

# Lottery stocks

## Project idea

This project attempts to provide a potential explanation to the value premium puzzle. The overall idea is to investigate the behavior and strategies of two different types of investors/analysts (agents hereafter). The first type includes agents who are thriving among their peers. They are the top performers for given tasks, and desire to maintain their status. Lets label this group “top performers”. The second type includes agents who are not top performers, and desire to improve their status and performance among their peers. Let this group be labeled as “status-seekers”.

Our hypothesis is two-fold. We assume that the top performers adopt strategies that are adequate enough to outperform their peers, and help maintain their status at the same time. In other words, they seek less risky choices (fearing a status loss in case of failure) with potential high returns (to continue outperforming). On the other hand, the status seekers are aware of the top performers’ strategy. Their skill level does not allow them to outperform their peers by adopting a similar strategy. Thus, we assume that they are forced to seek more risky choices to improve their status.

The next series of assumptions try to mock up a setup to help explain the observed trends of the value premium puzzle. Choices of higher risk could possibly have lower average returns compared to top performer strategies, a feature assumed in order to account for the top performers’ dominance on average. However, it is possible that the distribution of returns from high-risk choices asymptotically dominates the corresponding low-risk distribution. The high-risk pdf could either be more right-skewed or have thicker tails compared to the low-risk pdf. As a result, a status seeker being fully aware that cannot outperform his peers with low-risk strategies, attempts high-risk strategies with a “lottery”-type of payoff. The reason is that he values status more than average returns, and aims to win the lottery as the only alternative option to outperform his peers.

We would like to test these hypotheses in the data. The problem setup implies that top performers invest on value stocks, while status seekers invest on growth stocks with potential lottery payoffs. Currently, we plot these distributions and perform tests to verify our claim.

## Description of data

We choose two dates (cross-section points) to calculate stock returns. Growth and value stocks are identified through the market to book ratio. We consider the ratio’s value on the initial date only to label stocks as growth or value (those that qualify, not all stocks). There is the thought of using the M/B ratio on both dates, but that would exclude firms whose stock was initially growth and has evolved to value (or at least no longer growth) within the period considered.

In order to calculate the M/B ratio, we merge data from the Daily Stock File (DSF) and its header (DSFHDR) in CRSP with the Fundamental Annual (FUNDA) data in COMPUSTAT<sup>1</sup>. Firms that have been liquidated in between the given cross-section points are excluded, but care is taken to include events such as M&As that result in changes to data headers (tickers, names, etc). Penny stocks (below a set threshold, typically \$10.00) are excluded, but a limited number of stocks above the threshold price and negative book value are included. However, the ranking among stocks is done by the absolute value of M/B. Further filters include

1. Choose only Ordinary Common Shares (share codes “10”, “11”)<sup>2</sup>
2. Include only certain link types (“LC”, “LU”, “LS”, “LN”)
3. Exclude dummy IIDs (existing company, non-existent security)
4. Sanity check for price (allow negative for bid/offer average)
5. Sanity check for outstanding shares
6. Market cap upper and lower limits

The M/B ratio calculation includes unavoidably numbers reported on different dates. Since the accessible data on firm fundamentals are only those with annual periodicity (due to USC subscription to WRDS), we use the past fiscal-year-end reporting date that is closest to our chosen cross section date. This approach coincides most of the times with the standard method that Fama & French have established, when the reported fiscal-year-end date is close to December. However, I make sure that I still get the closest past fiscal-year-end date for cases where this date is end of March for some firms.

Given the difference between the selected date (first cross-section point) *CS* and the most recent past fiscal-year-end date *FYE*, a straightforward way to define the M/B ratio is

$$\begin{aligned} M/B^* &= \frac{\text{stock price}(CS) \times \text{outstanding shares}(CS)}{|\text{book value per share}(FYE)| \times \text{outstanding shares}(FYE)} \\ &= \frac{\text{market value CRSP}(CS)}{\text{book value COMPUSTAT}(FYE)} \end{aligned}$$

However, there are cases where a firm has more than one class of shares, where one or more classes are not traded. Data for such classes are not included in CRSP data, but are included in the calculation of the firm’s book value in COMPUSTAT. In order to correct approximately for such cases, we use the following M/B ratio

$$M/B = \frac{\text{market value COMPUSTAT}(FYE)}{\text{market value CRSP}(FYE)} M/B^*$$

---

<sup>1</sup>The linking set CCMXPF\_LINKTABLE from CRSP is used to properly link the libraries.

<sup>2</sup>Other authors include preferred stock, other tangible assets, etc.