ERGODICITY AND YOU:

ADAPTIVE HEURISTICS IN AN UNCERTAIN WORLD

Justin Zovas

University of Pennsylvania

Master of Behavioral and Decision Sciences

September 2019

ABSTRACT

ERGODICITY AND YOU: ADAPTIVE HEURISTICS IN AN UNCERTAIN WORLD

Justin Zovas

Life requires making decisions under uncertainty. Facing complex, dynamic environments, decision-making processes should focus on the consequences of choices with *time* as a fundamental consideration. To that end, I recommend honing adaptive heuristics through trial and error while maintaining a margin of safety from ruin.

Contents

Russian Roulette and You.…………….. …3

Ergodicity and You.……………………….3

The Environment and You………………...5

Predictions and You…………………….....6

Machines and You.………………………...7

Don’t Die (and other useful heuristics)……8

Investing and You………………………….9

References…………………………………13

Russian Roulette and You

Imagine a revolver pointed at your head – a bullet placed randomly in one of its six chambers. The gun holder offers a proposition: if you let her pull the trigger and nothing comes out, you win ten million dollars. Would you play?

Five out of six times you win big but the other outcome looms large. A profit-maximizing robot, undeterred by ruin, accepts the offer; the expected value is massive! But for nearly all of us humans the choice is a no-brainer. We are born, time passes, and then we die. Time is scarce and irreversible – it's all we have. Time is the currency of life.

Ergodicity and You

Consider another game where players wager money on a series of coin flips. A player chooses how much money to risk, which represents his initial wealth, and then a coin is flipped in succession. If the coin lands heads, wealth increases by 30%; if it lands tails, wealth decreases by 25%. Would you play?

I created a simulation of this game where 100 players, each starting with an initial wealth of $100, flip a coin 100 times. Figure 1 shows each player's wealth changes as well as the average and median wealth after each flip. On average, a player's wealth is expected to increase, but notice how many players performed worse than the average path.

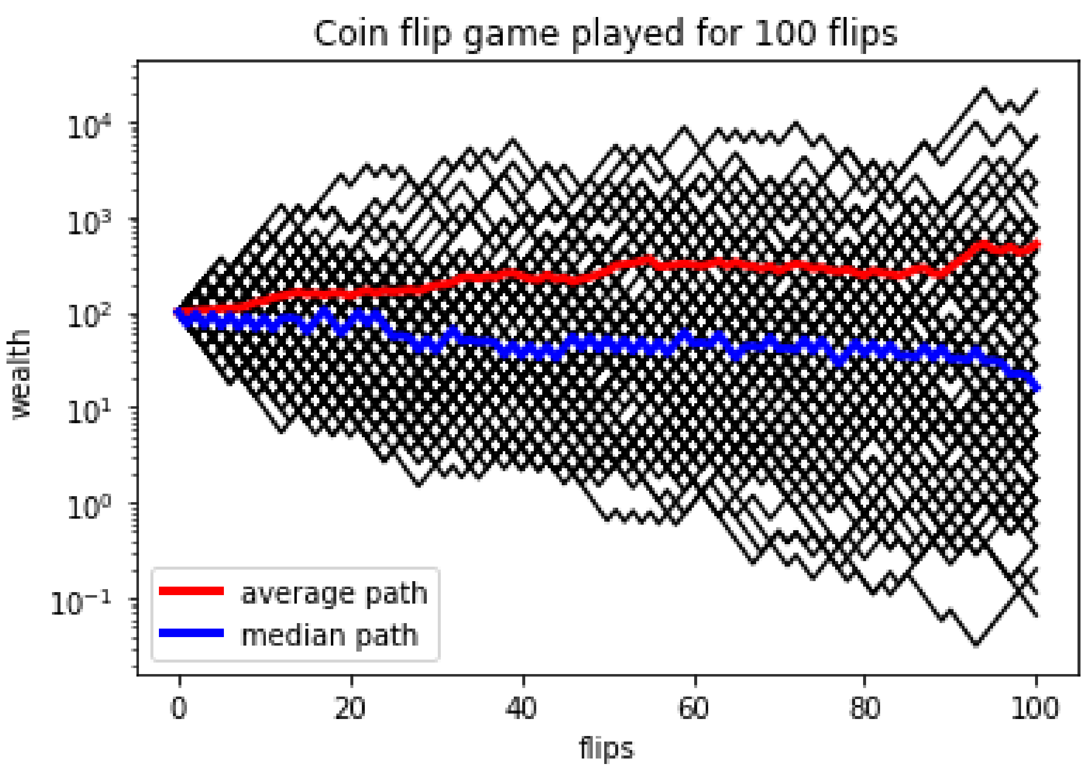

Figure 1: Coin flip game outcomes over 100 days (100 players each starting with $100)

In this simulation, the average final wealth after 100 flips is $519, yet the median final wealth is only $16. 87% of players performed worse than average and 68% of players lost money. How is this possible? Because wealth changes in percentage terms, the sequence of coin flips matters: flipping heads on the first flip and tails on the second flip is not the same as flipping tails then heads. The game's multiplicative nature creates a few players who accrue massive winnings that dwarf the majority of other players' earnings.

Russian roulette and this coin flip game are toy examples that highlight a concept called *ergodicity*. Ergodic theory recognizes that expected value – the ensemble average of the probability space – does not necessarily equal the average outcome over time[1].

An indexed sequence of random variables called a stochastic process can help evaluate ergodicity. These random processes model unique paths through time. A stochastic process is (perfectly) ergodic if its expected value equals the average outcome as time approaches infinity[1]. In a non-ergodic process like the two examples discussed, the time average differs from the expected value.

The coin flip game has a positive expected value, computed as follows:

$$Expected\ Value\ =\ wealth\ \times [\ P(heads) \times\ (1 +\ 0.3)\ + P(tails) \times\ (1 - \ 0.25)\ ]$$

$$Expected\ Value\ =\ wealth \times 1.025$$

But the game's time average expectation is negative:

$$Time\ Average = weath\ \times\ \sqrt{1.3\ \times\ .75}$$

$$Time\ Average = weath\ \times\ 0.987$$

As time passes – with each additional coin flip or trigger pull – the probability of losing all that was risked approaches one. Path dependence, where subsequent outcomes depend on prior results, creates an inequality between the ensemble average and time average. Ergodic theory emerged in the late 19th century in the field of statistical mechanics, particularly thermodynamics and the study of gas molecules, and is relevant to any dynamical system[1].

The critical flaw of expected value is its failure to account for a fundamental element of existence: time. By weighting all possible outcomes by their probability of occurrence, expected value views life as a collection of multiple simultaneous realities. In the real-world, the passage of time ensures we each experience one unique path and cannot access the other parallel universes that form expected value. Consequently, the weighted average of all possibilities may not equal individuals' average outcome over time. Time averages are relevant to individuals so relying on expected value when making decisions can be misleading.

I suspect humans innately understood ergodicity before we had a word for it.

The Environment and You

Evolution over millions of years has selected traits conducive to survival. Charles Darwin argued the environment influences human behavior through natural selection. Environmental characteristics made certain behaviors more favorable than others. Those better adapted to the environment had more babies than those less adapted[2]. The logic follows that humans who optimized time average performance were more likely to reproduce (while Russian roulette enthusiasts tended to exit the gene pool). From a Darwinian view, our choice behavior under risk reflects a disposition to prioritize time average outcomes over expected value.

Recent experimental evidence supports this claim. David Meder and his colleagues measured subjects' choice behavior by manipulating the ergodic properties of an experimental gambling environment. The researchers found subjects systematically adjusted their risk preferences to optimize time average performance[3]. These results support recent economic theory developed by Ole Peters and Murray Gell-Mann that argues decision-making should be sensitive to the particular risk dynamic[1]

The experiment compared decision-making under additive versus multiplicative gambling dynamics. Additive changes like your bi-weekly paycheck are ergodic: maximizing the expected value of additive wealth changes corresponds to maximizing time average growth. In contrast, under multiplicative dynamics like reproduction or investing, outcomes are non-ergodic due to path dependency[*]. Most of life's risks are multiplicative, notably the self-producing nature of evolution and capitalism.

Multiplicative environments require a logarithmic utility function to ensure ergodicity. In other words, when facing multiplicative risk, one's utility function should adjust so that it is more sensitive to wealth changes. The crucial implication on behavior is increased risk aversion under multiplicative dynamics (relative to additive dynamics)[1].

The experimental subjects may not know those technical details, but their behavior reflects an inherent understanding of ergodicity. The results suggested a time average optimization model best explained subjects' behavior[3]. Subjects adapted their utility functions to realize ergodicity: their choices reflected linear utility functions for additive gambles and non-linear utility functions for multiplicative gambles. In order to maximize time average wealth growth, subjects exhibited increased risk aversion when facing multiplicative risk.

These findings contradict traditional utility theory models that suggest humans are indifferent to the dynamics of the environment[1]. A behavioral model that assumes humans maximize expected value without regard to ergodicity is inaccurate. When a risky dynamic is non-ergodic, it is foolish to optimize for expected value.

---

*Multiplicative dynamics generate a non-stationary stochastic process. A random process is stationary if the variables are invariant to shifts in the time index; for a non-stationary process, the distribution of outcomes at time n may differ from the distribution at time $n + 1$[4].

Predictions and You

We seek order amidst chaos and view the world deterministically[5]. Determinism suggests outcomes can be linked to causes. In an effort to understand the past and predict the future, we seek clean, causal explanations. We prefer things happen for intelligible reasons and tend to minimize the role of chance[5,6]. To admit “I don’t know” makes you confront an uncomfortable truth whereas legibility reduces anxiety.

The problem is, the coherence of our explanatory narratives is far more persuasive than the quality of information supporting them[6]. Coherence requires logic and consistency, which problematically is function of our preexisting beliefs and (severely limited) past experiences[6]. These tendencies deceive us into thinking the world is more intelligible and predictable than it truly is. Yet, we are wired to make predictions; we must get out of bed and make uncertain choices daily. So what to do?

In domains with reliable mathematic relationships, deterministic models are helpful. But wherever human behavior plays a role, things get messy. For example, here’s a subtle distinction: scientists and engineers can develop sophisticated algorithms that excels at chess but creating an artificial intelligence that dominates a poker tournament is more difficult. There is a *human* element present in poker that makes it less tractable for a rules-based machine. As Richard Feynman famously said, “Imagine how much harder physics would be if electrons had feelings”.

An initial hurdle to making good decisions is embracing randomness and uncertainty. Thinking probabilistically helps you make informed decisions. It forces you consider multiple possible outcomes and assign likelihoods to them. You gain awareness for counterfactuals and recognize far more can happen than did happen. Research from Philip Tetlock finds human forecasters can benefit from the ability to distinguish up to twenty degrees of uncertainty, reflecting sensitivity to a 5% difference in likelihood[10]. Three degrees of uncertainty, for example, could view a future event as either likely, unlikely, or a 50/50 proposition. 20 degrees is impressive precision and provides hope that human judgement – the result of millions of years of us making predictions, genes entering and exiting the pool – can be effective in today’s environment.

How do you measure uncertainty? First, distinguish *known* unknowns from *unknown* unknowns. Known unknowns represent inherent, irreducible variability. Coin flips and casino games are true known unknowns. Epistemic uncertainty is uncertainty due to a lack of knowledge; it can be reduced over time but never eliminated[5].

We prefer known unknowns to unknown unknowns. Confidence in a variable’s probability distribution helps us avoid admitting “I don’t know” and enables statistical inference: we can measure variance, draw conclusions, and make predictions. When dealing with unknown unknowns, however, statistical tools become less effective, useless, or even harmful. Bad outcomes occur when you think you know something that you don’t – naïve beliefs.

The problem with unknown unknowns is two-fold: not only are they absent from (or not properly reflected in) the probability distribution, but disregarding their (invisible) presence is perilous. This is why we should be wary of relying too heavily on empiricism. A flaw in the evidence-based approach is absence of evidence does not provide evidence of absence[5]. Unprecedented events happen all the time. Morgan Housel writes, "The irony of history is that it's mostly the study of things changing, often used as a guide for what to do next"[7]. Or perhaps Daniel Kahneman put it bests, "that's the correct lesson to learn from surprises: that the world is surprising."[8]

It is essential to recognize the limits of your knowledge – what can versus what can't be known or measured. Here's what not to do: pretend unknown unknowns don't exist or use sophisticated math to try to compute something that's incomputable.

Machines and You

There is a well-known trade-off in modeling between bias and variance. Bias represents model error from being too simplistic; variance is error from over-fitting[5]. Today we capture and process more data than ever before. These advances in computing combined with math and ingenuity has produced novel ways to reduce bias. We are motivated to be precise because we want our models to capture reality, but at some point, removing bias creates unwanted variance.

Machines reduce bias and can do things we can't. For example, computer judgement is more consistent (less noisy) than human judgement. Yet, machines don't know everything humans know. As the end-users, we define knowledge – it comes from our perspective. Correlations are abundant but machines don't have the semantic understanding to determine which associations are meaningful[9]. A model is limited to the data it's built on. This is particularly problematic in domains that change over time.

Machine learning methods analyze vast data sets and identify patterns humans miss or take much longer to spot[9]. The challenge is, they're tough for us to audit. For example, it's difficult to understand how a neural network, with multiple layers and hundreds of nodes, transforms inputs to outputs. If the decision rules were knowable, we wouldn't need the complex model[9]. We developed techniques to gain understanding in ways we can't understand, so our fidelity to these creations should be fragile. This line of reasoning is from Benedict Evans, who summarizes the point well: "Machine learning finds patterns in data – what patterns depend on the data, and the data is up to us, and what we do with it is up to us"[9].

What is the proper trade-off between using relevant data and a sufficient sample size? Does this model pass the laugh test? How confident am I that this model, built on past data, will be useful now? Whether it's in model building and evaluation or making decisions based on outputs, the need for human judgement is unavoidable. We are more equipped than ever to answer questions; we're good at making machines do things. Now it's on us to ask better questions and carefully decide what to make them do. This is where intuition and emotions creep in, for better or for worse.

Michael Polanyi recently introduced the notion of tacit knowledge writing, “we can know more than we can tell”, perhaps inspired by the following excerpt from Osho’s *Intuition – Knowing Beyond Logic*[11, 12]

> In language it looks okay to ask, “Can intuition be explained?” But it means, “Can intuition be reduced to intellect?” And intuition is something beyond the intellect … The leap of intuition can be felt because there is a gap. Intuition can be felt by the intellect – it can be noted that something has happened – but it cannot be explained, because explanation needs causality.

Not all truths are legible. Words, definitions, equations, numbers are just proxies of reality. Enter intuition[*].

My intuitions and emotions may not always be reliable, but I use them constantly to make decisions. Emotion is a tool for learning. Feelings such as fear, joy or embarrassment provide helpful cues about our environment and feedback on our choices[13]. Research from Antonio Damasio found that patients are incapable of making decisions after having a type of brain surgery that prevents them from experiencing emotions[5]. So while emotions and intuitions may lead to bad choices, without them, we wouldn’t be able to make decisions at all. Emotions have been described as “lubricants of reason”[5].

When making uncertain decisions, I don’t know the right weights to give human judgement and objective, mathematic methods. The answer, if it exists, is domain specific so speaking in generalities doesn’t offer solutions. One approach is to develop algorithms that use data to produce base rate outputs. From there, you can tweak the output based on our judgements, intuitions, and emotions. A reverse approach has judgement as its anchor, tweaked by machines. Again, I don’t know the right balance but I’m convinced both tools are necessary.

Don’t Die (and other useful heuristics)

This essay builds on two principle beliefs: (1) we must make decisions in an uncertain world with incomplete information (2) time is our most precious resource. With time as a fundamental constraint, the ideal is to do what we want, when we want, with whom we want. To that end, I believe decision-making under uncertainty should prioritize the consequences of our actions with the paramount goal of avoiding ruin. Consequences can be mapped to future you (or your family, tribe, species, environment), which is more knowable than whatever future event you wish to predict.

---

*According to Daniel Kahneman, the reliability of intuition is function of feedback and changes to the environment[6].

I've taken this line of reasoning from Nassim Taleb who formalizes it as follows: let *X* represent what you care to predict and *F(X)* represent how that underlying variable affects you[5]. We have a greater ability to understand, adjust, control, and predict *F(X)* than *X*. For example, back to Russia roulette, whether there is a 1/6, 1/100 or 1/1000 chance the next shot will fire a bullet can be deemed irrelevant by refusing to play the game. In terms of monetary risks, how much you make when you're right and lose when you're wrong is far more important (and controllable) than the probability of being correct. We can't know the future but we can control our exposure to it.

This idea of focusing on consequences can be implemented using heuristics. Prioritizing consequences over predictions is a heuristic itself. Another heuristic is to err on the side of caution; the more uncertain you are about *X*, the less exposure you should have to it. Other simple rules involve comparing the benefits and harms associated with a decision. This requires having a rough understanding of the distribution of possible outcomes. Precision isn't necessary, but awareness for the possibility of extremely positive and negative tail events is important. Recognizing convexity and concavity – non-linear relationships between inputs and outputs – helps you detect asymmetric outcomes[5]. From there, the heuristic is intuitive: embrace opportunities where the upside is greater than the downside and avoid the opposite. By reducing or eliminating exposure to negative tail events, you are protected against your inability to predict them.

Both can be true: it is our default to be risk averse and we (both personally and at greater scales) can benefit from taking *calculated* risks. By calculated I mean conclusions driven by deliberate, effortful reasoning. Decision-time can be sped up by leveraging heuristics honed through trial-and-error. While maintaining a margin of safety, you can tinker with decision rules to find what suits you best.

Ultimately, it's your choice how you spend your time. Avoiding ruin is our default state, but from there our preferences diverge. An incredibly fulfilling experience to one person is viewed as a crazy risk by another. We may value time differently but it is our liberty to value time that is universally special. What we do, desire, and aspire to be gives meaning to our lives. Worry about what you can control and be careful exposing yourself to what you can't control. It's up to you.

Investing and You

This final section applies the ideas presented in this essay to investing. Investors must deal with both non-ergodicity and epistemic uncertainty. Non-ergodicity arises because price returns are multiplicative. After 50% price decline, a 100% price increase is required to get back to even; a 20% increase followed by a 20% decrease leaves you down 4%. And unlike the simple coin flip game, the probability distribution for future price returns is dynamic and complex. While historical price changes are best estimated using a power-law distribution, the future distribution is unknowable[*].

Going back to 1928, the S&P 500 has provided an average annualized return of 8% (ignoring dividends)[14]. Remember that averages of non-ergodic processes are misleading because they don't reflect individuals' experiences. No one realized that 8% average. An investor who bought in 1928 and held earned the geometric return of 6% annualized.

Other examples of information that lacks relevance comes from sexy headlines like, "If you invested $1,000 in Microsoft 25 years ago, you would have $45,000 today – a return of 16% per year!" Hypotheticals inform theory but distort reality. That flashy anecdote ignores the stock's volatile path during those 25 years. Long-term returns conveniently leave out the emotional parts of investing. For a Microsoft stock holder, a headline number doesn't tell you what it feels like to experience a 75% drawdown in 2009 or see the price flat-line for a decade while other stocks soar[14].

Investing will often be disappointing if you form short-term expectations based on long-term results. Sequence risk and path dependence causes individual investors' returns to detach from the long-run average. The market doesn't know or care about your personal ergodic needs. For this reason, its critical investors are aware of how prices behave in the short-term.

Benoit Mandelbrot discovered price changes of risky assets like stocks, bonds, and commodities are discontinuous and interdependent[15,16]. Unlike the temperature outside, a stock can jump from $70 per share to $75 in an instant. Interdependence means markets have a memory: what has happened influences what will happen[15,16]. Related, he also found that volatility tends to cluster; volatility begets volatility. For example, the 10 worst daily price declines in the S&P 500 over the past 50 years occurred within two periods: the 3 worst days happened in October 1987 and the next 7 worst within a 63 day window in late 2008. These characteristics help explain why markets have tail risks and roughly follow a power-law distribution[†].

The historical relationship between the frequency and magnitude of price changes can be estimated. This scaling parameter, however, is highly sensitive to the most extreme events, which occur least often, and is subject to perpetual revision. One data point – the 20% price drop that occurred on October 19th 1987 – has a disproportionate influence on the estimated power-law[17]. What's more, it would be naive to conclude that October 19th, 1987 represents the most stocks can lose in a single day. If you don't believe me, consider this (another insight from Taleb): the worst event ever, when it happened, exceeded the worst event at the time. In the

---

[*] To be exact, the distribution of stock returns is fat tailed and the tails follow a power-law.

[†] A power-law distribution is characterized by a non-linear, proportional relationship between the frequency and magnitude of a random variable. In simpler terms, this means most price changes are small and a small number of price changes are very large. As a result, deviations from the center of the distribution are far from "standard".

context of investing, the key implication is massive price swings are rare and fundamentally alter the statistical properties of the historical data. Exponentially more impactful price movements are exponentially more difficult to predict. What impacts us most, we know about least. As a result, the *true* mean and variance of returns can never be known[17].

To recap, in the relative short-term (daily, weekly or monthly periods), price changes are interdependent, discontinuous and follow a power-law. A non-ergodic process with these attributes means investors' unique paths have significant downside risk, at least in the short-run. Counteracting this risk requires choosing a proper time horizon and exposure: deliberately investing an appropriate amount of money for a predetermined amount of time.

Start by answering: why am I investing? Mapping the motives for investing to future needs and desires gives your decisions purpose. Amidst short-term volatility, your purpose should drive your decisions opposed to fleeting emotions that may not align with your goals. If you are investing in stocks to help pay for your kid's college tuition in 20 years, a big weekly price decline shouldn't compel you to sell. Investing for retirement using IRAs can be effective beyond just enforcing a long time horizon because the tax benefits and penalties for early withdrawals act as a behavioral commitment tool. The structure incentivizes investors to continuously contribute and avoiding selling over multiple decades.

More generally, there *may* be benefits to investing with a longer time horizon. I emphasize *may* because it's important not to mistake absence of evidence for evidence of absence. The evidence suggests for yearly (or longer) price changes, negative tail events are both less frequent and less extreme. The stock market has gone up over time. The challenge with measuring increasingly longer time periods, however, is data scarcity; longer periods have less non-overlapping occurrences so the sample size is much smaller. Regardless, investing with purpose and choosing a future date when you plan to sell is important.

Position sizing is the other critical component to successful investing. A carefully chosen time horizon becomes meaningless if you are forced to prematurely sell your investment. For example, money needed to pay recurring and unexpected expenses should not be invested in stocks. Assuming you hold a sufficient amount of cash that prevents you from having to sell, the next consideration is your risk appetite. What would you do if your investments were in a 30% drawdown? If your answer is to sell, you are likely risking too much money. If weekly volatility elicits visceral emotions, you are also probably over-exposed.

John Kelly created a neat formula to make our money gambles ergodic. Resting on two critical inputs – the probability of winning and payoff of the investment – the Kelly criterion recommends risking a percentage amount that maximizes long-run growth[18]. By construction, the formula prescribes risking only a fraction of one's wealth. Because bet sizes change in response to wealth changes, following the formula keeps you in the game by making it impossible to run out of money.

Of course, in most situations the probabilities and payoffs can't be known with certainty. It's difficult to follow a formula's output that rests so heavily on inputs soaked in epistemic uncertainty. For this reason, practitioners typically only risk a fraction of what Kelly recommends.

Ultimately, there's no formula to determine the right amount to risk; the answer is deeply personal. The Kelly criterion provides a percentage and it's up to you to determine your bankroll size and whether you are comfortable with that percentage. The formula cannot ensure that its user will maintain consistent behavior during a drawdown or after multiple losses in a row. Your personal answer emerges through trial-and-error.

Investing is inextricably emotional; the role psychology plays can only be understood through direct experience. How you *think* you will react to a 30% drawdown can be much different than your actual reaction. Tinker with different amounts of exposure and over time you will develop a sense for the amount you're comfortable with. The key is to start with small exposures; invest an amount you are confident you won't be compelled to sell before your time horizon is up. From there, you can consider slightly increasing your exposure. Risk taking is only worth it if you're able and willing to make the next bet.

References

1. Peters, Ole, and Murray Gell-Mann. "Evaluating gambles using dynamics." *Chaos: An Interdisciplinary Journal of Nonlinear Science* 26.2 (2016): 023103.
2. Todd, Peter M., and Gerd Gigerenzer. "Environments that make us smart: Ecological rationality." *Current directions in psychological science* 16.3 (2007): 167-171.
3. Meder, David, et al. "Ergodicity-breaking reveals time optimal economic behavior in humans." *arXiv preprint arXiv:1906.04652* (2019).
4. Cover, Thomas M., and Joy A. Thomas. *Elements of information theory*. John Wiley & Sons, 2012.
5. Taleb, Nassim Nicholas. "Incerto: Skin in the Game, Antifragile, The Black Swan, Fooled by Randomness, the Bed of Procrustes." (2018).
6. Kahneman, Daniel. *Thinking, fast and slow*. Macmillan, 2011.
7. Housel, Morgan. "The Thin Line Between Bold and Reckless." Collaborative Fund, 25 Jan. 2018, www.collaborativefund.com/blog/the-thin-line-between-bold-and-reckless/.
8. Housel, Morgan. "The Psychology of Prediction." Collaborative Fund, 21 Jul. 2019, www.collaborativefund.com/blog/ the-psychology-of-prediction/.
9. Evans, Benedict. "Notes on AI Bias." *Benedict Evans*, 15 Apr. 2019, www.ben-evans.com/benedictevans/2019/4/15/notes-on-ai-bias.
10. Tetlock, Philip E., and Dan Gardner. *Superforecasting: The art and science of prediction*. Random House, 2016.
11. Polanyi, Michael. *The tacit dimension*. University of Chicago press, 2009.
12. Osho. *Intuition: Knowing beyond logic*. St. Martin's Press, 2001.
13. Lo, Andrew W. *Adaptive markets: Financial evolution at the speed of thought*. Princeton University Press, 2019.
14. Bloomberg, L.P. "Daily prices for SPX Index and MSFT US Equity" Bloomberg database. University of Pennsylvania, Lippincott Library**.** Jul. 15, 2019
15. Hudson, Richard L., and Benoit B. Mandelbrot. *The Misbehavior of Markets: A Fractal View of Risk, Ruin, and Reward*. Profile, 2008.
16. Mandelbrot, Benoit B. *Fractals and scaling in finance: Discontinuity, concentration, risk. Selecta volume E*. Springer Science & Business Media, 2013.
17. Taleb, Nassim Nicholas. "The Statistical Consequences of Fat Tails." (2018).
18. Brown, Aaron*. Red-blooded risk: the secret history of Wall Street.* John Wiley & Sons*, 2011.*