Internet Spaceships are Serious Business:

Econophysics of EVE Online

Jonathan Wurtz

5/5/2015

Abstract

In this paper the economic system of the on-line multiplayer game “EVE Online” is analyzed and compared with real economic systems. Using a publicly-available dataset of asks and bids in the primary in-game market hubs, timeseries of items were created to answer the fundamental question “How is this system the same and different from real markets?” Two methods are used: The first inspects the power law returns of different items, returning a characteristic power law across about an order of magnitude. The second inspects the efficient market hypothesis via autocorrelations, and finds correlated fluctuations below the scale size set by a transaction tax.

Introduction

In this report, the novel economic system of the on-line multiplayer game EVE Online will be compared to real-life markets. This report will first discuss some background of the game, focusing primarily on its market and economic aspects. Then the data set will be introduced, and two results will be presented: first, power law returns for price fluctuations compared to the stylized fact $P(R) \sim R^{-2}$, the probability scales as the returns to the -2nd power. Second, the efficient market hypothesis will be put to the test in the system, by way of autocorrelations of price fluctuations. Finally, this report will conclude with a discussion of results as well as possibilities for future work.

System Introduction: EVE Online

EVE Online is a Massively Multiplayer Online game, established in 2004. The game itself is multi-faceted, with normally around 30,000 players online on the games single server. The nature of the game is very emergent and player-driven; as such a strong economy and trading system has emerged. However, it should be emphasized that the trading system is not the primary focus of the game, and as such can be influenced by different events (for example, major player alliances declaring war). Daily there are between 1-2m transactions that occur across a range of about 10,000 different items. The market itself is primarily a commodities market and acts mainly as a middle-man to support otherwise impossibly large manufacturing or logistical chains necessary to build items to interact with the game world or other players.

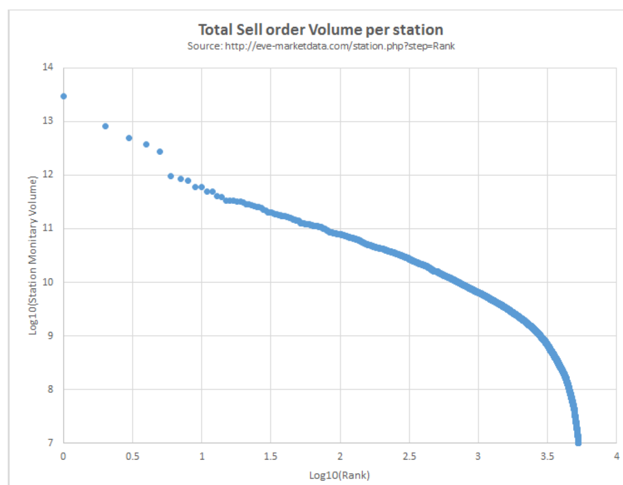

Figure 1: Zipf plot of total item volume per station, on a log-log plot showing a power law across about 4 orders of magnitude, characteristic of a scale free network.

The game world itself consists of about 2,000 different "star systems" interconnect by a locally connected network; inter-system trading interactions between players are limited to communication and physically moving items. As such, there emerged in 2005-2006 localized "trade hubs" that have grown to conduct a large portion of all in-game trades. The study of the emergence of these trade hubs was one of the first intentions of this project; however due to the historical nature and size of the data it was unfortunately unfeasible and was scrapped. A zipf plot of total market volume per system is shown in figure 1. The top-ranked system, "Jita", has about an order of magnitude more market volume then there would otherwise be via a characteristic power law, suggesting some other mechanism causing trade hub emergence. Because trade is localized to five main trade hubs, it is possible to look at the time series of different items at those trade hubs independently and glean information about the nature of the economic system. All data analysis done in this project is confined to the trades done in the five main trade hubs, primarily the top one.

There is a tax of about 0.75% of the total value of an order to create an order, on top of about 0.75% for both parties to complete an order; at least 2.25% of the total monetary value is always inherently lost in any transaction. This mechanic creates a natural money sink which curbs inflation, which otherwise historically has been around 15-20% per year[1].

It should be emphasized that the game is very rich and emergent, and this overview's intent is primarily to give context to the system in question. I have personally played this game for about 5 years and as such understand some of the subtleties of the system, which can easily be papers in and of themselves.

Data

The trading system in EVE Online is fundamentally the same as most stock markets. If an agent has an item and wants to sell it, they can do one of two things. They can list the item on the market as a gives order, and wait some time for another agent to purchase it. Or, they can instantaneously sell to another agent who put up an asks order, normally for a slightly decreased price. Equivalently, if an agent is willing to buy an item, they can either immediately purchase from a gives order, or put up an asks order and wait some time for another agent to instantaneously sell. As such, for each different item and at each time there are a list of asks and gives orders (anywhere from 10-20 for less popular items to many more) ranked by their price; an agent automatically purchases or sells to the best price available. An example of some asks/gives data is shown to the right in Figure 2.

Gives Orders		Asks Orders	
Price	Amount	Price	Amount
5.22	3,894,006	5.21	150,174,909
5.23	126,995,273	5.19	99,211,625
5.23	169,217,143	5.18	43,263,110
5.23	2,028,399	5.17	68,089,718
5.24	120,474,052	5.16	99,999,999
5.24	73,496,128	5.15	1,003,896,750
5.24	182,776,681	5.12	42,674,309
5.24	5,970,436	5.11	10,000,000
5.25	200,000,000	5.1	999,300,985
5.25	32,036,788	5.09	830,782,857
5.25	4,747,702	5.08	423,422,981
5.25	90,634,351	5.07	250,000,000
5.25	149,957,993	5.07	1,000,000,000
5.25	384,268,147	5.06	340,588,205
5.26	67,130,239	5.06	200,000,000
5.26	4,157,525	5.06	5,000,000
5.27	64,416,426	5.05	200,000,000
5.28	120,474,052	5.02	250,815,164
5.29	10,782,130	5.01	20,000,000

Figure 2: Partial data for item "Tritanium" in Jita on 5/3/2015.

The data used for this analysis is data dumps from a third-party website[2] which uses cache scraping from volunteer players to pull data about ask and gives lists for different items. However, because this resource is third party it is inherently incomplete; more popular items and locations are more heavily sampled then less popular ones. Even with these restrictions the data is very non-trivial; the more popular items have about 4-5 samples per hour per hub, with the most popular item having about 40 samples per hour per hub. The service has been running since 2006. The basic data structure is appended data dumps of different items' ask and gives lists for different locations. Each day has about 300MB of data. Data was acquired in a 300 day region from May 2014 to March 2015, limited primarily by slow download and compiling speed. Each list of ask and gives orders was compressed into the best ask and gives price, the number of items listed to gives, and the monetary capital listed on asks. In total the compression yielded individual data vectors of:

[1] Based on price increases of PLEX, which is assumed to have approximately equal value across time.
[2] http://www.eve-central.com

(Item, Time, Location, Buy Price, Buy Volume, Sell Price, Sell Volume)

As mentioned above, only the 5 trade hubs were included in order to better compress data. Furthermore, only the top ~1,000 most popular items were included, in order to simplify the database back-ends (which utilized SQL). Each day had about 150,000 entries, or about 45m points over the 300 day period.

The individual data vectors for every item and location was then converted into a time series. Each hour was assigned a bin, and each data vector was respectively filled into each bin. Blank bins were assigned the values of the backwards-closest filled data bin. For more popular items, this meant that the entire time series was well-sampled with 4-5 samples per hour; however less popular items had artificial periods of zero hourly returns per hour due to under sampling. Some sample time series of more some more popular items are shown to the right in figure 3. Note the close ask-bid spread, as well as the occasional sharp spikes which are due to a sudden transient lack of supply or demand in the system.

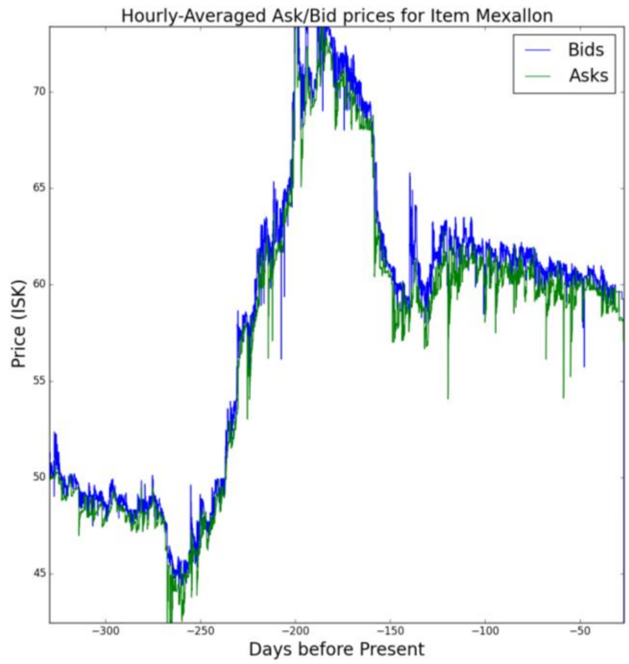

Figure 3: Bid and Ask time series for “Mexallon”, an equivalent to Steel or other basic building material.

Results

Power Law Returns

An empirical law found in many economic data sets is the power law nature of returns for different items on a market. For small returns, the price of a stock is assumed to be an uncorrelated random walk; however, there are sometimes large jumps in a market which are part of a strongly non-Gaussian nature. It is found that the probability of a return as a function of the magnitude of that return is a function of a power law, primarily $P(R){\sim}R^{-2}$, where R is the return (normally the log return) and P is the probability. A similar plot could be found by aggregating all of the time series data of this set. 1073 different items in the main trade hub were used to find 24 hour returns at a 1 hour resolution over a 300 day time period (a total of 9.3m points). Returns were normalized to be of unit variance, and histogrammed into linearly spaced bins to create a probability density function.

A 24 hour window was used over an hourly one because of two reasons. First, not all data is well-sampled and thus a wider sample window allows for less worry about the incomplete data set. Second, a 24h window is a “natural” frequency for the system; agents in the system are simply players and as such are not avid high-frequency traders (perhaps during the day!). Each agent would nominally interact with the system for a few hours per day, for example, at night after work. The result is shown below in Figure 4.

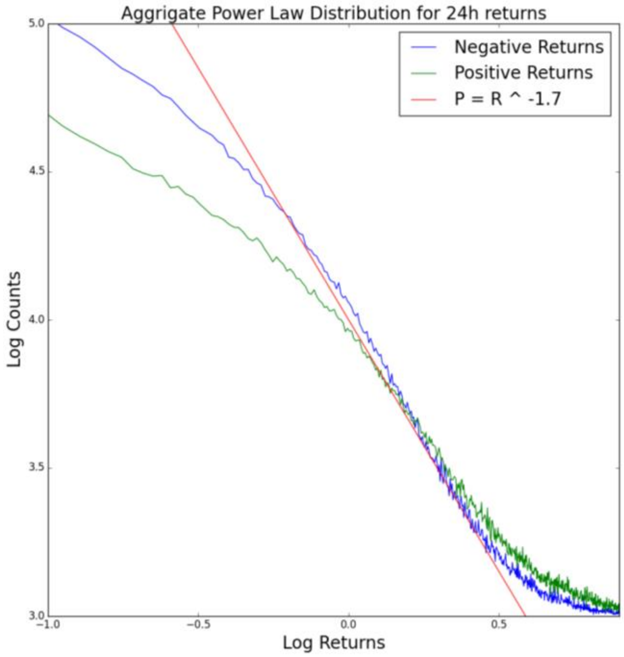

Figure 4: Aggregate power law distribution of price returns in the primary trade hub. Red is a line attempting to match the most linear part of the data.

The first clear result of the aggregate histogram is the apparent non-linearity in the PDF, suggesting that this market system may not be completely scale free. At low returns the distribution is below its power law expectation; however, this may be in the random walk region which instead of having a characteristic linear shape would instead have a shape similar to that of Figure 4. At high returns the system also diverges from the power law regime. This may have several possible reasons. The primary reason could be the finite nature of the data set, coupled with the 24-hour returns with a 1-hour resolution. If there is an anomalous jump from one price to another on a timescale smaller than 24 hours, then there would be about 24 returns which have a similar and large return. This would mean that singular events that cause large jumps in price would be over-represented in this plot, as we see here. In fact, In the limiting case for returns approaching infinity (Numerically, 16 standard deviations) the counts approach a constant of about 2,000, independent of return power, suggesting there may be about 2,000 of these "jump" events in the 1073 items in 360 days.

Another possible reason might be due to external actions characteristic to this system—for example, large-scale price manipulation by groups of players is completely legal within the game rules, and might explain the high-return data. This is one of the areas that research into this system may be fruitful, as it may be unique to this particular system but have characteristics generalized to real-world systems.

Autocorrelations of Returns

An axiom, or at least strong assumption, in economics and econophysics is the efficient market hypothesis. This hypothesis states that any market quickly reaches an equilibrium where no "free money" can be made via arbitrage. There are several consequences of this assumption, which are strongly backed by stylized facts in data. The primary consequence is that the autocorrelation of returns for a stock is zero for non-zero time lags. If there was a non-zero autocorrelation, that would imply an agent could look at the historical price of the stock and predict if its price would rise or fall; of course, that agent would then take advantage of the knowledge and trade in that stock, changing the price to where there would be no profit opportunity.

The question is then to ask if the autocorrelation of different items are zero for non-zero time lags. Data for different groups of items (For example, the group of all "Basic Minerals", which may correspond to the group of all fast-food corporation stocks) was used to calculate the 48-hour returns normalized to a standard deviation of one. Item groups were used to compress data onto the same graph, because they should have theoretically similar behavior. Some example data is shown below in figure 5[a,b].

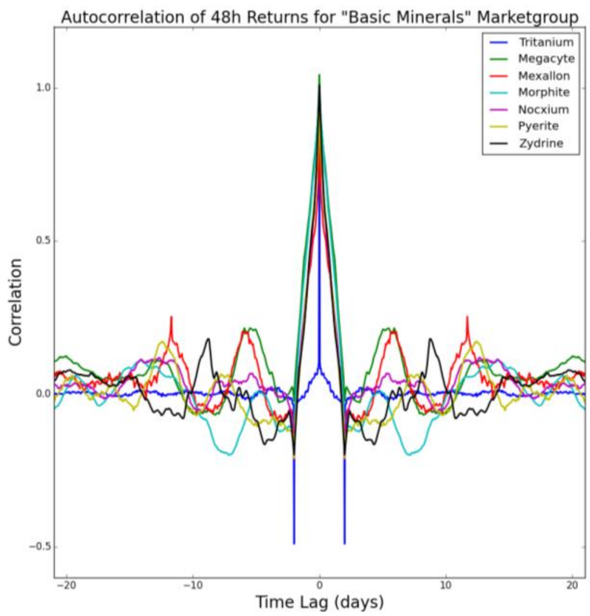

Figure 5a: Autocorrelation for "Basic Minerals" Market group, equivalent to basic commodities such as steel or concrete.

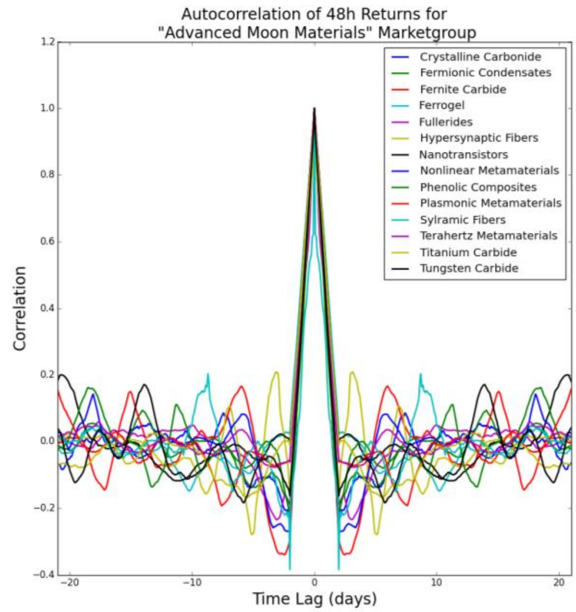

Figure 5b: Autocorrelation for "Advanced Moon Materials" Market group, equivalent to refined goods such as silicon or textiles.

It is clear that there is a consistent non-zero correlation for non-zero lags. There is a natural "blurring" of 48 hours corresponding to the 48 hour returns with a 1 hour resolution; the central peak has a width of 2 days. However, for lags greater than 48 hours, there are significant non-zero correlations, consistent across different market groups (these two groups were picked due to high data density). The left-hand correlation for "Tritanium" is anomalous because of an event in late 2014 which caused prices to increase 20% over the course of a few hours, with the price spiking by almost an order of magnitude for an hour, causing anomalous renormalization of the returns to unit variance.

As is mentioned above, a non-zero correlation for non-zero lags is a signature of an inefficient market. However, this is not necessarily the case. In this market system, there is a tax overhead which comes out to be about 1.5% of the total cost to complete the transaction; in order to take a buy-and-hold approach the returns of the item must thus be more then ~3.0% to make a profit. This means that the market is not scale free—there is some region where there are allowed to be correlated fluctuations, but outside that region those fluctuations must be small. One must then look at the average magnitude of fluctuations and ask if the fluctuations are in fact below the characteristic scale size.

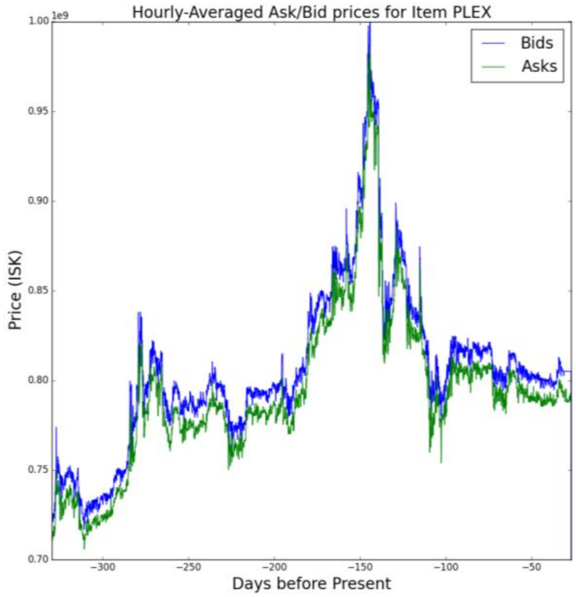

Figure 6a

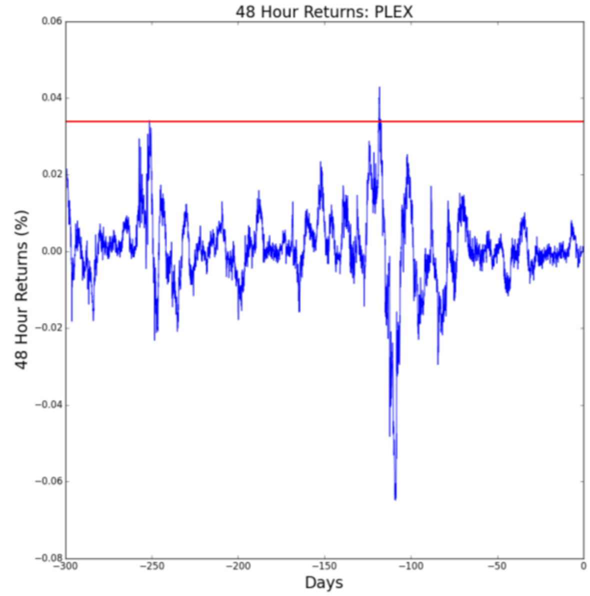

Figure 6b

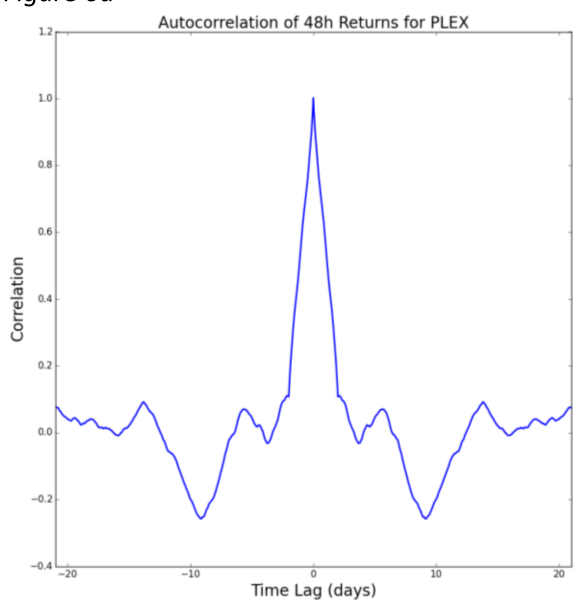

Figure 6c

Figure 6

a. *Timeseries for item "PLEX", which can be tied to real-life cash value. As such, PLEX can be used as an investing item and an index for in-game monetary value. Note the bubble at -150days.*
b. *48 hour returns for Bids. The blue line represents the cutoff for a buy-and-hold trade strategy.*
c. *Autocorrelation for 48h returns. There is a strong negative correlation at about -8 days, which may hint at a weekly trading cycle; however this may also be an artefact of the before mentioned bubble.*

As can be seen with the example item, the fluctuations of returns is primarily below the scale size imposed by taxes, and for this item only becomes profitable 2 or 3 times in a 300-day period. This is consistent across different items; this item was picked because of its high data density (around 40 samples per hour) due to its popularity as an investment item. The magnitude of the fluctuations means that the transaction tax creates a scale size in the system such that fluctuations in the system can be correlated, as long as the fluctuations are smaller than the transaction tax. A subject of further study is how a transaction tax effects generalized economic systems.

Conclusion

In this report, two findings were discovered in the economic system of EVE Online. The first finding was the empirical power-law distribution of returns, with divergences at low and high returns. The low return limit was concluded to be a crossover to the correlated gaussian random walk, as the returns are possibly below the scale size of the transaction tax. The high return limit was concluded to be an outstanding question, with the possibility of it being either 1) an artifact due to the finite data set coupled with oversampling of returns, or 2) Some secondary mechanism such as individuals or groups manipulating the market, causing the ~2,000 anomalous events that would be poorly represented with this data reduction method.

The second finding was the autocorrelations of returns for items on the market. It was found that there was a non-zero autocorrelation at nonzero lags, indicative of a market inefficiency. However, due to the transaction tax inherent to the system, price fluctuations under this scale could be correlated. For most items, most price fluctuations were under this scale size.

Further research can go in several directions. One of the primary goals of the project was to examine the emergence of market hubs in 2005, but was limited by the historical nature of the data. Another question may be inspecting the source of high-return events that are anomalous with a power law. Another question to be asked could be inspecting the relationships between the volume, buy price, and sell price time series; the result may be a standard example of supply-and-demand economics, where with reduced volume the price increases and vice versa. A final question is more general: In what other ways does this system behave like a real economic system, and more interestingly, where does it deviate? EVE Online has potential for having a rich and unique data set which acts as a “pocket model” for real economic systems. It is hoped that research on this subject will continue, especially on the question of comparing this system to real life.