most used gaming platforms. Being popular and accessible, we use it as a representative sample of the larger gaming community. We acquired the stock value for our chosen companies through Yahoo Finance (https://finance.yahoo.com), and use the opening price for each day as the daily stock price. Stock prices are only available for non-holiday weekdays, but we will continue to refer to it as daily stock data. The companies we chose for our sample of the industry are Microsoft, Sony, Activision, Nintendo, and EA. We used the CDC's reports (https://covid.cdc.gov/covid-data-tracker/datatracker-home) to gather the daily Covid-19 data from March 1, 2020 to August 31, 2021 as their reporting should be the most accurate.

2.2 Data Collection and Programming

For our data collection, both the CDC and Yahoo Finance provide easily accessible tables that can be downloaded as a csv file. Collecting daily Steam player data was similar but more involved. Steam db is a 3rd party tool, which interfaces with Steam by using Steamkit to collect all of its data. It contains data on every game available on steam, including player data, release date, patch notes and updates, etc. Daily Steam players is available as a csv file under graphs, concurrent Steam Users, Lifetime concurrent users on Steam.

Once our data was collected in excel, we converted the files from csv into xlsx. We then trimmed it to only contain data between September 1st, 2018 to August 31st, 2021, or 18 months prior to and following March 1st, 2020. The main manipulations we did to our data was to split the data into two columns, before and after March 1st, 2020. The reason we converted to xlsv file type was to allow for convenient importing into Rstudio. All programming for this project took place in R.

Using R we wrote scripts to import our data from excel, perform preliminary tests on our data, perform hypothesis tests, finally to graph our data. Our preliminary tests were to determine the normality of our data distribution, which we determined through a Shapiro-Wilk hypothesis test. We also created Q-Q plots of our tested data to show graphically where they correspond to a normal distribution curve, however we do not include these in this paper.

For each hypothesis test, we determined which was most appropriate and then passed it parameters for the data we wanted to use and for what our alternative hypothesis was.

2.3 Statistical Analysis

We took our Steam player data and stock price data and split both into two groups, the 18 months prior to March 1st, 2020 and the 18 months including and following the same date. We label these groups as before Covid-19 and during Covid-19, respectively. We also took our stock data and created two additional data sets, 30 days before March 1, 2020 and 30 days including and following the same day. These are 60 days of available data, not calendar days, and go from January 16, 2020 to April 13, 2020. All data and data manipulations are stored in excel files.

We proceed by performing a series of hypothesis tests to test our data. We use the statistical software R for all tests and models. Several of our hypothesis tests will use the T-Test to compare the means of two sets. We use the T-test because we do not know the population variance. In order to use the T-Test, both sets sample means need to be approximately normally distributed. We satisfy this condition using the Central Limit Theorem, or CLT, which holds that for any population with sample size n $\geq$ 30 the sample means follow a normal distribution. Since all of our samples have at least 30 data points, we use the CLT for all T-tests.

2.3.1 Hypothesis Test 1:

The number of daily Steam players has increased since the pandemic began.

$$H_0 : \mu_0 = \mu_1$$
$$H_A : \mu_0 < \mu_1$$

H_0 represents our null hypothesis, H_A our alternative hypothesis, and μ_0 and μ_1 represents the mean of our daily player data before Covid-19 and during Covid-19, respectively. Since both data sets have a few hundred data points, we apply the Central Limit Theorem and use a T-test to compare the two sets. We use a significance level $\alpha = 0.05$. If our p-value $p < 0.05$, we reject our null hypothesis.

$\bar{x}_1$	$\bar{x}_2$	t	df	α	p-value
15296300	22918550	-73.632	850.6	0.05	$<$ 2.2e-16

Figure 1: T-test results for player data

Since our p-value $p < 0.05$, we reject our null hypothesis. It is worth noting that 2.2e-16 is the minimum value that R can calculate for p-values.

2.3.2 Hypothesis Test 2:

The number of daily concurrent Steam players has increased since the pandemic began.

$$H_0 : \mu_0 = \mu_1$$
$$H_A : \mu_0 < \mu_1$$

H_0 represents our null hypothesis, H_A our alternative hypothesis, and μ_0 and μ_1 represents the mean of our concurrent daily player data before Covid-19 and during Covid-19, respectively. Since both data sets have a few hundred data points, we apply the Central Limit Theorem and use a T-test to compare the two sets. We use a significance level $\alpha = 0.05$. If our p-value $p < 0.05$, we reject our null hypothesis.

$\bar{x}_1$	$\bar{x}_2$	t	df	α	p-value
4459485	6022452	-43.676	1057.4	0.05	$<$ 2.2e-16

Figure 2: T-test results for concurrent player data

Since our p-value $p < 0.05$, we reject our null hypothesis.

2.3.3 Hypothesis Test 3:

There exists a correlation between change in daily Steam players and new daily Covid-19 cases.

Before we test our hypothesis, we need to determine which test we will use to test for correlation. If both sets of data are normally distributed, we will use the Pearson Correlation test. Otherwise, we will use Kendall Rank Correlation test.

We test this condition using the Shapiro-Wilk Normality test on both data sets. The null hypothesis is that the population is normally distributed. We use $\alpha = 0.05$, and if $p < 0.05$ we reject the null hypothesis.

W	α	p-value
0.93446	0.05	9.211e-15

Figure 3: Shapiro-Wilk results for change in users

Because our p-value for both sets $p < 0.05$, we reject our null hypothesis and proceed with Kendall Rank Correlation.

W	α	p-value
0.91337	0.05	$<$ 2.2e-16

Figure 4: Shapiro-Wilk results for new Covid-19 cases

$$H_0 : \tau = 0$$
$$H_A : \tau \neq 0$$

H_0 represents our null hypothesis, H_A our alternative hypothesis, and τ is our correlation coefficient between the change in daily Covid-19 cases and the change in daily players. We use a significance level $\alpha = 0.05$. If our p-value $p < 0.05$, we reject our null hypothesis.

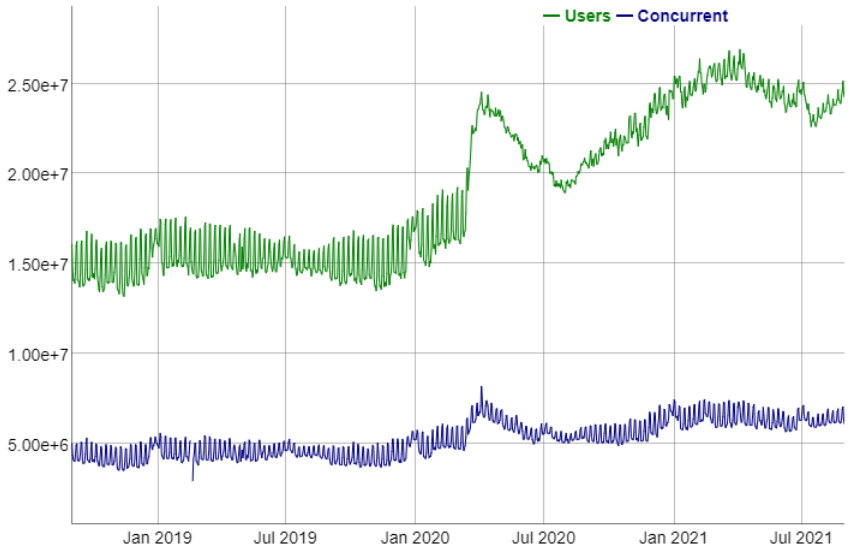
z	τ	α	p-value
-5.1772	-0.1480494	0.05	2.252e-07

Figure 5: Correlation test results for User-Case data

Since our p-value $p < 0.05$, we reject our null hypothesis.

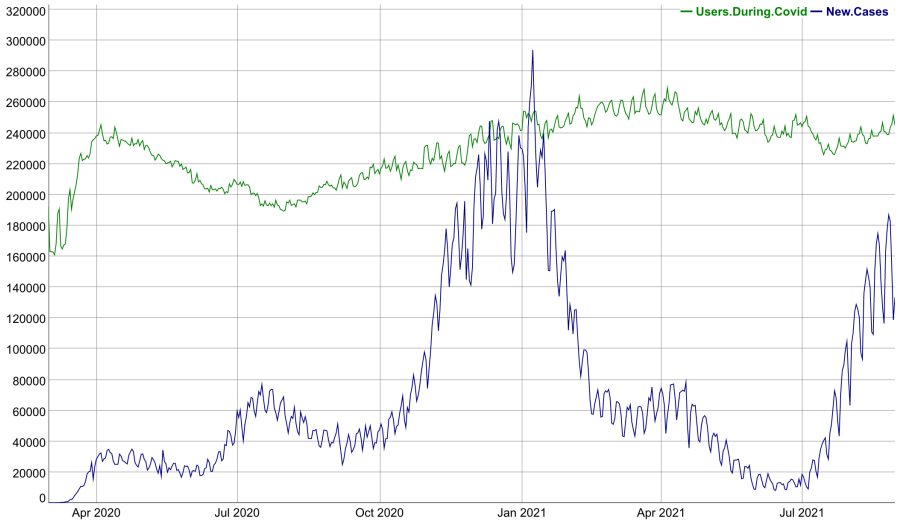
2.3.4 Hypothesis Tests 1,2 and 3 Interpretation

First, there is undeniably a significant increase in the number of Steam users following the pandemic. Both the number of overall daily users and the number of concurrent players have increased. This is apparent both from our hypothesis tests and graphically.



As our representative sample for players in the industry, it suggests that video gaming as a whole has had a dramatic increase in players. This is in line with our initial reasoning that more people turned to video games as a means of entertainment and/or social interaction during the pandemic.

When looking at the correlation, we reasoned that new cases of Covid-19 would translate to new players, as quarantined individuals are more likely to play video games. However, there is only a weak, negative correlation between new cases and change in player.



This suggests that the beginning of the pandemic introduced many new people to video games, but does not have a significant impact on maintaining those players.

2.3.5 Hypothesis Test 4:

The stock price for game companies has increased since Covid-19 began.

$$H_0 : \mu_0 = \mu_1$$
$$H_A : \mu_0 < \mu_1$$

H_0 represents our null hypothesis, H_A our alternative hypothesis, and μ_0 and μ_1 represents the mean of our gaming companies stock prices before Covid-19 and during Covid-19, respectively. Since both data sets have a few hundred data points, we apply the Central Limit Theorem and use a T-test to compare the two sets. We use a significance level $\alpha = 0.05$. We use a significance level $\alpha = 0.05$. If our p-value $p < 0.05$, we reject our null hypothesis.

$\bar{x}_1$	$\bar{x}_2$	t	df	α	p-value
379.5541	589.2899	-44.605	571.68	0.05	$<$ 2.2e-16

Figure 6: T-test results for stock prices

Since our p-value $p < 0.05$, we reject our null hypothesis.

2.3.6 Hypothesis Test 5:

The growth rate of stock prices for game companies has increased since Covid-19 began.

$$H_0 : \mu_0 = \mu_1$$
$$H_A : \mu_0 < \mu_1$$

H_0 represents our null hypothesis, H_A our alternative hypothesis, and μ_0 and μ_1 represents the mean of our gaming companies stock price growth rate before Covid-19 and during Covid-19, respectively. Since both data sets have a few hundred data points, we apply the Central Limit Theorem and use a T-test to compare the two sets. We use a significance level $\alpha = 0.05$. We use a significance level $\alpha = 0.05$. If our p-value $p < 0.05$, we reject our null hypothesis.

$\bar{x}_1$	$\bar{x}_2$	t	df	α	p-value
0.02836586	0.14360207	-0.96509	726.31	0.05	0.1674

Figure 7: T-test results for stock rate

Since our p-value $p > 0.05$, we fail to reject our null hypothesis.

2.3.7 Hypothesis Test 6:

The variance in the growth rate of stock prices for gaming companies changes during Covid-19.

To compare the variance of both data sets, we need to first determine if both sets are distributed normally. If they are we will proceed by F-Test. Otherwise we will proceed by Fligner-Killeen test.

We test this condition using the Shapiro-Wilk Normality test on both data sets. The null hypothesis is that the population is normally distributed. We use $\alpha = 0.05$, and if $p < 0.05$ we reject the null hypothesis.

Because our p-value for both sets $p < 0.05$, we reject our null hypothesis and proceed with Fligner-Killeen test.

W	α	p-value
0.96522	0.05	9.457e-08

Figure 8: Shapiro-Wilk results for change in stock price growth rate before Covid-19

W	α	p-value
0.90742	0.05	1.848e-14

Figure 9: Shapiro-Wilk results for change in stock price growth rate during Covid-19

$$H_0 : \sigma_0^2 = \sigma_1^2$$
$$H_A : \sigma_0^2 \neq \sigma_1^2$$

H_0 represents our null hypothesis, H_A our alternative hypothesis, and σ_0^2 and σ_1^2 represents the variance in our stock price growth before Covid-19 and during Covid-19, respectively.

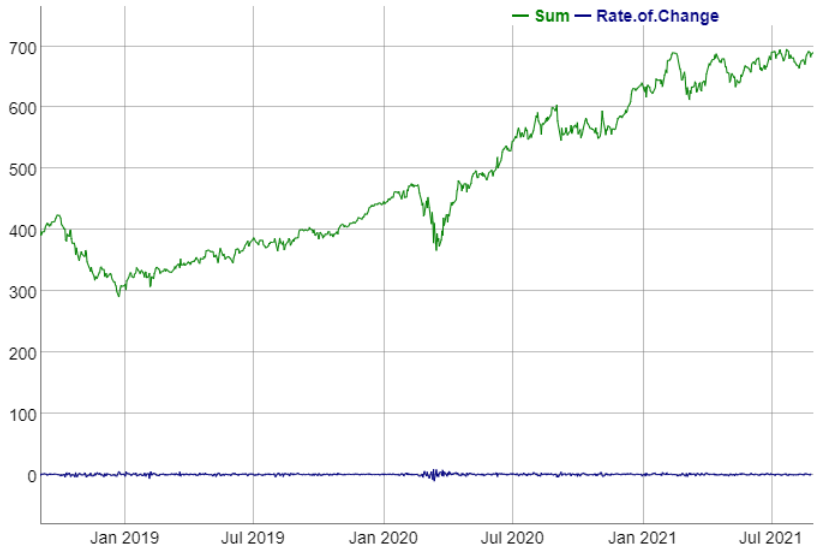
χ^2	df	α	p-value
1.4078	1	0.05	0.2354

Figure 10: Fligner-Killen results for change in stock price growth rate during Covid-19

Because our p-value for both sets $p > 0.05$, we fail to reject our null hypothesis.

2.3.8 Hypothesis Tests 4,5 and 6 Interpretation

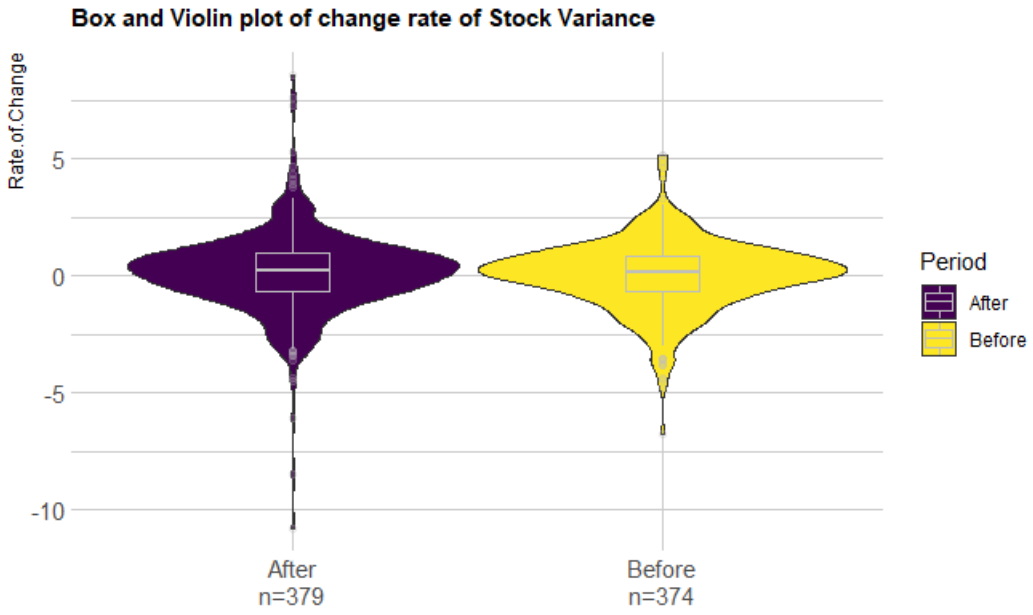
Our representative sample of video game companies has continued to grow since the pandemic began. However, the growth rate of the stock prices does not appear to have increased significantly. One possible interpretation of this is that player usage does not significantly affect stock prices. This seems unlikely, as a market is typically affected by its consumer interaction. Also, since we follow this trend for 18 months before and after March 1st, 2020, it is unlikely that the effect of increased players has yet to propagate into a change in stock prices.



Another more likely interpretation is that the significance of other negative events directly or indirectly linked to the pandemic - such as the ongoing microchip shortage - has countered the positive effects from the increase in the player base.

Since the growth rate of the gaming industry has not changed significantly, we tested the variance to see if the prices are more volatile. Our results show similar variance during Covid-19 as before, suggests little change in the volatility of the stock prices.

The lack of change in growth rate or volatility is in line with the Cowen reports prediction that the industry would be resistant to recession.



2.3.9 Hypothesis Test 7:

The short term growth rate of stock prices for game companies has changed since Covid-19 began.

$$H_0 : \mu_0 = \mu_1$$
$$H_A : \mu_0 \neq \mu_1$$

H_0 represents our null hypothesis, H_A our alternative hypothesis, and μ_0 and μ_1 represents the mean of our short term stock price growth rate before Covid-19 and during Covid-19, respectively. Since both data sets have a 30 data points, we apply the Central Limit Theorem and use a T-test to compare the two sets. We use a significance level $\alpha = 0.05$. If our p-value $p < 0.05$, we reject our null hypothesis.

$\bar{x}_1$	$\bar{x}_2$	t	df	α	p-value
-0.07013647	0.21875702	-0.3205	37.617	0.05	0.6248

Figure 11: T-test results for stock rate

Since our p-value $p > 0.05$, we fail to reject our null hypothesis.

2.3.10 Hypothesis Test 8:

The variance in the short term growth rate of stock prices for gaming companies changes during Covid-19.

To compare the variance of both data sets, we need to first determine if both sets are distributed normally. If they are we will proceed by F-Test. Otherwise we will proceed by Fligner-Killeen test.

We test this condition using the Shapiro-Wilk Normality test on both data sets. The null hypothesis is that the population is normally distributed. We use $\alpha = 0.05$, and if $p < 0.05$ we reject the null hypothesis.

W	α	p-value
0.94851	0.05	0.1542

Figure 12: Shapiro-Wilk results for change in stock price growth rate before Covid-19

W	α	p-value
0.97629	0.05	0.7206

Figure 13: Shapiro-Wilk results for change in stock price growth rate during Covid-19

Because our p-value for both sets $p > 0.05$, we fail to reject our null hypotheses and proceed with F-Test test.

$$H_0 : \sigma_0^2 = \sigma_1^2$$
$$H_A : \sigma_0^2 < \sigma_1^2$$

H_0 represents our null hypothesis, H_A our alternative hypothesis, and σ_0^2 and σ_1^2 represents the short term variance in our stock price growth before Covid-19 and during Covid-19, respectively.

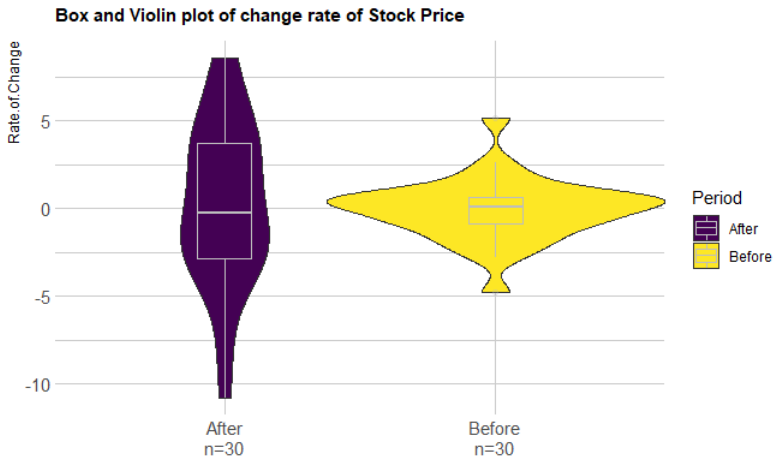
S_0^2	S_1^2	F	num df	denom df	α	p-value
3.215967	21.15861	6.5792	29	29	0.05	1.181e-06

Figure 14: F-Test results for change in stock price growth rate during Covid-19

Because our p-value $p < 0.05$, we reject our null hypothesis.

2.3.11 Hypothesis Tests 7 and 8 Interpretation

Our previous results showed little change to the long term trend of growth in the video game industry, but in the short term the following the advent of Covid-19 the effects were more significant. Like the long term, there was not a significant change in the growth rate of stock prices in the short term, but there was a significant increase in the variance. This increased fluctuation in price is evidence that the pandemic did have an effect on the video game stock market, but combined with the short and long term growth rate remaining relatively unchanged, it appears that Covid-19 has not had a strong impact on the industry. This provides further evidence supporting that the video game industry is resistant to recession.



3 Conclusion and Discussion

The increase in Steam players is a promising development in the video game industry. As our representative sample, it indicates total real players have increased since the pandemic began. Despite an apparent lack in change to the stock prices, it is likely that post-pandemic the industry will grow more dramatically as the negative impacts of Covid-19 dissipate while the number of players remains higher than before. Since there is only a weak correlation between players and Covid-19 cases, it seems likely that the trend of player growth will remain following the pandemic. This is significant because it indicates an increased interest and larger interaction

with the video game industry. Game developers who are aware of this trend would likely benefit from releasing new product or advertising prior products.

There was not much change in growth rate of stock prices, but this still bodes well for the video game industry. Similar to the 2001 and 2008-09 recessions, these results suggest that the industry is resistant to recession. This is significant because it provides further evidence that the video game industry will remain stable during recessions, and is thus a relatively safe area to invest in when worried about stock decline.

Steam is currently the most used platform for playing video games in the world. For this reason it was chosen as the source for our player data. However, Steam users only represent part of the gaming community, and future research should include more sources such as Playstation or Xbox players.

Covid-19 has had other effects on the industry. Further research will help determine the change in demographics among game users, such as gender and race distribution. Furthermore, what game genres are gaining popularity during the pandemic is another point of interest. Following these trends will help developers decided what audiences to target and help them determine what kind of game they want to build.

References

Barr, M., Copeland-Stewart, A. (2021). Playing Video Games During the COVID-19 Pandemic and Effects on Players' Well-Being. Games and Culture, 17(1), 122–139. https://doi.org/10.1177/15554120211017036

Creutz, D. (2020, 03 16). VIDEO GAMES AND COVID: IN WHICH WE STATE. Cowen: https://cowen.bluematrix.com/sellside/EmailDocViewer?encrypt=511852dd-fc31-49b8- 8870-277abe4f1526mime=pdfco=Cowen&id=georg.szalai@thr.comsource=mail

Nichols, C. N. (2021, June 8). Fact-check: Have one-third of US small businesses closed during pandemic? Austin American-Statesman. https://eu.statesman.com/story/news/politics/politifact/2021/06/08/kamala-harris-small-business-closures-covid-fact-check/7602531002/

Smith, S. M. (2021, June 10). Unemployment rises in 2020, as the country battles the COVID-19 pandemic : Monthly Labor Review: U.S. Bureau of Labor Statistics. U.S. Bureau

of Labor Statistics. https://www.bls.gov/opub/mlr/2021/article/unemployment-rises-in-2020-as-the-country-battles-the-covid-19-pandemic.htm

Statista. (2021, June 4). COVID-19: global video game and console sales increase as of March 2020. https://www.statista.com/statistics/1109979/video-game-console-sales-covid/

World Health Organization. (2020). Coronavirus disease 2019 (COVID-19): situation report, 105. World Health Organization. https://apps.who.int/iris/handle/10665/332059