On the prominence of number:

Using large datasets to explore the categorical representation of number

Junyang Guo & Neil Stewart

Join the Google Meeting here
or use the link: https://meet.google.com/mvx-yraf-iub

Summary

While virtually all economic models implicitly assume that decisions are made on a continuous number line, the empirical distributions of choices largely cluster on some particular numbers.

This observation prompted us to develop a new theory of the representation of number. This complements other accounts of the representation of number including:

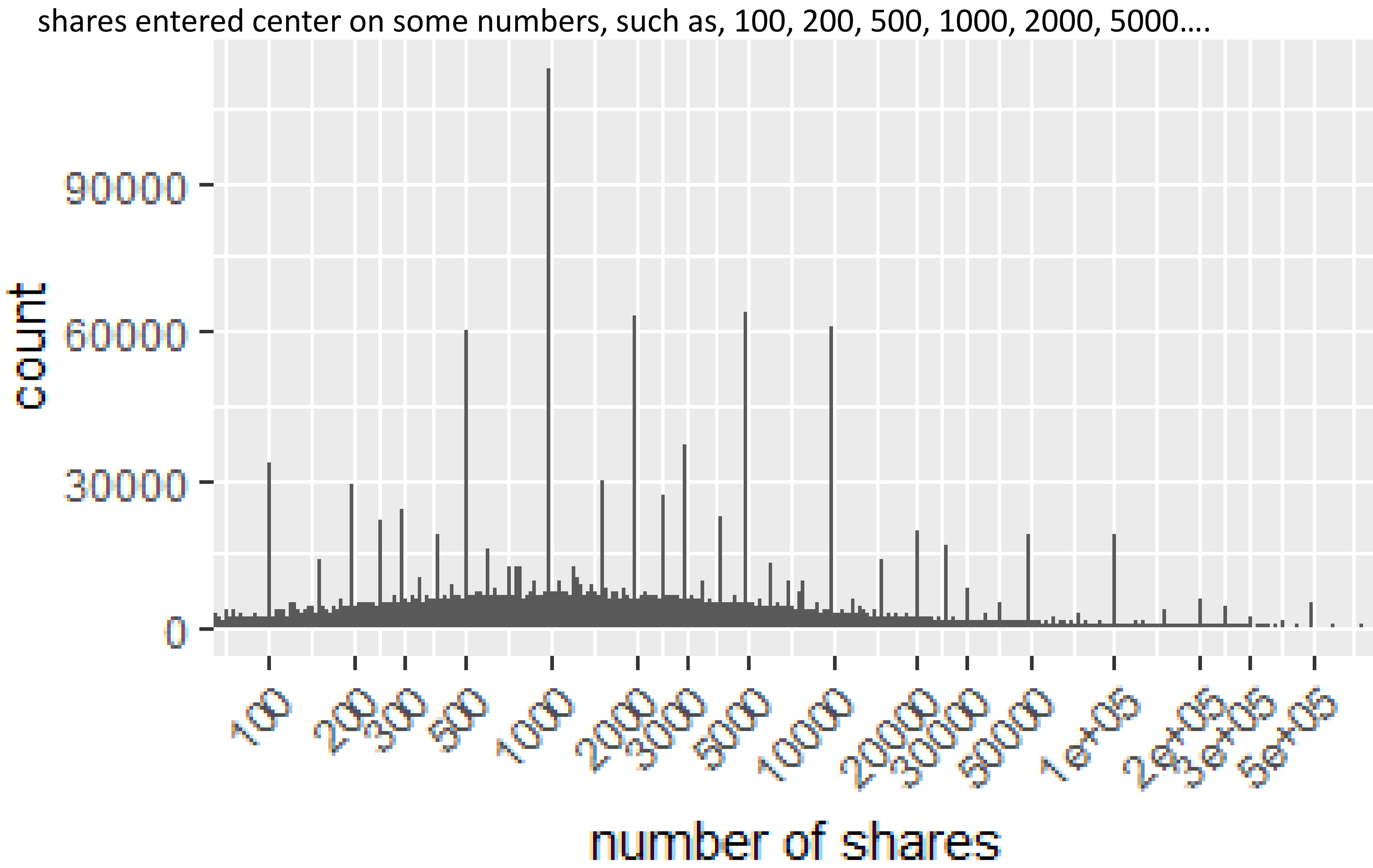
- The logarithmic mental number line (Dehaene, 2003)
- Subitizing small numbers from visual patterns (Kaufman, Lord, Reese, & Volkmann, 1949)
- Symbolic representation (Gelman & Gallistel, 1986)

We propose that numbers are also represented in a discrete, categorical representation, such that free choices over the number scale are dominated by a small set of prominent numbers (Albers, 1997)

There are two major parts of the model. First, a continuous distribution captures the latent desire, which can be in any continuous forms. The second part is a mental number line. Each prominent number has a lognormal distribution. After adjusting the distributions by a set of weight signaling the prominence of each prominent number, any numbers along the real number line have certain coverages from each distribution. The relative coverages from these prominent numbers represent the probability that any number will be pulled towards a particular prominent number. A concrete example will be provided in the model session.

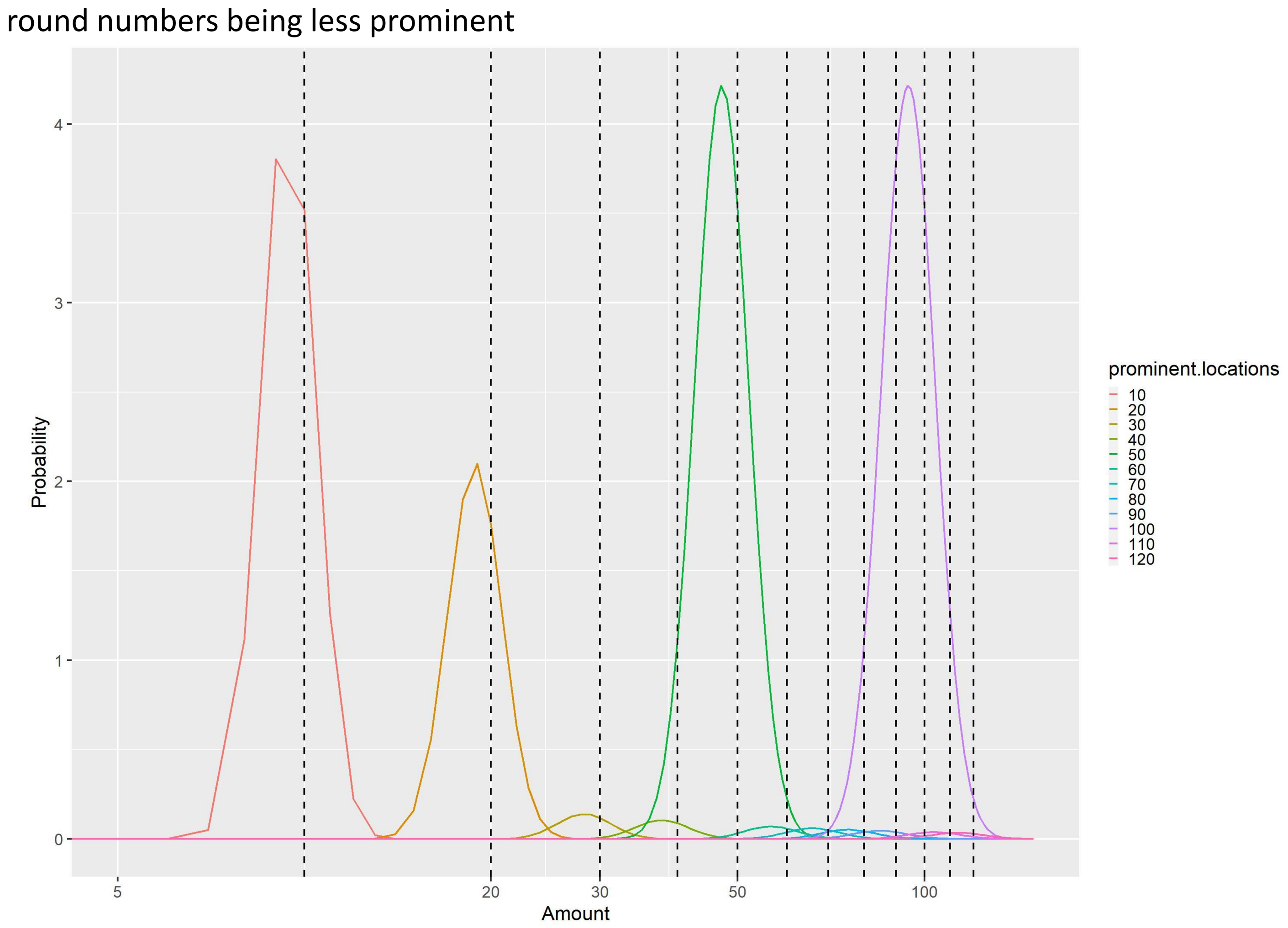
Prevalent Use of Prominent Numbers

When individuals purchase a stock from a stockbroker, they can either enter an amount of shares or an amount of money. Here we show that, although there is no restrictions on choosing from any integers, numbers entered cluster on some particular numbers. The figure below shows that the number of shares entered center on some numbers, such as, 100, 200, 500, 1000, 2000, 5000....

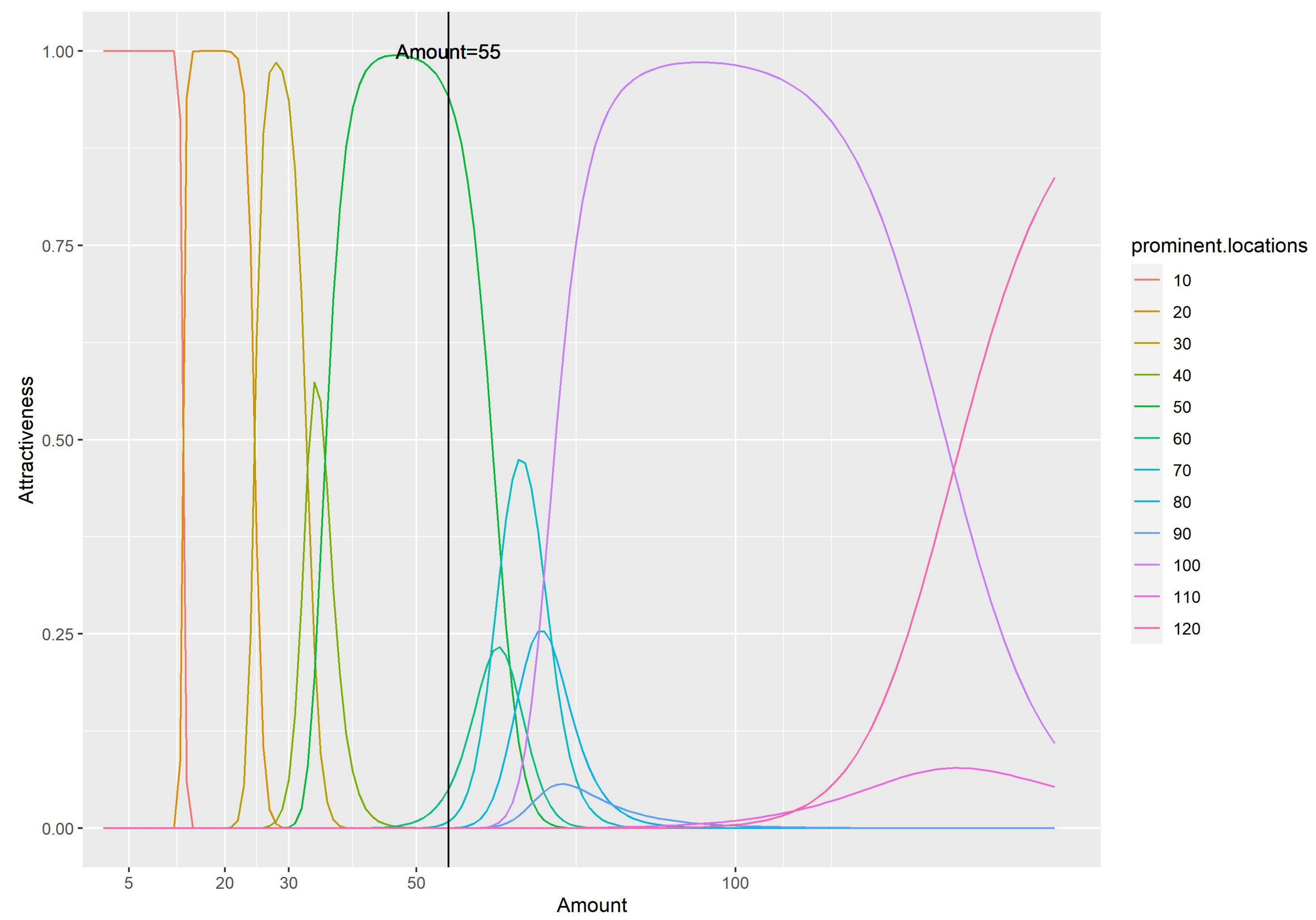


Model

We propose that each prominent number has a lognormal generalisation function which captures the gravity of that prominent number. Here we see that 10, 50, and 100 are particularly prominent, with other round numbers being less prominent

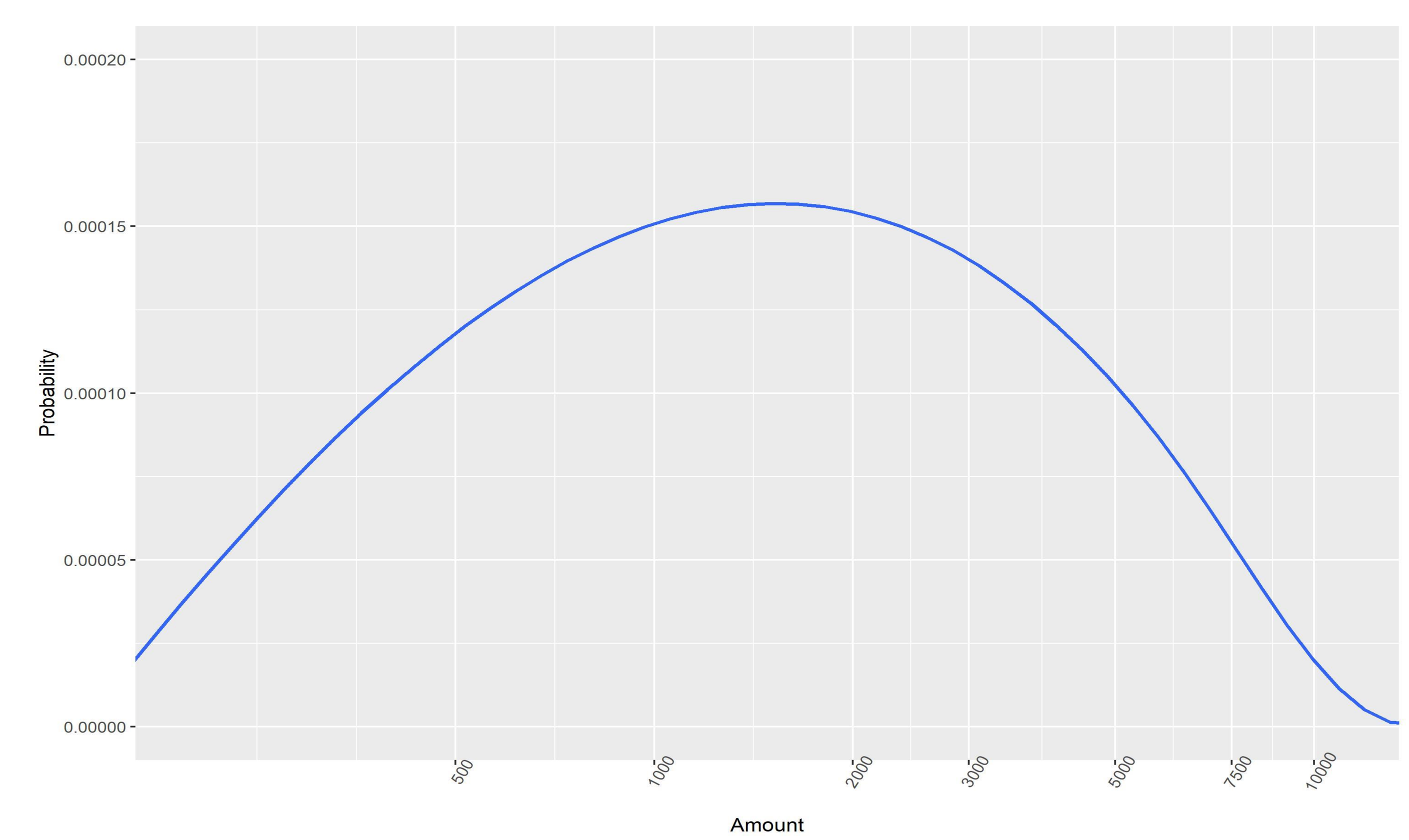


We normalise these generalisation functions to get the probability that any number will be pulled towards a particular prominent number. For example, for the desire of amount, 55, the probability of choosing 50 is 0.95 while the probabilities of choosing 60 and 70 are about 0.04 and 0.01.

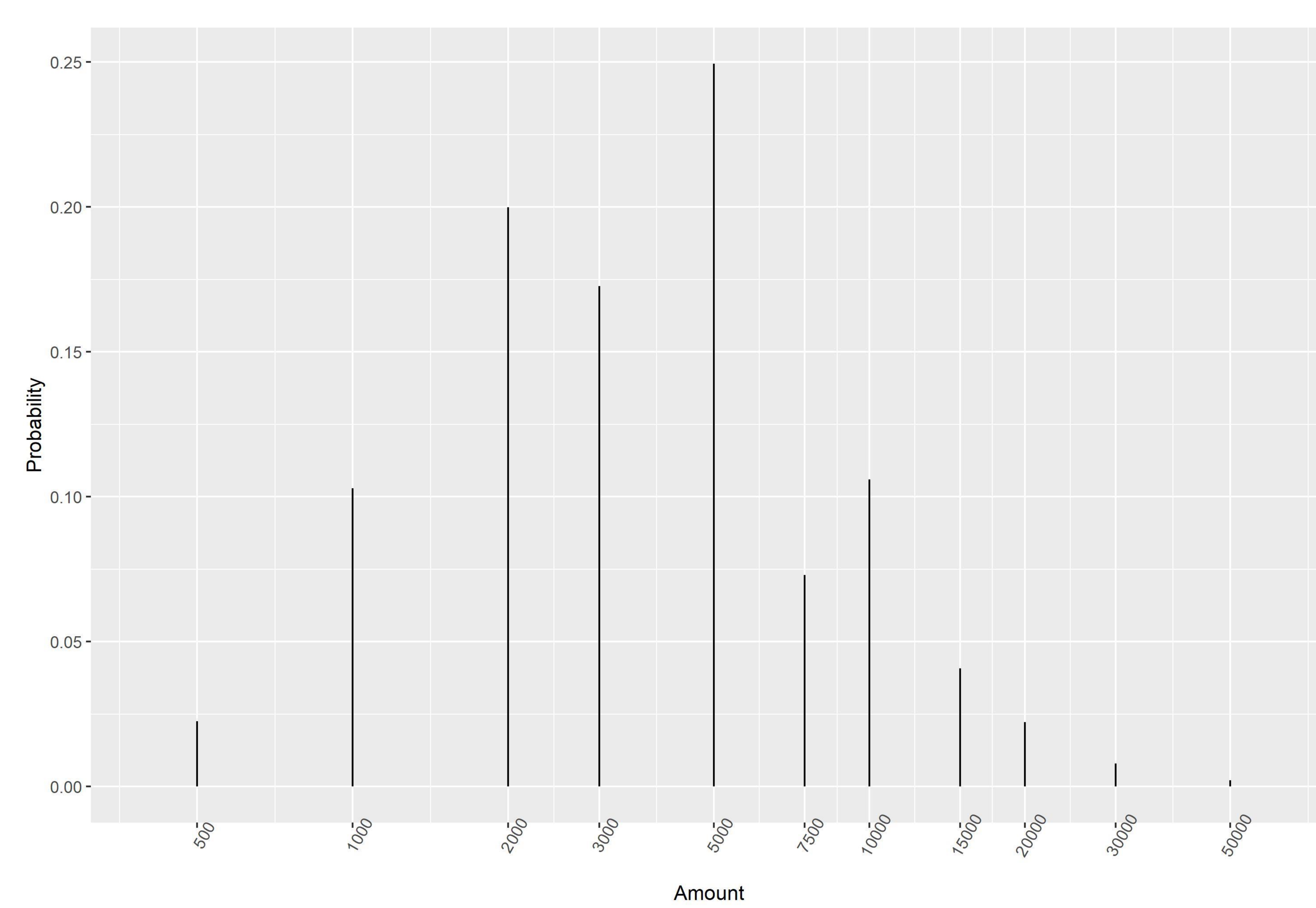


Applying the Model

We assume a latent desire over the quantity of stock purchased, which will be a result of the factors that draw people to purchase a stock (e.g., price movements, industry sector, wealth and liquidity, etc.)

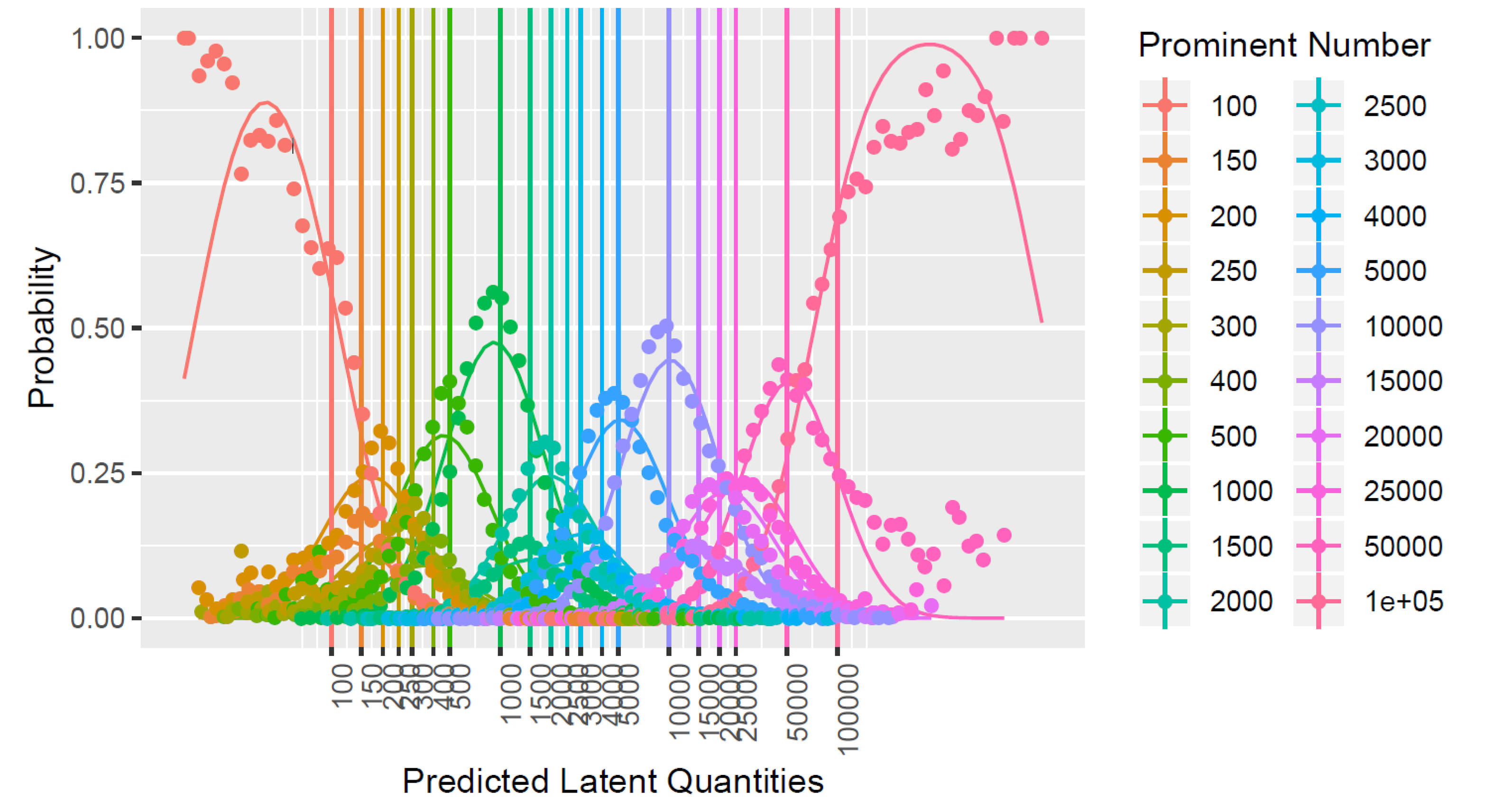


We then allow this latent desire to be drawn to prominent numbers, to make predictions about how much of each quantity will actually be purchased



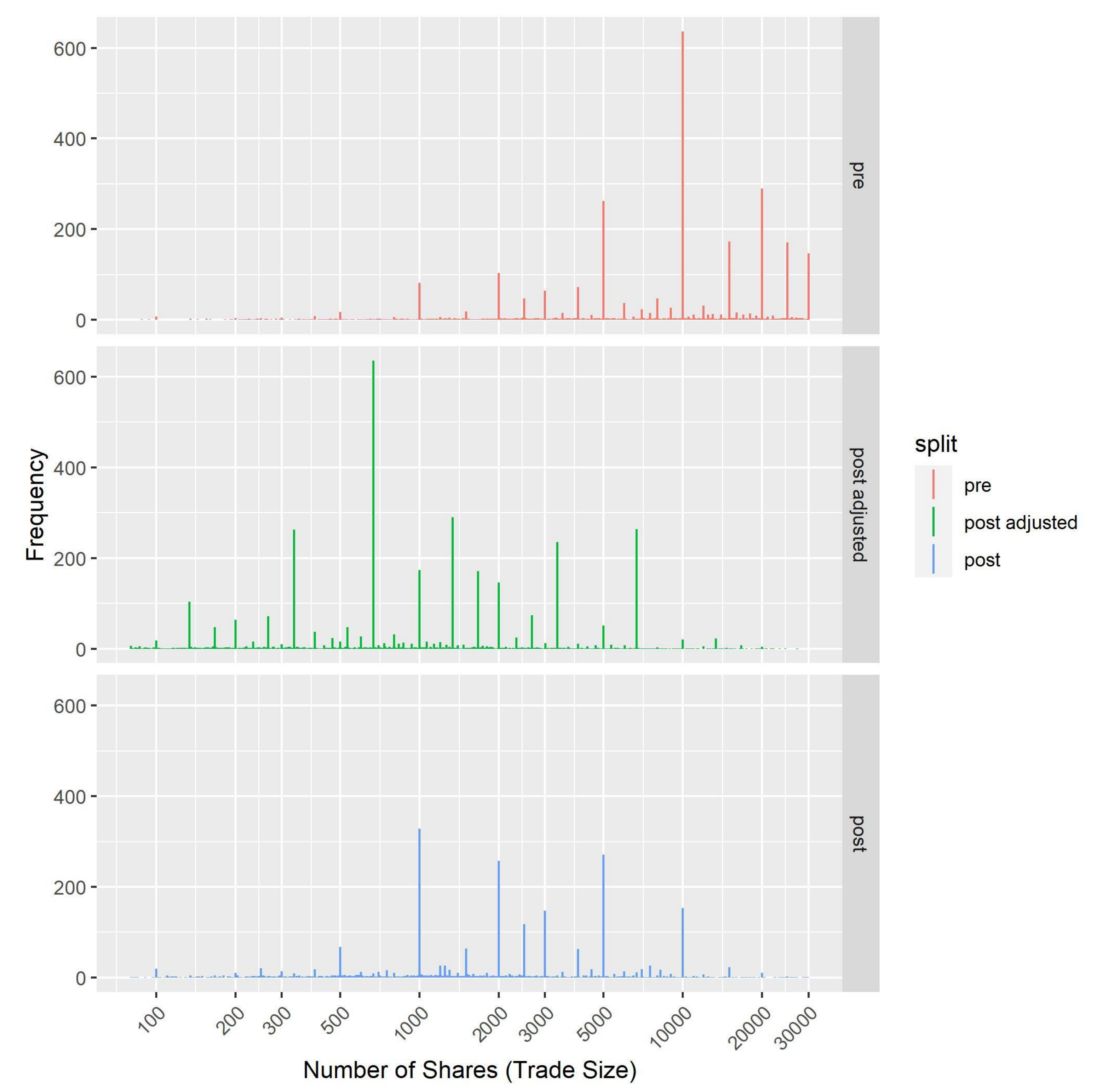
Model Fitting

For each purchase we predict a (continuous) number of stocks purchased from economic fundamentals. Below, the x-axis is that predicted latent quantity. The curves show how often each latent quantity is resolved to each prominent number in the data (dots) and in the best-fitting model (lines)



Real Life Example – Stock Split

On 2014-06-20, Watchstone Group plc took place a case of reverse split, each 15 shares is replaced by 1 share each worth 15 times of the original price. This means a purchase of 10,000 stocks before the split is equivalent to purchasing 667 stocks after the split. While purchases of exactly 10,000 stocks are extremely common before the split (top), the distribution after the split(bottom) is very different from the distribution we'd expect (middle) under the assumption that people spend the same amounts after the split as they did before. Almost no one purchases 667 stocks after the split. Instead, individuals purchase a prominent number of stocks, most commonly, 500 or 1,000 stocks instead.



Reference

Dehaene, S. (2003). The neural basis of the Weber–Fechner law: a logarithmic mental number line. *Trends in cognitive sciences*, *7*(4), 145-147.

Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkmann, J. (1949). The discrimination of visual number. *The American journal of psychology*, *62*(4), 498-525.

Gelman, R., & Gallistel, C. R. (1986). *The child's understanding of number*. Harvard University Press.

Albers, W. (1997). Foundations of a theory of prominence in the decimal system. Part I: Numerical response as a process, exactness, scales, and structure of scales.