

High Frequency Trading: Assignment 1

Matthew Dicks (DCKMAT004)

University of Cape Town, Department of Statistical Sciences



① Visualizing Trading

Naspers - NPN

Anglo American - AGL

② Order-flow Auto-correlation

Naspers - NPN

Anglo American - AGL

③ Inter-arrival Times

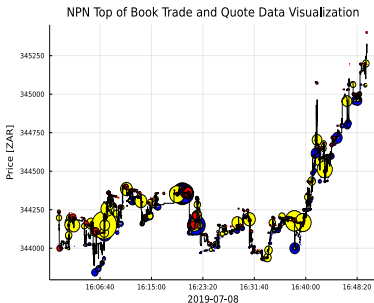
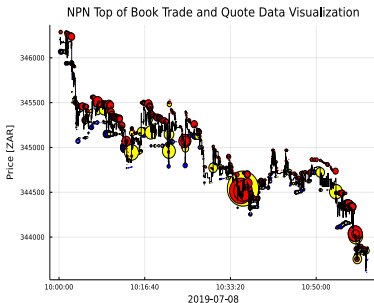
Naspers - NPN

Anglo American - AGL

④ Conclusion

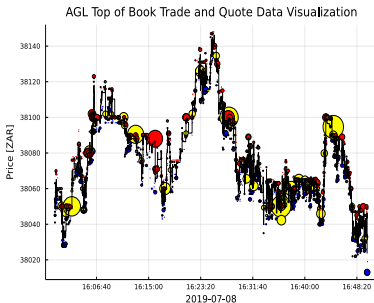
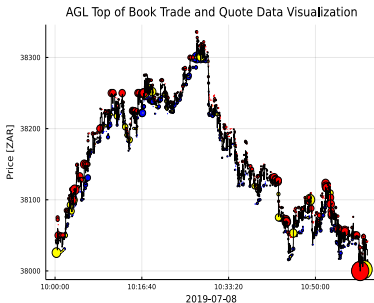
Visualizing NPN Top of Book Data

- In both time slots there are multiple large orders that are likely to have consumed the entire volume of the best bid and ask.
- These orders are going to have a large impact on the mid-price.



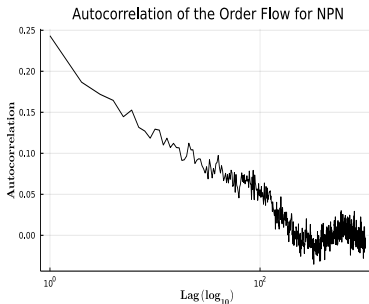
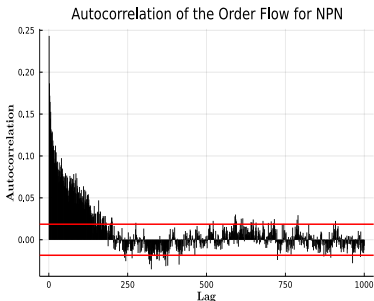
Visualizing AGL Top of Book Data

- The relative size of the volumes are slightly lower for AGL. There are also fewer trades and quotes with relatively high volume.
- Plot on the left shows some large trades occurring close to the peaks and troughs of the price. Could be a random artifact or well executed trades.



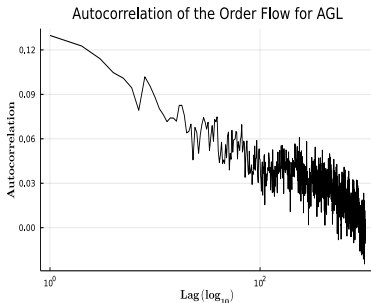
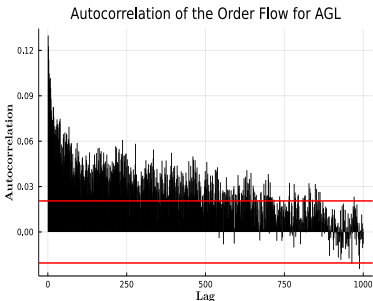
Auto-correlation in the Order Flow for NPN

- The plot on the left shows that the order-flow is significantly correlated up to a lag of approximately 200.
- The stylised fact of order-flow auto-correlation has been shown to occur for NPN over the week of trading.



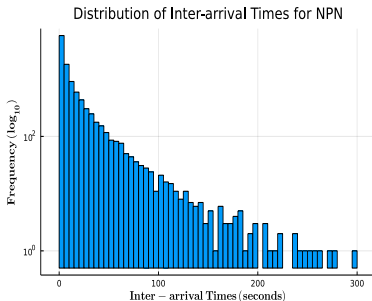
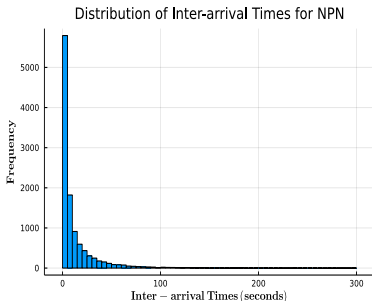
Auto-correlation in the Order Flow for AGL

- We are getting a much lower value for the auto-correlation, and thus a lower significance, but we see that the significance of the auto-correlation persists for a longer period of time.
- We observe that significant auto-correlations are present up to 500 lags.



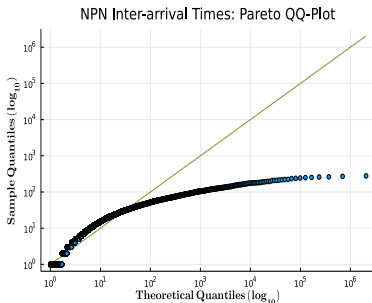
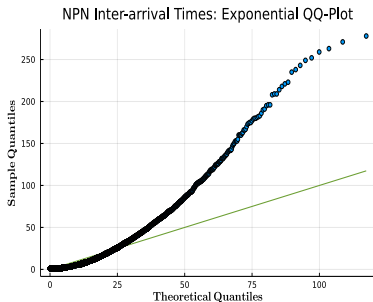
Inter-arrival Time Distributions for NPN

- The histogram on the left shows that the distribution of inter-arrivals is likely going to be heavier tailed than an exponential distribution.
- The log scaled plot shows that the inter-arrivals should not follow a power law distribution.



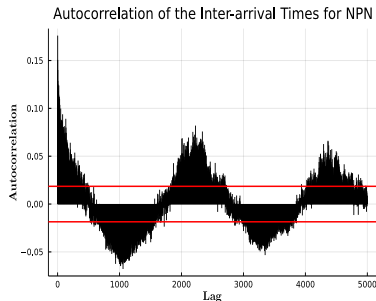
Inter-arrival Time QQ-plots for NPN

- We can see that the inter-arrival times are fatter tailed than the exponential distribution but far lighter tailed than the Pareto distribution.



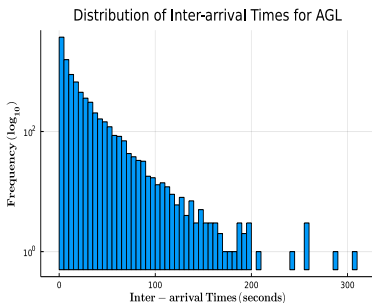
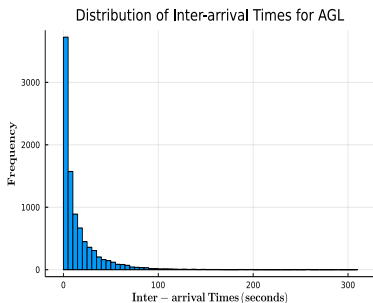
Auto-correlations in the Inter-arrival Times for NPN

- The auto-correlations show that there is a clear periodic pattern of trading NPN.
- Initially trading is occurring quickly. Then the inter-arrival times start to increase until it a peak at midday. After that there is once again an decrease in inter-arrival times.



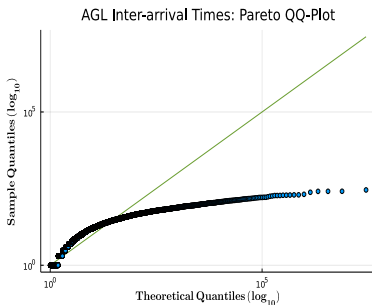
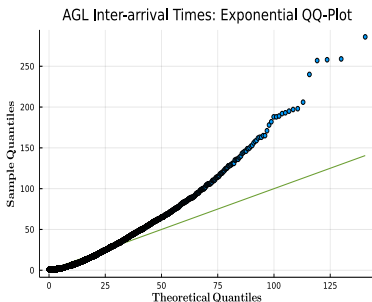
Inter-arrival Time Distributions for AGL

- Very similar to what we saw for NPN.
- The log scaled plot has a steeper slope indicating a slightly lighter tailed distribution when compared with NPN.



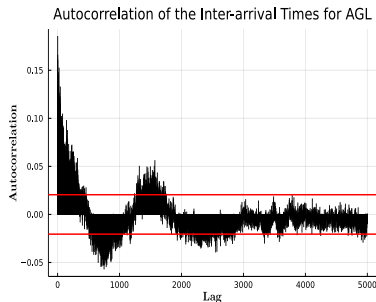
Inter-arrival Time QQ-plots for AGL

- Once again similar to what was seen for NPN.
- The QQ-plots show that the inter-arrival times distribution for AGL is lighter tailed when compared with NPN. It more closely resembles the exponential distribution.



Auto-correlations in the Inter-arrival Times for AGL

- The periodic pattern is also present for AGL. However, the pattern is far less pronounced when compared to what we saw for NPN.



Key Takeaways

- NPN had more relatively large trade and quotes when compared with AGL.
- Auto-correlation in the order flow was shown to be present in both equities.
- Both equities' inter-arrival times had distributions that were heavier tailed when compared with the exponential distribution and lighter tailed when compared with the Pareto distribution.
- The auto-correlation in the inter-arrival times showed a clear pattern in the rate of trading for both stocks.