PHILOSOPHICAL ECONOMICS

Profit Margins, Bayes' Theorem, and the Dangers of Overconfidence

Posted on September 4, 2017 by philosophicalecon@gmail.com

It's the fall of 2011. Investors are caught up in fears of another 2008-style financial crisis, this time arising out of schisms in the Eurozone. The S&P 500 is trading at 1200, the same price it traded at in 1998, roughly 13 years earlier, despite the fact that its earnings today are almost three times as high as they were back then. The index's trailing price-to-earnings (P/E) ratio sits at around 12, significantly below the historical average of 16.

Suppose that I'm a hypothetical value-conscious investor who has been taking a cautious stance on the market. I look at the market's valuation, and think:

> "Stocks look cheap here. I've been holding all this cash, waiting for an opportunity. Maybe I should get in."

I then remember:

> "Stocks only *look* cheap because earnings have been inflated by record-high corporate profit margins–10% versus a historical average of 6%. When profit margins revert back to the mean, as they've done every time they've reached these levels in the past, S&P 500 earnings will shrink from $100 down to $60, lifting the index's P/E ratio from a seemingly cheap 12 to a definitively not cheap 20."

With that concern in mind, I hold off on buying stocks, and decide to instead wait for profit margins, earnings and (true) valuations to come back down to normal historical levels.

The year 2012 comes and goes. Profit margins stay elevated, so I keep waiting. 2013 follows–again, profit margins stay elevated, so I keep waiting. 2014 after that–again, profit margins stay elevated, so I keep waiting. Then 2015, then 2016, then 2017–each year I wait, and each year I end up disappointed: profit margins fail to do what I expect them to do. But I'm a disciplined investor, so I keep waiting. During the total period of my waiting, the stock market more than doubles in value, trouncing the returns of my cash-heavy portfolio and leaving me with an ugly record of underperformance.

To evolve as an investor, I'm eventually going to have to be honest with myself: I got something wrong here. Rather than fight that fact, I'm going to need to open up to it and learn from it. I'm going to need to re-examine the potentially mistaken beliefs that brought it about–in this case, potentially mistaken beliefs about the inner workings of corporate profitability.

> "I was confident that elevated corporate profit margins would revert to the mean, which is what happened every time they were elevated in the past. But that didn't happen here. Why didn't it happen? Where did my analysis go wrong? What did I fail to see?"

These questions are all well and good, but there is a more important question that I'm going to need to ask, a question that often gets missed in post-mortem investigations of this type. Specifically:

> "Why did it take me so long to update my beliefs in the presence of repeated disconfirmation? I had a thesis: that the elevated corporate profit margins I was seeing in 2011 would fall back down to past averages. Reality told me that this thesis might be wrong in 2012, when the prediction failed to come true. Then it told me again in 2013. Then it told me again in 2014, and again in 2015, and again in 2016, and again in 2017. Was all of this repetition really necessary? Could I have been more receptive of the message the *first* time it was presented?"

Winning in the investing game isn't simply about having true prior beliefs about the world. It's also about efficiently updating those beliefs in response to feedback from reality. The primary mistake that I made in the above scenario was *not* the mistake of having incorrect prior beliefs about the likely future direction of corporate profit margins–from the perspective of what I knew in 2011, those beliefs were reasonable beliefs to have. Rather, my primary mistake was my failure to properly update those prior beliefs in response to the steady stream of disconfirmation that kept coming in. The updating process should have moved me to a different stance *sooner*, which would have allowed me to participate in a greater share of the returns that the market went on to produce.

The Importance of Updating in Investing: Analogy From a Coin-Flipping Game

To better appreciate the importance of updating in investing, we can explore the following investing analogy, expressed in the terms of a coin-flipping game.

Coin-Flipping Game. Suppose that you and a small group of other people are about to compete with each other in a coin-flipping game. Each player will start the game with $1,000 of play money. Rankings in the game will be determined based on how much each player is able to grow that money over the course of the game. At the end of the game, real monetary prizes and penalties will be assigned to players based on where in the rankings they end up.

Two types of coins can be flipped in the game: green coins and red coins. Green coins are physically designed to have a 70% probability of landing heads and a 30% probability of landing tails. Red coins are physically designed to have the opposite profile: a 30% probability of landing heads and a 70% probability of landing tails.

The game is divided into 20 separate rounds, each consisting of 50 coin flips (1,000 flips in total). At the beginning of each round, the game's referee fills a large bucket with an unknown quantity of red and green coins. He then randomly draws a single coin from it. He uses that coin for all 50 flips in the round, making sure to keep its color hidden from the participants. When the next round comes along, he empties the bucket, refills it with a random number of red and green coins, and draws a coin to use for the ensuing 50 flips.

Before each flip, the referee auctions off "ownership" of the flip to the player that offers to pay the highest price for it. For each bid that a player puts out, other players are given the option of either putting out a higher bid, or stepping aside. If everyone steps aside, then the flip is declared sold.

Once the "owner" of the flip is set, the referee flips the coin. If it lands on heads, he pays out $2.00 in play money to the owner. If it lands on tails, he pays the owner nothing (and therefore the owner loses whatever amount of play money she paid for it). After each round is over, the referee reveals the result of the flip to the participants and opens up bidding for the next flip. The game goes on like this until the end, at which point the final rankings are tallied and the associated monetary prizes and penalties disbursed.

The key to performing well in this game is having an accurate picture of what each flip is "worth." If you have an accurate picture of what each flip is worth, then you will know when the other players are bidding too little or too much to own it, and therefore you will know whether you should increase your bid and buy it, or stand back and let it be sold.

Suppose that the referee is flipping a green coin in a round. The "worth" of each flip, which we take to be the expected payout to the owner, will be $1.40. In general, you should buy the flip if it's being offered at a price below this price, and you should refrain from buying it if it's being offered at a price above it. Of course, any given flip will either land heads and pay out $2.00, or land tails and pay out nothing, so *with hindsight* you will be able to say that a given flip was a good buy even though it was priced above $1.40, or that it was a good sell even though it was priced below it. But in this game you don't have the luxury of making decisions in hindsight. All you can do is look forward. If you do that, you will realize that over a large number of flips with a green coin, heads will tend to occur 70% of the time, and tails 30% of the time. The payout per flip will therefore tend to *average* out to: 0.70*($2.00) + 0.30*($0.00) = **$1.40**, which is the highest price that you should generally be willing to pay. By the same logic, if it turns out that the referee is flipping a red coin in a round, then the expected payout to the owner of each flip, which we take to be it's "worth", will be: 0.30*($2.00) + 0.70*($0.00) = **$0.60**. If the coin is red, then you generally should be willing to buy a flip up to that price, but not above it.

(Note: There are other considerations, beyond the mere "worth" (expected payout) of a flip, that may prove relevant to your decision of how much to bid for it. If you know that other players are likely to try to outbid you, you might want to continue to place bids even after the price has risen above your estimate of the worth, purely in order to force those players to pay higher prices. You might also become rationally risk-seeking, in the sense that you're willing to buy flips at prices above their worth precisely because you're looking for a "gamble"–consider, for example, a scenario near the end of the game in which the person directly ahead of you in the rankings is only slightly ahead, but the person directly behind you is very far behind. In that case, you might have a lot to gain and nothing to lose from a gamble, so you may be willing to take it even at odds that are against you. Finally, given that your expected return from buying a flip will depend on the difference between the worth and the price you pay, you will technically need to stop bidding when the price is some distance below the worth, so

that your expected return stays positive, and also so that you are able to conform with the Kelly Criterion. That necessary distance will usually be tiny, but it could become significant, depending on how much money you have left in the game. These considerations, while interesting, are beyond the scope of what we're trying to explore here.)

To form an accurate picture of what each flip in the game is worth, you're going to need to find out whether the referee is using a green coin or a red coin for the flip. Unfortunately, you can't directly find that out—he's intentionally keeping it a secret from you. However, you might be able to assign a probability that he's using a green coin or a red coin in any given round based on other information that is available to you. Combining that probability with the probability that each type of coin will land heads or tails will allow you to build a second-order estimate of the worth of each flip. That estimate will be some number between \$0.60 (the worth of a red flip) and \$1.40 (the worth of a green flip), scaled based on how likely you think it is that the referee is flipping one type of coin versus the other.

So that's where the real challenge in the game lies. You need to do a better job than the other participants of using available information to form a second-order estimate of the worth of the various flips that are taking place. If you can do that, then over time and across many different rounds and flips, you will tend to buy at better prices than the other participants, which will allow you to earn more money than they earn. It's in this sense that the game represents an apt analogy to investing. The challenge in investing is the same challenge: to do a better job than the rest of the market of using available information to form an estimate of the likely returns of the various investment securities that are on offer, across whatever time horizon you happen to be focusing on. If you can do that, then over the course of many different investment periods, you will tend to invest in securities that produce better returns than the average, which will cause you to outperform the "market."

Returning to the game, to proceed intelligently in it, you're going to need to be given information to use. So let's assume that at the beginning of each round, after the referee draws his coin, he lets you and the other participants dig around inside the bucket to determine how many green coins are in it versus red coins. Knowing how many green coins are in the bucket versus red coins will allow you to assign a probability to the prospect that he drew a green coin or a red coin, given that he randomly drew the coin from the bucket.

To clarify, there are two different senses of "probability" that we've been using here. The first sense is the frequentist sense, in which "probability" is taken to refer to the *frequency* at which something will tend to happen over a large number of trials. For example, over a large number of flips, green coins will tend to fall on heads 70% of the time, and red coins will tend to fall on heads 30% of the time, so we say that green coins have a heads probability of 70%, and red coins have a heads probability of 30%. The second sense is the Bayesian/Laplacian sense, where "probability" is taken to refer to our *degree of belief* in something. For example, suppose that I count the coins in the bucket and determine that there are 9 green coins for every one red coin. The referee drew his coin from the bucket. If he drew it randomly, without preference, then I can say that there's a 9 out of 10 chance, or a 90% probability, that he drew a green coin. But this number only reflects *my* degree of belief that he drew a green coin—in reality, the matter has already been settled, he either drew a green coin or he didn't. These two senses of the term may seem incompatible, but they need not be. In fact, if Laplace is right that the universe is deterministic, then they will essentially reduce to the same thing. All perceived randomness in the world will simply be the result of ignorance.

Suppose that prior to the start of the first round of the game, you dig around in the bucket and come to estimate that roughly 90% of the coins in it are green and that roughly 10% are red. From your perspective, this means that there's a 90% chance that the referee drew a green coin, and a 10% chance that he drew a red one. Combining those probabilities with the green coin's 70% heads probability and the red coin's 30% heads probability, your new second-order estimate of the worth of each flip in the round will be 0.90*0.70*\$2.00 (case: coin is green, referee flips heads) + 0.10*0.30*\$2.00 (case: coin is red, referee flips heads) = **\$1.32**.

Let's now start the first round of the game. For the first flip, you're able to successfully buy it for \$1.20, which is an attractive price from your perspective, below your worth estimate of \$1.32. The result of the flip comes back "Tails", so you lose \$1.20. On the next flip, you're again able to successfully buy it at \$1.20. The result of the flip again comes back "Tails", so you again lose \$1.20. Eight more flips follow. In each case, you successfully outbid the rest of the market, buying at \$1.20. The results of the flips come back TTHTTTTH. After ten flips, the result then stands at 2 heads and 8 tails, leaving you with a cumulative loss of **\$8.00**. Ouch.

You begin to wonder: if the coin is, in fact, green, i.e., 70% biased to land heads, then why is it landing so much on tails? Uncomfortable with the situation, you pause the game to investigate. It seems that you are being confronted with two possible outcomes, both of which are unlikely, and one of which must have actually taken place.

Outcome #1 — The Referee Drew a Red Coin: You determined that the bucket contained 9 green coins for every one red coin. On that basis, there was a 90% chance that when the referee drew the coin for the round, he drew a green coin. Did he, in fact, draw a red coin? It's possible, but unlikely.

Outcome #2 — The Referee Drew a Green Coin but, by Chance, the Flip Results Have Come Back Tails-Heavy: If the first unlikely outcome did not take place–that is, if the referee is, in fact, flipping a green coin as initially expected–then a different unlikely outcome will have taken place. Specifically, the referee will have conducted 10 flips of a coin with a 70% chance of landing heads, and the coin will only have landed heads twice–20% of the time. The flipping process has an element of random chance to it, so this outcome is possible. But it's unlikely.

What you have, then, are two unlikely possible outcomes, one of which actually happened. To properly "update" your beliefs about what color of coin the referee is likely to be using, you're going to have to weigh these two unlikely possible outcomes against together. The correct way to do that is through the use of Bayes' Theorem, which we will now take a detour into to explain. Readers that are already fresh on Bayes' Theorem can feel free to skip the next section–but let me say that I think the explanation that I give in it is a pretty good one, likely to be worth your time, even if you're already strong on the topic.

Bayes' Theorem Explained

Bayes' Theorem expresses the following relationship:

$$P(H|D) = P(D|H) * P(H) / P(D)$$

We can think of the letter H here as referring to some hypothesis or belief, and the letter D as referring to some data or information that is obtained subsequent to that hypothesis or belief. Bayes' theorem tells us how to "update" the probability that the hypothesis or belief is true in light of the data or information that has been obtained. The intuitive basis for the theorem is difficult to grasp, and even more difficult to retain in memory in a clear form. To help make it clear, I've concocted the following spatial analogy.

Imagine a square of area 1, shown below. Inside the square is a circle H of area P(H). Ignore the weirdness of the term P(H) for a moment–just assume that it's a number representing an area. You're standing above the square with a single speck of sand on your finger. You flick the speck down onto the square. It lands somewhere inside the square. You don't know where that is because the speck is too small to see from a distance. It could be anywhere.
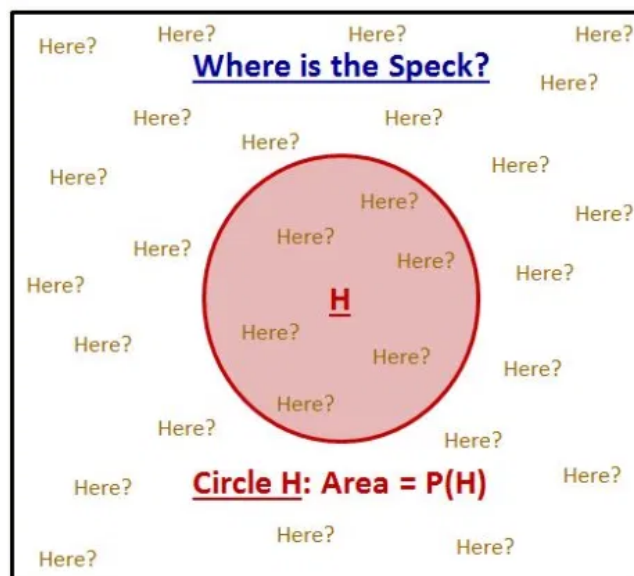
The question we want to ask is, what is the probability that the speck is inside circle H? Given that it was flicked onto a random spot inside the square, the answer has to be: the area of H, denoted by P(H), divided by the area of the square, which is 1. Think about that for a moment and you will see why it has to be true: the only factor that can impact the probability that a randomly located speck is inside a given space is the area of the space. P(H) / 1 = P(H), so the probability that the speck is inside H is simply P(H).

Square: Area = 1

The speck is located at some _random_ spot inside this square, but you don't know where. The probability that it is inside Circle H is simply: P(H) / 1 = P(H).
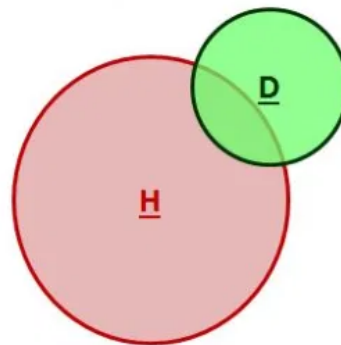
H
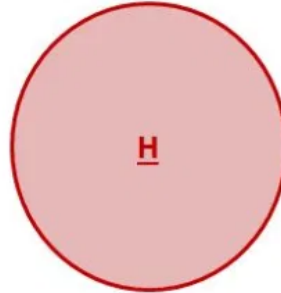
Circle H: Area = P(H)

Now, suppose that I draw another circle inside the square and label it circle D, with an area of P(D). I then reveal to you that when you flicked the speck onto the square, it landed somewhere inside circle D. To repeat, the speck of sand is located somewhere inside circle D–you now know this for a fact.

Square: Area = 1

Circle D: Area = P(D)

D

H

Circle H: Area = P(H)

The question then becomes, _knowing_ that the speck is located somewhere inside circle D, how does your estimate of the probability that it is inside circle H change? In other words, what is the probability that the speck is inside H given that it is known to be somewhere inside D? The way we express this latter value is with the term P(H|D), which means the probability of (the speck being in) H given (that we know the speck is in) D.

Intuitively, we can see that the value of P(H|D) is simply the area of the overlap between circle H and circle D, which we label as P(H&D), divided by the area of circle D, which is P(D).

Square: Area = 1

Overlap b/t H and D:
Area = P(H&D)

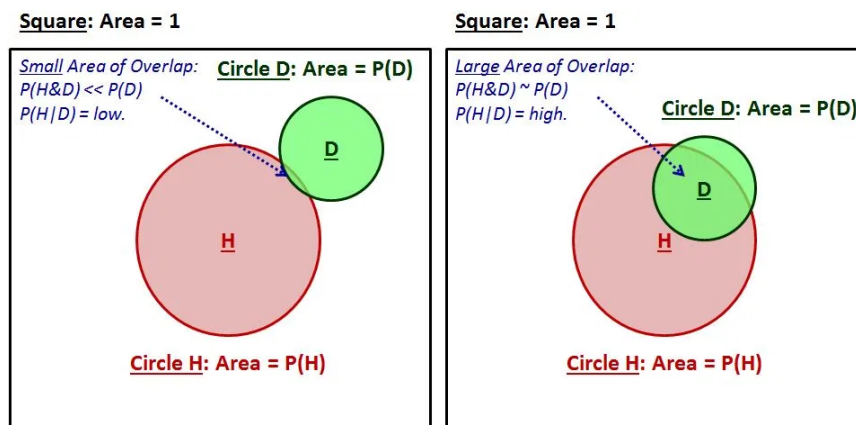**Circle D**: Area = P(D)

D

H

Circle H: Area = P(H)
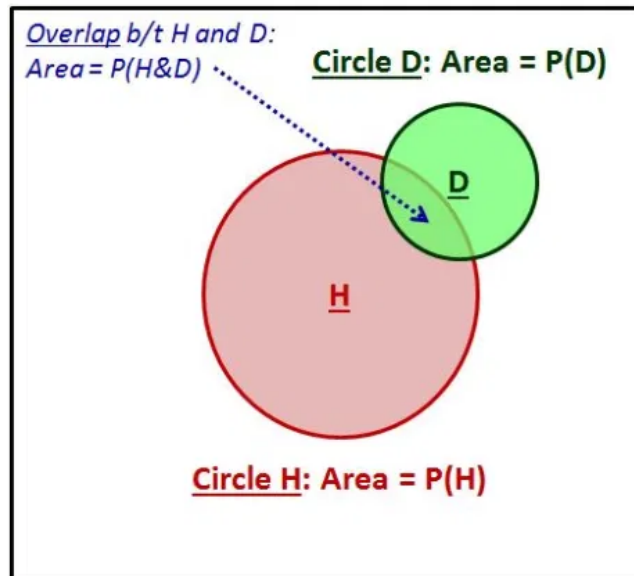
Expressing this intuition formally, we get at a simplified version of Bayes' Theorem:

$$P(H|D) = P(H\&D) / P(D).$$

What the theorem is saying is that the probability of (the speck being in) H given (that the speck is in) D is equal to the area of overlap between H and D (denoted by P(H&D)), divided by the area of D (denoted by P(D)).

Notice that if the area of overlap between H and D is small compared to the area of D, then the probability of (the speck being in) H given (that the speck is in) D will be low (see left schematic). And if the area of overlap between H and D is large relative to the area of D, then the probability of (the speck being in) H given (that the speck is in) D will be high (see right schematic).

Square: Area = 1

Small Area of Overlap:
P(H&D) << P(D)
P(H|D) = low.

Circle D: Area = P(D)

D

H

Circle H: Area = P(H)

Square: Area = 1

Large Area of Overlap:
P(H&D) ~ P(D)
P(H|D) = high.

Circle D: Area = P(D)

D

H

Circle H: Area = P(H)

To make the theorem useful for quantitative applications, we incorporate the following equality:

$$P(H\&D) = P(D|H)*P(H)$$

To understand this equality, recall that P(H&D) is the probability that the speck is inside both H and D. Intuitively, that probability is equal to the probability that the speck is inside H—which is P(H)—times the probability that it is inside D given that it is inside H—which is annotated P(D|H).

Substituting the above equality into the simplified version of the theorem, we arrive at the more familiar version, presented at the beginning of the section:

$$P(H|D) = P(D|H)*P(H) / P(D)$$

In Bayesian applications, the term P(H) is called the "prior probability." It's the initial probability that we assign to our hypothesis being true. Subsequent to that assignment, we will receive data with implications for the truth of the hypothesis. The term P(D|H), called the "likelihood function", expresses how likely it is that we would receive that data assuming that the hypothesis is true. To "update", we multiply the prior probability times the likelihood function. We then divide by P(D), sometimes referred to as the "normalizing constant", which ensures that a measure of 1 is obtained across the overall probability space.

Our "speck of sand" analogy provides a useful intuitive illustration of how the process of Bayesian updating works. We start with a hypothesis: that the speck of sand is located inside circle H (note that we chose the letter 'H' to symbolize 'hypothesis'). We assign a prior probability P(H) to that hypothesis being true. It is then revealed to us that the speck of sand is located inside a second circle, D. This fact obviously has implications for our hypothesis–it is relevant data, which is why we labeled the circle with the letter 'D'. Upon receiving this data, we update the probability that the hypothesis is true to be a new number. Specifically, we set it to be equal to "the area of overlap between H and D" divided by "the area of D." Intuitively, that's what it immediately changes into, once we know that the speck is inside D.

To extend this "speck of sand" intuition to broader applications, we need to understand that for any data that we obtain subsequent to a hypothesis, the hypothesis will exhibit some "overlap" with the data, which is to say that the truth of the hypothesis will represent one possible pathway through which that data might have been obtained. To estimate the probability that the hypothesis is true given that the data was obtained, we need to quantify how prevalent that pathway is relative to all pathways through which the data could have been obtained, including alternative pathways that conflict with the hypothesis. That is what Bayes' theorem does.

The Dangers of Overconfidence

To return to the coin-flipping game, recall that you were struggling with a dilemma. On the one hand, after digging through the bucket, you estimated that 9 out of 10 coins in the bucket were green, and therefore that there was a 90% chance that the referee, who randomly drew his coin from the bucket, was using a green coin. On the other hand, after several rounds of the game, you noticed that a string of tails-heavy results had been accumulating, an outcome that you would not have expected to see if a green coin were being used. The solution to this dilemma is to update your initial estimate of the probability that the referee is using a green coin to reflect the implication of the tails-heavy result that you've since observed.

In truth, you should have been doing that the entire time–the fact that you weren't is part of the reason why you've been losing money in the game. Recall that the coin-flipping game, like the game of investing, is ultimately a game about who is able to do the best (most efficient, most accurate) job of using available information to build an estimate of what things are worth. Here, "available information" isn't limited to your "prior", i.e., your initial estimate of the probability that the referee was using a green coin. It also includes the actual results of the flips that have been accumulating since the round began– those results contain valuable information about what type of coin the referee is likely to be using, information that you cannot afford to ignore.

The table below shows what a proper updating process would look like during the round, assuming that we start out with 90% confidence (prior) that the coin is green. The two important columns in the table are "Rnd Gen (H/T)", which shows the cumulative results of the flips in the round, and "Updated Worth ($)", which shows how our estimates of the worth of each flip evolve in response to them.

H = GreenCoin, D = Results, 1 = Heads, 0 = Tails		Rnd Gen (H/T)	Rnd Gen (1/0)	Sum Wins	Total Flips	P(D|H)	P(H)	P(D|~H)	P(~H)	P(D)	P(H|D)	Updated Worth ($)
GreenCoin Win Odds	70%	T	0	0	1	0.300	0.900	0.700	0.100	0.340	0.794	1.2353
RedCoin Win Odds	30%	T	0	0	2	0.090	0.900	0.490	0.100	0.130	0.623	1.0985
		T	0	0	3	0.027	0.900	0.343	0.100	0.059	0.415	0.9317
		T	0	0	4	0.008	0.900	0.240	0.100	0.031	0.233	0.7863
Win Value	2	H	1	1	5	0.028	0.900	0.360	0.100	0.062	0.415	0.9317
Lose Value	0	T	0	1	6	0.010	0.900	0.303	0.100	0.039	0.233	0.7863
		T	0	1	7	0.004	0.900	0.247	0.100	0.028	0.115	0.6921
Prior Prob of GreenCoin	0.900	T	0	1	8	0.001	0.900	0.198	0.100	0.021	0.053	0.6423
		T	0	1	9	0.000	0.900	0.156	0.100	0.016	0.023	0.6187
		H	1	2	10	0.001	0.900	0.233	0.100	0.025	0.053	0.6423
		T	0	2	11	0.001	0.900	0.200	0.100	0.020	0.023	0.6187
		H	1	3	12	0.001	0.900	0.240	0.100	0.025	0.053	0.6423
		T	0	3	13	0.001	0.900	0.218	0.100	0.022	0.023	0.6187
		H	1	4	14	0.001	0.900	0.229	0.100	0.024	0.053	0.6423
		H	1	5	15	0.003	0.900	0.206	0.100	0.023	0.115	0.6921
		T	0	5	16	0.001	0.900	0.210	0.100	0.022	0.053	0.6423
		T	0	5	17	0.001	0.900	0.208	0.100	0.021	0.023	0.6187
		H	1	6	18	0.001	0.900	0.187	0.100	0.020	0.053	0.6423
		T	0	6	19	0.001	0.900	0.192	0.100	0.020	0.023	0.6187
		T	0	6	20	0.000	0.900	0.192	0.100	0.019	0.010	0.6081
		T	0	6	21	0.000	0.900	0.188	0.100	0.019	0.004	0.6035
		T	0	6	22	0.000	0.900	0.181	0.100	0.018	0.002	0.6015
		H	1	7	23	0.000	0.900	0.178	0.100	0.018	0.004	0.6035
		T	0	7	24	0.000	0.900	0.176	0.100	0.018	0.002	0.6015
		T	0	7	25	0.000	0.900	0.171	0.100	0.017	0.001	0.6006
		H	1	8	26	0.000	0.900	0.167	0.100	0.017	0.002	0.6015
		H	1	9	27	0.000	0.900	0.150	0.100	0.015	0.004	0.6035
		H	1	10	28	0.000	0.900	0.126	0.100	0.013	0.010	0.6081
		T	0	10	29	0.000	0.900	0.135	0.100	0.014	0.004	0.6035
		T	0	10	30	0.000	0.900	0.142	0.100	0.014	0.002	0.6015

Assuming that the referee has, in fact, been using a red coin with a 30% heads probability (the assumption that we used to generate the above data), it will take our updating process around 9 flips to sniff that fact out. After those nine flips, our worth estimate will have effectively converged onto the correct value of $0.60, even though we started the process with a belief that was incorrect.

To summarize, the proper way to play each round of the game is as follows:

(1) Assign a prior probability to the hypothesis that the referee is using a green (or a red) coin, and use that probability to calculate the worth of each flip. To assign a good prior probability, we need information. There are many ways to get it. We can sample the contents of the bucket and use the observed ratio of green coins to red coins to infer a probability, which is what the referee was allowing us to do. We can study the bidding patterns of the other participants, which might contain valuable clues as to the color of coin being used. We can install hidden cameras in the place where the flips are being conducted, which will allow us see the color of the coin for ourselves. We can try to convince insiders who know what color the coin is to reveal that information to us. We can even pay the referee to tell us the color directly. Any piece of information will potentially be valuable here, if it can improve our estimate of the probability that the referee is using a given type of coin.

(2) Update that probability as actual coin flip results accumulate. For this, we use Bayes' Theorem.

If we're more efficient and accurate in performing (1) and (2) than our fellow participants, then over many rounds and many flips, we will tend to earn more money than they earn. The same is true in the game of investing.

Now, suppose that we're in a new round where a red coin is actually being used, but we initially think it's likely to be a green coin. The following chart shows how our estimates of the worth of each flip will evolve in that case. The different lines show the different worth estimates that we would arrive at using different prior green coin probabilities: 0.500 (no idea), 0.900 (likely green), 0.990 (very likely green), and 0.999 (virtually guaranteed to be green). The correct worth estimate, of course, is $0.60, because the coin is, in fact, red. By updating properly, we will eventually get to that estimate, on each of the assumed priors. The difference, of course, will be in how many flips it will take for us to get there, and how much we will lose in the interim period from our resulting willingness to overpay.

(Note: Y-axis is the worth estimate, X-axis is the flip number in the round. Each line begins after the results of the first flip, so the first worth estimate is already an updated number.)
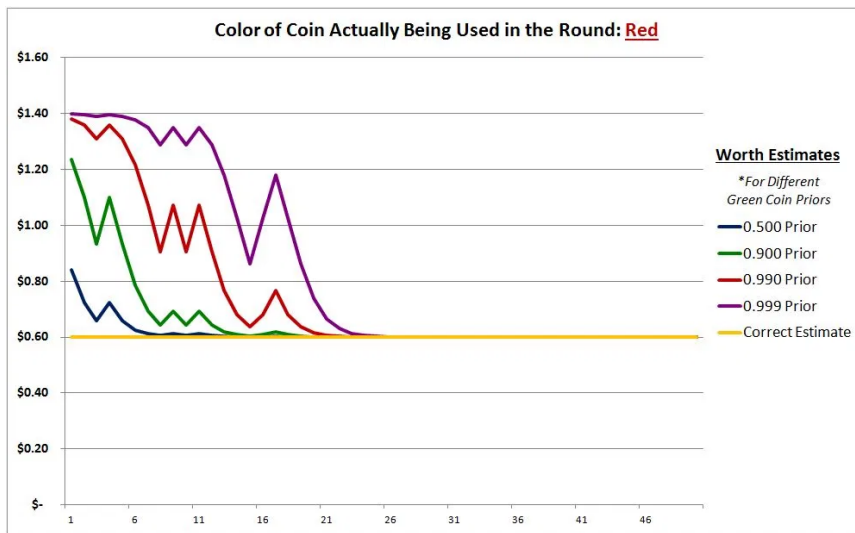
Notice that if we assign a 0.500 prior probability (blue line) to the coin being green, which is a way of expressing the fact that we have no information about the coin's likely color, and the coin ends up being red, we may still do OK in the round. That's because the updating process will efficiently bring us to the correct worth estimate, even though we'll be starting from an incorrect estimate. The process won't take long, and our worth estimates won't spend too much time at values far away from the true worth.

But if we assign higher probabilities to the coin being green—say, 0.990, or 0.999, per the above—and the coin ends up being red, our performance in the round is going to suffer. The updating process that will be needed to move us to a correct estimate will end up taking significantly longer, and we'll be significantly overpaying for each flip along the way. The reason that the updating process will take significantly longer on these more confident priors (0.990, 0.999, etc.) is that a large number of unexpected tails will have to accumulate before the ensuing result will be "unlikely" enough (on a green coin) to outweigh our strong green coin priors and sufficiently alter our stance. Each one of the tails that has to build up will come at a cost—a substantial cost, given how far off our worth estimates (and our bids) are going to be.

To see the inefficiency play out, consider the performance of the 0.999 prior, shown in the purple line above. That prior corresponds to an assigned 99.9% probability that the coin is green. Even after 10 flips, where 80% come back tails, we're still going to be assigning a very strong probability to the coin being green—93.5% to be exact. Our estimate of the worth will have hardly budged, sitting at roughly $1.35, in comparison with the actual worth of $0.60.

The next chart shows how our estimates of the worth of each flip might proceed in a round in which a green coin is used.
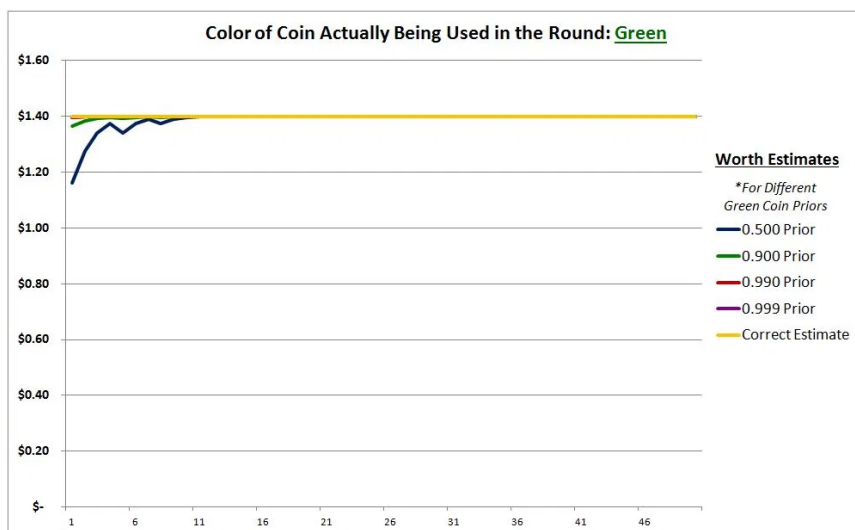
As in the previous case, the blue line, which is the worth estimate that we arrive at using a 0.500 prior (no knowledge either way), starts out at an incorrect value (technically $1.00, though the chart begins after the first update, when the estimate is roughly $1.18). Despite this incorrect starting point, the estimate quickly converges onto the right answer ($1.40) through

the updating process. We can't really see the green line, the red line, or the purple line because they essentially start out on the correct worth estimate from the get-go, at values close to $1.40. "Updating" them ends up not really being required.

The contrast between these cases highlights the asymmetric risks associated with overconfidence in the game. If we assign a very high prior probability to the coin being green–a highly aggressive number such as 0.999–and the coin ends up being red, we're going to retard the updating process and create significant losses for ourselves. At the same time, if we assign that number and the coin ends up being green, we aren't going to gain that much in effeciency or accuracy relative to what less aggressive assignments might have produced. Now, to be fair, this apparent risk asymmetry is a corollary of the fact that <u>if we are actually correct</u> in assigning a high prior probability to the coin being green, then a situation where it ends up being red <u>isn't going to happen</u> (except maybe once in a blue moon). But if it does end up happening more often than that, suggesting that we were *too confident* in our assignment, we're going to pay a heavy price for the mistake.

Now, I want to be clear. If we're <u>genuinely confident</u> that the coin is green, then we should assign a strong prior probability to it and calculate the worth of each flip accordingly. That's how we'll win the game. But we need to make sure that we have a sound basis for our confidence. If our confidence turns out to be unfounded, such that we end up assigning a high prior probability to the wrong color, it's going to be significantly more difficult, mathematically, for us to "update" back to the right answer. Our strong prior is going to effectively tie us down into believing that the coin is the color we initially thought it was, even as the incoming evidence *screams* otherwise.

Insights for the Profit Margin Debate

The primary mistake that those who were bearish on profit margins made in earlier phases of the current market cycle–and I would have to include myself in that group, at least for a time–was not the mistake of having "wrong" beliefs about the subject, but rather the mistake of assigning too much confidence to those beliefs. There wasn't a sound basis for being confident in them, first because the subject itself was inherently murky, and second because the arguments that were being used were of types that tend not to be reliable (the arguments may have been *persuasive*, but persuasive and reliable are not the same thing).

Looking specifically at the <u>theoretical</u> arguments, those who were bearish on profit margins argued that competition would eventually force a mean-reversion to occur. But what competition were they talking about? Competition where? In what sectors? Among what companies? Competition has always been around. If it represented the antidote to elevated profit margins, then why had it allowed profit margins to become elevated in the first place? If it was capable of reversing them, then why hadn't it been capable of <u>stopping</u> them from forming?

Abstract theoretical arguments such as the one presented above tend to miss important details. Granular examinations, conducted rigorously from the bottom up, are usually more reliable. If such an <u>examination</u> had been conducted in this case, it would have shown that the profit margin expansion that took place from the mid 1990s to 2011 was *not* broad-based, but was instead concentrated in select large-cap companies, most notably those in the Tech industry (think: companies like Apple, Microsoft, Google, etc). Inside specific sectors, the profit margin expansion was <u>skewed</u>, with companies in the highest tiers of profitability seeing large profit margin increases, and companies in the lower tiers seeing no increases at all, or even decreases. These are exactly the kinds of signs that we would expect to see if increased <u>monopolization</u> were taking place in the competitive landscape. Something appears to be making it easier for large best-of-breed corporations, particularly those in the Tech sector, to earn high profits without being threatened by competition. Whatever that something is (and it is likely to be *multiple* things), there was little reason to be confident, in 2011, that it was about to go away.

Looking specifically at the <u>empirical</u> arguments, those who were bearish on profit margins pointed out that every time profit margins had been at current levels in the past, they had always eventually fallen back down to the mean. But what was the sample size on that observation? Two historical instances? Three? Maybe four? A hypothesis inferred from a small sample may be worth embracing, but it should be embraced with caution, not confidence. And what about the data from the mid 1990s to 2011, data that, with the exception of brief recession-related drops in 2001 and 2008 (both of which quickly reversed themselves), had been showing a clear and persistent tendency towards profit margin elevation? This is what the chart of profit margins looked like from 2011's vantage point:
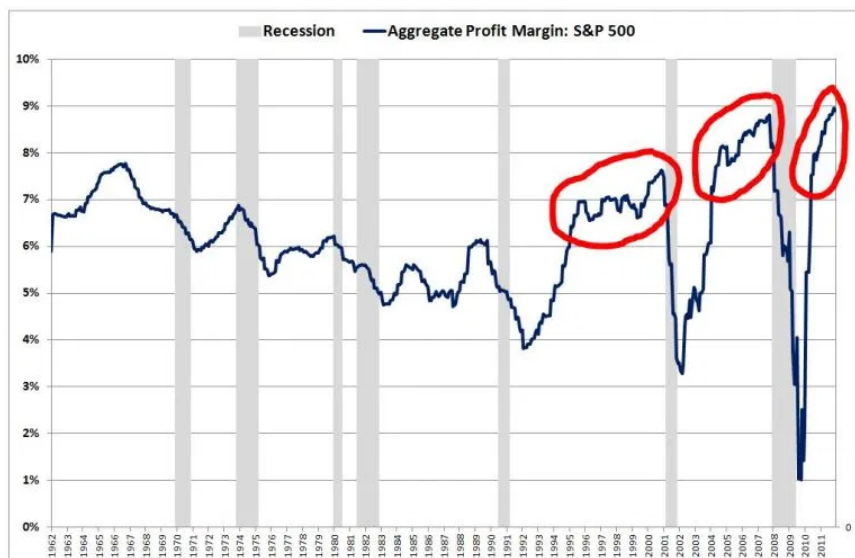
If the goal was to accurately predict what was going to happen from 2011 onward, then the data from the mid 1990s to 2011 should have been weighted more heavily than data from more distant periods of history, given that that data was obtained from a period that was temporally closer to (and therefore more likely to share commonalities with) the period of interest.

Granted, it's easy to make these points in hindsight, given that we know how the result ended up playing out. But I would nonetheless maintain that a sound evaluation of the theoretical and empirical evidence for the mean-reversion hypothesis, carried out from the perspective of what could have been known at that time, would have led to the assignment of significant uncertainty to the hypothesis, even if the hypothesis would have been retained. If that uncertainty had been appreciated, the updating process would have been completed more quickly in response to the disconfirming results that ensued, which would have allowed those investors who initially embraced the hypothesis to have participated in a greater share of the returns that the market went on to deliver.

Share this:

Tweet

This entry was posted in Uncategorized. Bookmark the permalink.

PHILOSOPHICAL ECONOMICS
Proudly powered by WordPress.