Exploring Economic Shocks Through the Use of Statistical Computing

Luis Sebastian Weekes

Abstract

There are various types of economic shocks. There are shocks to price, supply, demand and pretty much any metric in economics that you can think about. Investopedia describes economic shocks as, "any change to fundamental macroeconomic variables or relationships that has a substantial effect on macroeconomic outcomes and measures of economic performance". In essence, an economic shock occurs as a short-term event that causes a disruption in trends that usually reverts back to a normal tendency. In this paper, we will be primarily concerned in exploring price shocks and consumptions shocks.

Introduction:

I will be exploring price shocks through the lens of the phenomenon called "rockets and feathers" otherwise known as asymmetric pricing. The rockets and feathers phenomenon occurs when prices in a market go up far quicker than the rate at which they restabilize. Apparently around two in three markets exhibit rockets and feathers behavior (Tappatta 2009). I will be looking to confirm the rockets and feathers phenomenon in the American oil market. I plan on confirming the presence of rockets and feathers phenomenon in the American Oil Market by focusing on historical periods of rapid price increase such as the oil embargo of the early 1970s and the Iraqi invasion of Kuwait in 1990. The way in which I will be investigating the rockets and feathers phenomenon is by separating the upswings and downswings in price during the price shocks and seeing if the upswings are on average bigger than the downswings for the given time period. Furthermore, I will be using the resampling technique of bootstrapping to bolster my results.

In addition, I will be exploring consumption shocks in the US economy. In particular, I am interested in finding the probability of having a consumption shock comparable to those during historical events such as economic recessions, the Covid Pandemic and random consumption shocks. In order to calculate the probabilities mentioned above, I used four probability distributions—a regular normal distribution and a double exponential or Laplace distribution, a modified Laplace distribution and a mixed Normal-Laplace (NL) distribution. All percent changed mentioned in the following paper are based on monthly observations.

Rockets and Feathers:

Procedures:

First off, I collected time series data on American Oil Prices from FRED. The data set that I worked with is the "Consumer Price Index for All Urban Consumers: Gasoline". Firstly, I loaded my data set into my environment and named my data frame "US_OIL_PRICES". I then mutated my data frame by renaming the oil price index to something more familiar to work with—"OIL_PRICE_BASE_1982_1984", and changing the format of the DATE variable into a form that the ggplot2 package could work with through the use of the lubridate function "mdy()". In order to get a grasp of the movement in oil prices, I plotted the index over time—this allowed me to pinpoint potential oil price shocks to examine. Furthermore, it was essential for me to convert the data into percent changes for a given month so that we could compare time periods. In order to do convert the index into percent changes from month to month, I needed to use the "diff()" function on the "OIL_PRICE_BASE_1982_1984" variable, and then divide by the index of the given month. While trying to perform the aforementioned task, I ran into the problem that I could not mutate a new variable into my data frame, because the differences added up to one less than the number of index values—there were only 669 differences for the 670 observations. In order to solve this issue, I saved my “US_OIL_PRICES” into a temporary data frame. I then loaded the differences of the oil price index from the temporary data frame into a new variable called “differences_in_CPI”. My next step was to create a vector called “diff_container” of size 700 to store these differences. I then loaded my “diff_container” with my “differences_in_CPI” variable using a for loop and added a 0 for the 670th index. I was then able to mutate a percent change in price variable using my “diff_container”.

The next step that I had to perform was to separate the positive percent changes in oil price and negative percent changes in oil price. I created another temporary data frame of the updated “US_OIL_PRICES” data frame that had the percent change in oil prices. Furthermore, I defined two empty vectors called “pos_diff_container” and “neg_diff_container” to store the appropriate values. I then ran the percent change in prices stored in the temporary dataframe through a for loop and saved the appropriate percent changes in price to the appropriate container. We could now mutate a positive and a negative percent change in price variable into our “US_OIL_PRICES” data frame.

In addition, I made two new data frames called “Positive_Percent_Changes” and “Negative_Percent_Changes” by copying the “US_OIL_PRICES” and selecting the appropriate percent change variables and filtering out the rows that had NAs. I could then use the summarize function to find the mean of the negative and positive percent changes in prices for the entire time period.

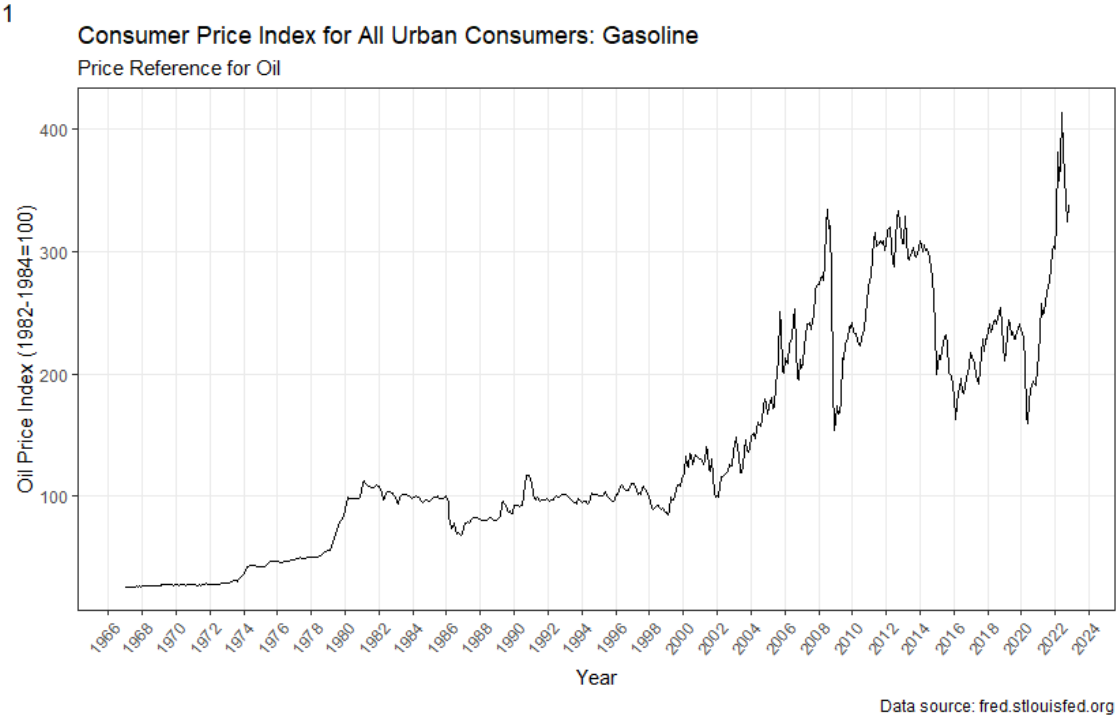
Moving along, I decided to examine two periods in time as mentioned prior. In order to do this, I created a new data frame “OilPrices_1972_sample” that copied the “US_OIL_PRICES” data frame, mutated a year variable by converting the date variable into a workable form and using the lubridate year() function to extract the year. Furthermore, I was then able to filter the appropriate years for the price shocks that I wanted to examine. I then plotted the oil price indices over time for the given time periods in order to get a better understanding of the movement in oil prices during the given time period. Following a similar procedure as mentioned above, I created a new data frame called

“Positive_Percent_Changes_1972_sample” that copied “OilPrices_1972_sample” and selected the positive percent changes variable and removed the NAs-this step was repeated for the negative percent changes. I was then able to take the mean of the upward movements in price by month and the downward movements in price by month by using the summarize function and the mean function. The process mentioned in this paragraph was repeated for the oil shock that occurred in 1990 due to the invasion in Kuwait.

Moreover, I bootstrapped the mean upward swings in price and downward swings in price because there are not many observations in either of the price shock samples. For instance, in the 1972 oil embargo sample there are only 43 observations and in the 1990 Kuwait invasion sample, there are only 36 observations. Furthermore, bootstrapping allowed me to somewhat simulate similar behavior as observed by the price shock time periods.

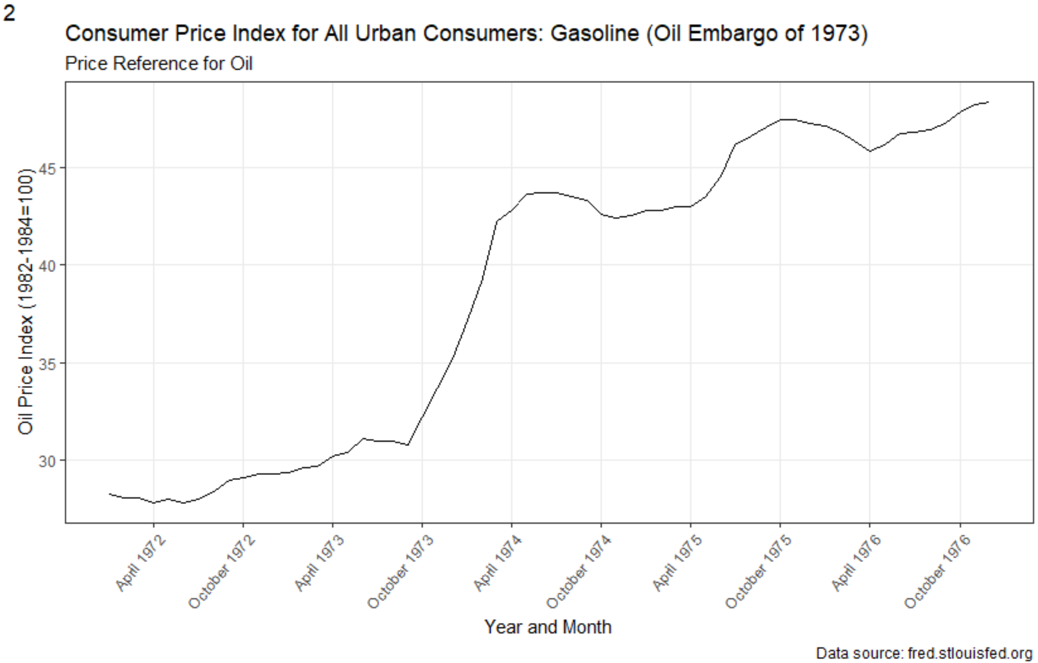
Results and Discussion:

Fig 1



As you can see from figure 1, there are sharp increases and stabilizations in prices around 1973 and 1991. In order to capture the entire time period from which the price shocks were occurring, I filtered the date ranges to be a little bit before and a little bit after the large spikes—to allow for the data taken to reflect a stabilization period. The time period that I filtered for the Oil Embargo data was from 1972 to 1976 and for the invasion of Kuwait I filtered data in the time period from 1990 to 1992. For the overall time period of Figure 1, the upward swings in oil prices were approximately 2.85% on average, while the downward swings in oil prices were approximately -2.6% on average. This makes sense because if you combine the aforementioned finding with the fact that there have been more months with price increases, it will confirm that the price of oil has been increasing over time. The average price swings for the entire time period of figure 1 will also give us a comparison point for the swings in price during our selected time periods.

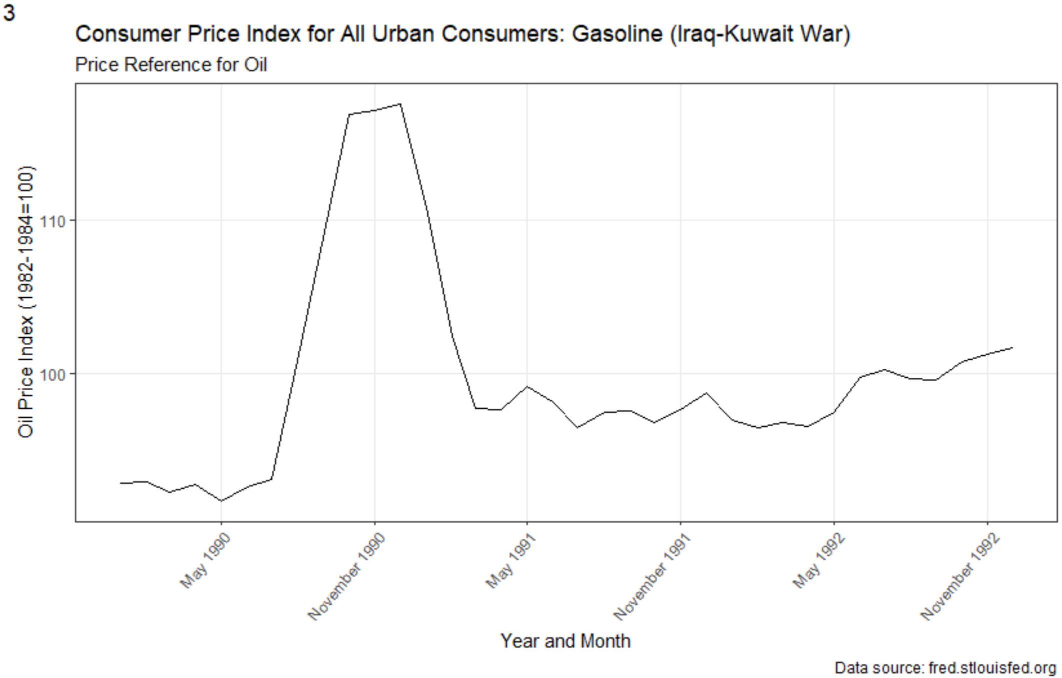
Fig 2



Looking at figure 2, it is clear that there is a shock in prices around late September in 1972—furthermore after around march of 1974 oil prices seem to revert to a similar trend as to before the shock. On average, for this time period, prices increased by 1.66% by month and decreased by 0.71% by month.

Furthermore, our boot strapped estimates exhibit an average price increase per month of 1.67% and an average price decrease per month of 0.71%. These results are in accordance with the hypothesis of rockets and feathers, because on average, prices increased at a faster rate than which they decreased. Unfortunately, this time period might not be a great case study because in figure 2 you can see that prices never came back down to a level comparable to that before the shock. That being said, prices did go up quicker on average than they fell.

Fig 3



Upon examining figure 3, it is clear to see that there is a clear jump in price around the summer of 1990. This is around when Iraq invaded Kuwait which cut the world supply of oil significantly--causing prices to jump starkly. In contrast to figure 2, figure 3 does portray that price comes back down to a similar price level comparable to before the price shock. Therefore, studying this time periods' price shock might be beneficial. On average, for this time period, prices increased by 1.78% by month and decreased by 1.79% by month. Furthermore, our boot strapped estimates exhibit an average price increase per month of 1.76% and an average price decrease per month of 1.82%. These results contradict the hypothesis of rockets and feathers, because on average, prices increased at about the same rate at which they decreased.

Consumption Shock Probabilities:

Procedure:

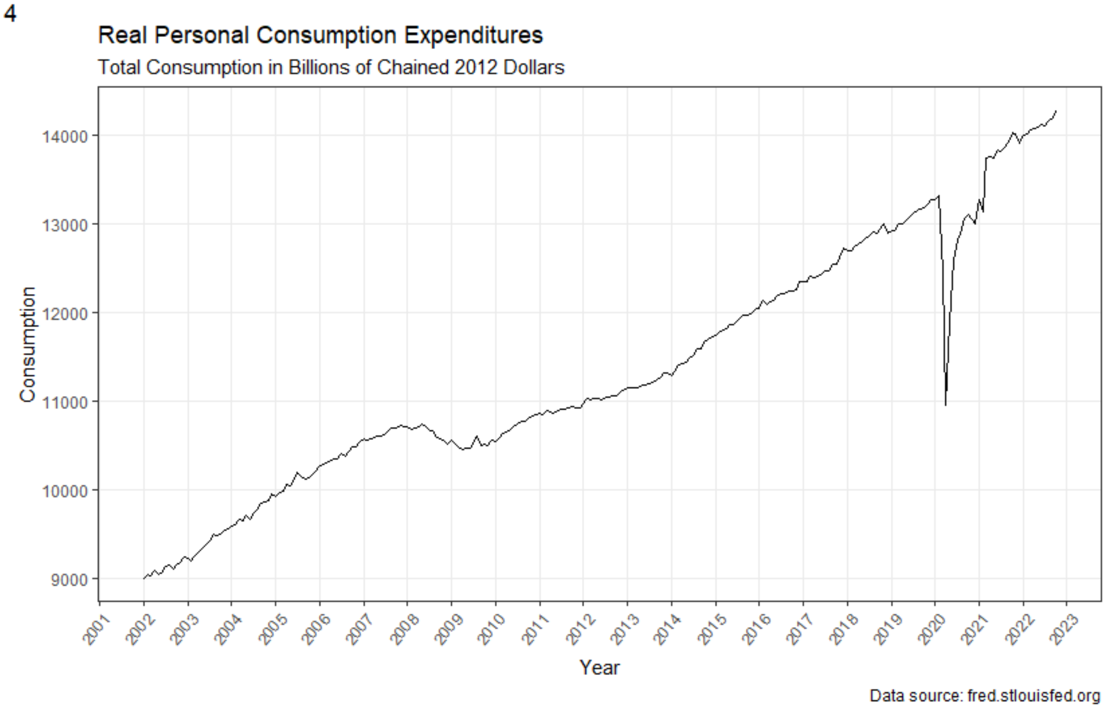
I obtained my data from the “Real Personal Consumption Expenditures” time series data set on FRED. I then loaded my data into R studio and named the data frame “Consumption”. Furthermore, I used the lubridate function to mutate the Date variable into a form suitable for the ggplot2 package. I then plotted the Consumption figures in order to get an idea of where there were large movements in consumption. Next, I loaded my data into a temporary data frame in order to take the differences between each of the months, I then loaded the differences into a vector to add zero to the last month so that no errors would occur in attempted data frame mutations due to mismatching sizes of vectors. My next step was to create a new data frame called “Consumption_pct_change” where I copied the “Consumption” data frame and created a new variable called “Percent_change_in_cons” by using the vector that contained the differences in consumption by month and dividing them by the total consumption for that month. In order to get a grasp of the distribution that I would be working with, I plotted the percent change in consumption by month using ggplot2’s geom_density() function. I also used the fav_stats() function in order to know the mean, standard deviation and quantiles of the distribution. In order to get an even better grasp of the distribution in consumption shocks, I filtered out the outliers and replotted the density curve. Furthermore, I checked the qqnorm() and qqline() plots for the consumption shocks to check normality.

When it came to simulating the probability of getting a monthly consumption shock larger than during a given event, I followed the following procedure. I created a logical vector the size of how many random numbers I wanted generate. I then used a for loop with 1000 iterations to generate 1000 numbers from a normal distribution with the mean as .001916097 and standard deviation as 0.01177263. Inside the loop, I used an ifelse statement to check whether the randomly generated number was greater than the consumption shock that we were concerned with. If the random number was greater, true was assigned to the index of the logical vector we declared earlier and false if not. I then took the proportion of trues against falses to estimate the probability of having a monthly consumption shock larger than during a given event.

I underwent the process described in the paragraph above for the three aforementioned events (Great recession, Covid pandemic and a randomly generated shock) using the normal distribution and thrice using the Laplace distribution, modified Laplace distribution and NL distribution for the three different events. The monthly shock sizes that I will be examining are: -6.41% for the Covid pandemic, -.92% for the great recession and 0.23% respectively. I have chosen to simulate the probabilities using the Laplace distribution, because in economic literature economists have used exponential distribution to describe the amplitude of economic growth shocks (Stockhammer and Oller 2011). Furthermore, Laplace distributions are used in particular in order to account for both negative and positive shocks (Stockhammer and Oller 2011). Moreover, I increased the Laplace's scale parameter by a large factor in order to see if increasing the amount of variability would lead to more realistic results--increasing the scale parameter allowed for a slightly better fit to the real-world data. Lastly, In Stockhammar and Oller's paper "On the Probability Distribution of Economic Growth (2011)" the writers mention that mixing a normal distribution with an asymmetric Laplace might yield the best results. Given that the data didn't seem to be as obviously asymmetric as in Stockhammar and Oller's paper, I decided to go ahead and simulate probabilities using a normal distribution and the modified Laplace distribution with weights of 0.2 and 0.8 respectively. Lastly, I set the seed and used a uniform distribution to generate a random consumption shock that could happen any month.

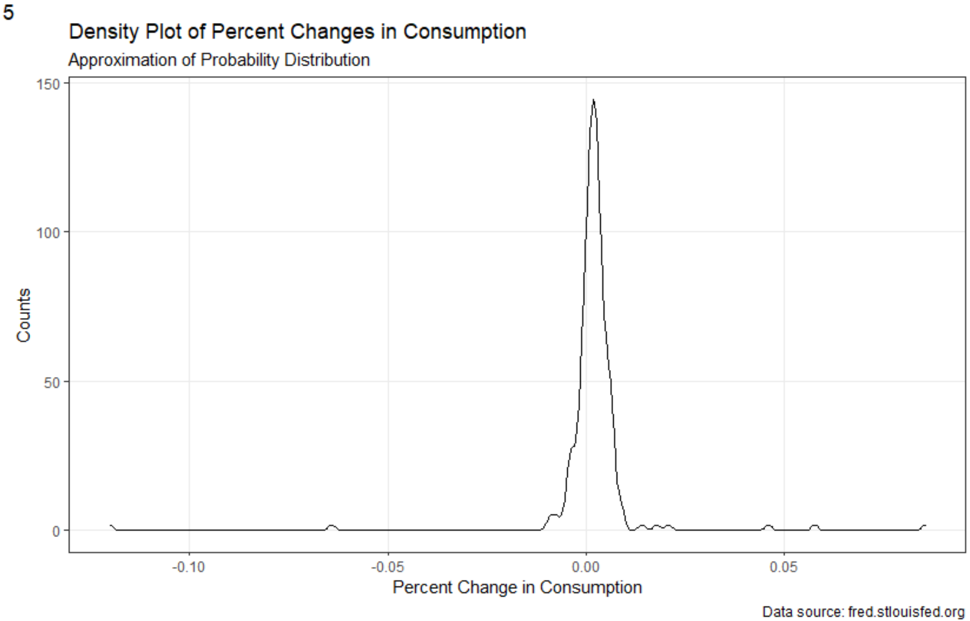
Results and Discussion:

Fig 4



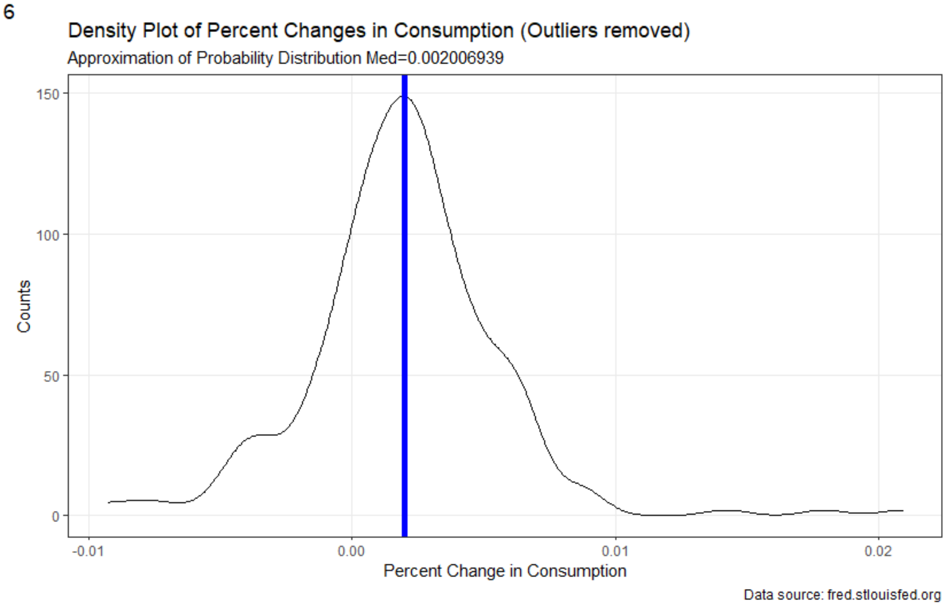
Looking at figure 4, you can see that there are periods of large negative consumption shocks in the years 2008-2009 and incredibly large negative consumption shocks in 2020. These consumption shocks occurred due to the financial recession in 2008-2009 and the Corona virus pandemic. Therefore, I would like to find the probability of observing a monthly consumption shock larger than observed during the mentioned time periods. I chose the months 2009-08-01 and 2020-02-01 to represent the events of interest.

Fig 5



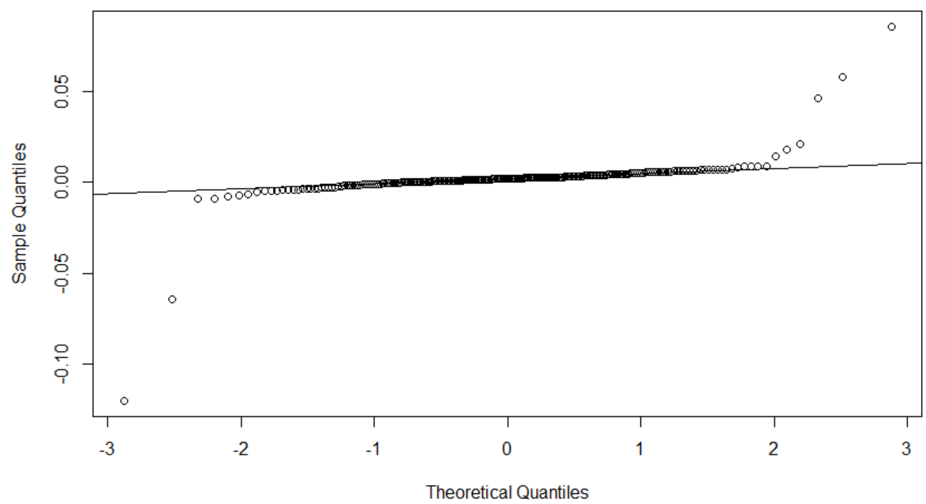
We can tell that figure 5 looks fairly normally distributed, but we do not have a great view of the distribution due to outliers.

Fig 6



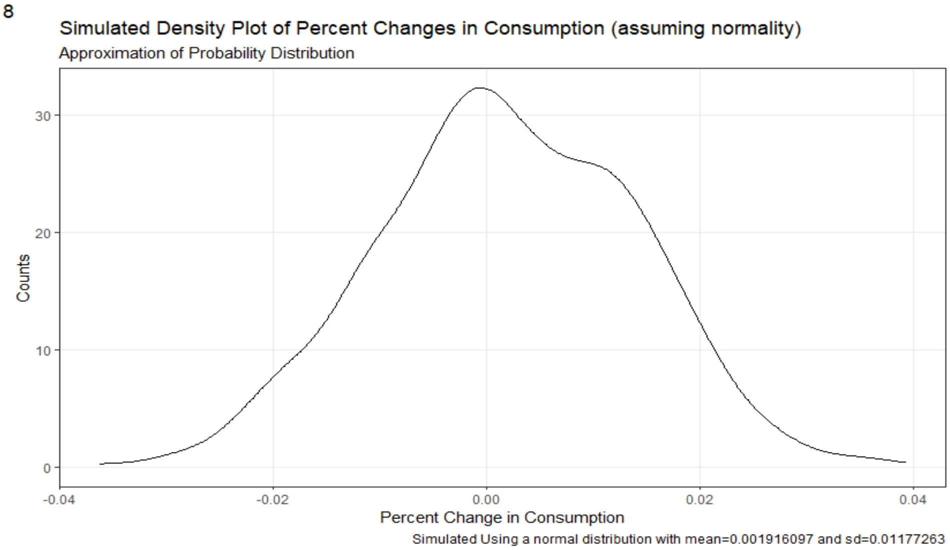
Upon examining the density plot of percent changes in monthly consumption, we can now see that the probability distribution is somewhat normal and fairly symmetrical.

Figure 7: Normal Q-Q Plot



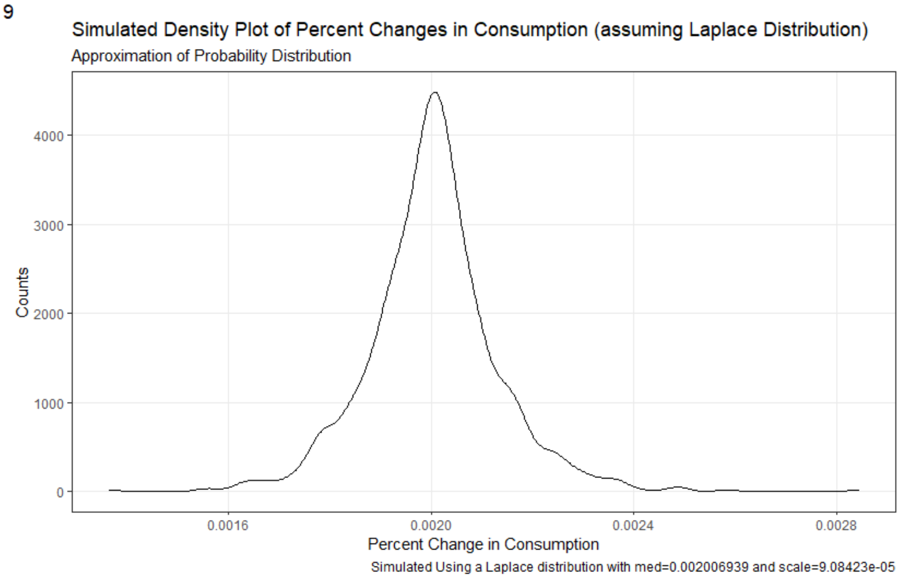
Upon examining figure 7, we see that the data follows a very normal trend in the center but has strange behavior at the tails. Therefore, for my first simulation, I will generate probabilities using a normal distribution.

Fig 8



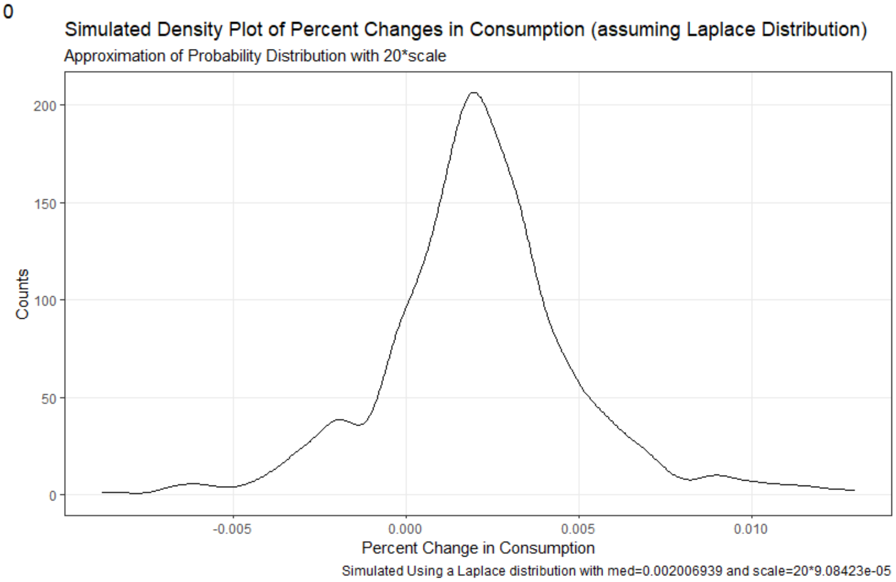
Upon comparing figure 8 to figure 6, we can see that the normal distribution is much wider than the data generated distribution which may lead to inaccuracies, but it is a good starting point. While simulating the probability of getting a monthly consumption shock larger than during the Covid Pandemic using a normal distribution, I got a result of 0. This makes sense because we have never had an event like this happen in modern history. For the great recession, I generated a probability of getting a greater consumption shock in a given month to be 0.245, which seems to be pretty high considering that the great recession is deemed to have been a rare event. For observing a consumption shock larger than the randomly generated consumption shock, I generated a probability of 0.493.

Fig 9



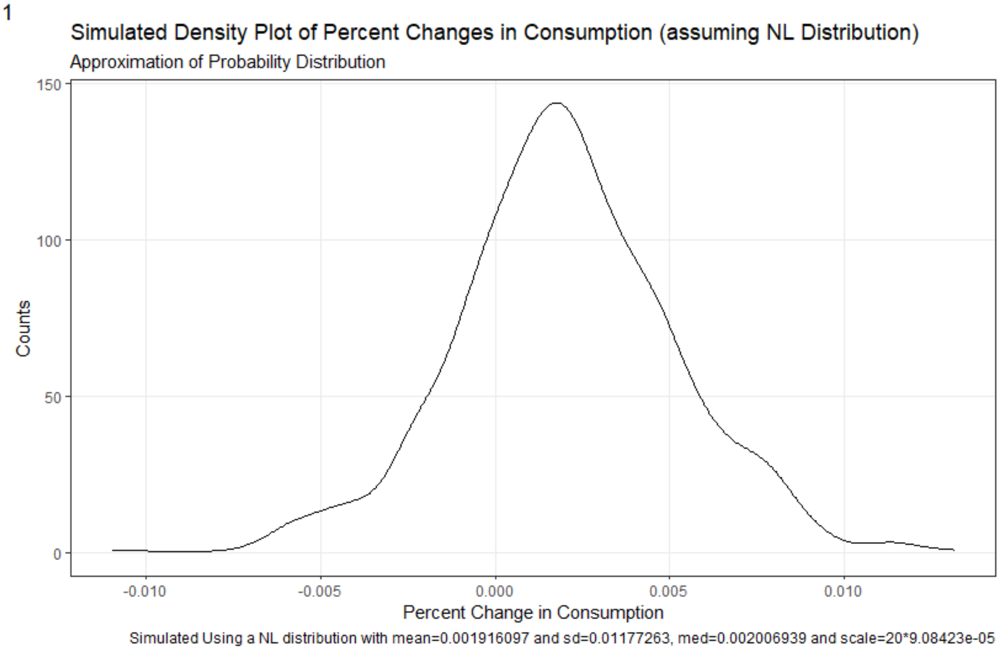
In Figure 9, we can see that the Laplace distribution follows a more similar shape to the data, but the density is too concentrated around the median which may lead to inaccurate results as well. While simulating the probability of getting a monthly consumption shock larger than during the Covid Pandemic using a Laplace distribution, I got a result of 0. Again, this makes sense. For the great recession, we generated a probability of getting a greater consumption shock in a given month to be 0. This probability seems to be inaccurate given the fact that the data accounts for time periods such as the great recession and more. For observing a consumption shock larger than the randomly generated consumption shock, I generated a probability of 0.014. This too seems to be strikingly too low since this probability should have represented the probability of a regular monthly consumption shock. Therefore, I decided to examine the same probabilities with a much larger scale parameter.

Fig 10



When comparing figure 10 to figure 6, we see that the Laplace distribution with 20 times the scaling factor seems to be more realistic than the previous Laplace distribution and normal distribution, but the peak still seems to be a little bit too high and the density is somewhat still too concentrated. While simulating the probability of getting a monthly consumption shock larger than during the Covid Pandemic using a Laplace distribution with the scale multiplied by 20, I got a result of 0. Again, this makes sense. For the great recession, we generated a probability of getting a greater consumption shock in a given month to be 0.008. This probability seems to be possible given that monthly consumption shocks that occurred during the great recession were very rare, but not unheard of in recent history. For observing a consumption shock larger than the randomly generated consumption shock, I generated a probability of 0.416. This probability makes sense, because if you generate a random consumption shock, you expect it to land somewhere around the middle of the probability distribution.

Fig 11



Comparing figure 11 to figure 6, we can see that we get similar peaks and corners in the distributions, therefore, I believe that the NL distribution might generate more accurate results than using the modified Laplace distribution. The weighting of the distribution is 0.2 normal and 0.8 Laplace. While simulating the probability of getting a monthly consumption shock larger than during the Covid Pandemic using an NL distribution, I got a result of 0. Again, this makes sense. For the great recession, we generated a probability of getting a greater consumption shock in a given month to be 0.018. This probability seems to be possible given that monthly consumption shocks that occurred during the great recession were very rare, but not unheard of in recent history. The aforementioned probability doubled compared to when it was generated using the modified Laplace distribution. For observing a consumption shock larger than the randomly generated consumption shock, I generated a probability of 0.449. Again, this probability makes sense, because if you generate a random consumption shock, you expect it to land somewhere around the middle of the probability distribution.

Conclusion:

Investigating both parts of this paper lead to some interesting results and raised other questions. For the Rocket's and Feathers part of this paper, it was found that the oil embargo event of the early 1970s might have confirmed the phenomenon but might have also not met the criteria that we were looking for—prices never came back down. In contrast, the Iraq-Kuwait war price shock on oil did not confirm the Rockets and Feathers hypothesis. The mean upswing and down swing in prices for both price shocks were much lower in value than the average upswing and down swing for the entire data set; this is because prices varied extensively after the year 2000. This might be the case, because after 2000, there were 3 major economic recessions that could have greatly affected gas prices. Furthermore, deregulation was on the rise after 2000. A additional step would be to compare an analogous prices shock for a small economy and a large economy to see whether it takes longer in a small economy for prices to stabilize due to there being less competition.

When simulating the probability of observing a larger consumption shock than during an event, it was shown how important it is to choose a reasonable probability distribution and spend time justifying it. Two glaring examples prove that choosing the right distribution is important. When I calculated the probability of getting a monthly consumption shock larger than a monthly consumption shock during the great recession while using a normal distribution, a probability of 0.245 was generated. It is not very likely that ¼ and the great recession go together in many sentences. In addition, we got very miniscule numbers across the board under the Laplace distribution. Lastly, one major limitation in this exercise is that the data only started in 2002, and hence we are not fully accounting for a large number of historical events. Furthermore, this exercise has limitations in its predictability capacity, because future events are independent from past events. That being said, another interesting addition to this exercise would be to use the studied distributions to stress test the potential gains and losses in portfolio returns.

Citations:

Stockhammar, Par, and Lars-Erik Oller. “On the Probability Distribution of Economic Growth.” *Journal of Applied Statistics*, Sept. 2011, https://doi.org/10.1080/02664763.2010.545110.

Mariano Tappata “Rockets and Feathers: Understanding Asymmetric Pricing.” *The Rand Journal of Economics*, winter 2009